

NanoScience and Technology

Hilmi Ünlü

Norman J.M. Horing

Jaroslav Dabowski *Editors*

Low-Dimensional and Nanostructured Materials and Devices

Properties, Synthesis, Characterization,
Modelling and Applications

NanoScience and Technology

Series editors

Phaedon Avouris, Yorktown Heights, USA

Bharat Bhushan, Columbus, USA

Dieter Bimberg, Berlin, Germany

Klaus von Klitzing, Stuttgart, Germany

Cun-Zheng Ning, Tempe, USA

Roland Wiesendanger, Hamburg, Germany

The series NanoScience and Technology is focused on the fascinating nano-world, mesoscopic physics, analysis with atomic resolution, nano and quantum-effect devices, nanomechanics and atomic-scale processes. All the basic aspects and technology-oriented developments in this emerging discipline are covered by comprehensive and timely books. The series constitutes a survey of the relevant special topics, which are presented by leading experts in the field. These books will appeal to researchers, engineers, and advanced students.

More information about this series at <http://www.springer.com/series/3705>

Hilmi Ünlü · Norman J.M. Horing
Jaroslaw Dabowski
Editors

Low-Dimensional and Nanostructured Materials and Devices

Properties, Synthesis, Characterization,
Modelling and Applications

Editors

Hilmi Ünlü

Department of Physics Engineering, Faculty
of Science and Letters

Istanbul Technical University

Istanbul

Turkey

Jaroslav Dabowski

Innovations for High Performance

Microelectronics (IHP) GmbH

Frankfurt

Germany

Norman J.M. Horing

Department of Physics and Engineering

Physics

Stevens Institute of Technology

Hoboken, NJ

USA

ISSN 1434-4904

NanoScience and Technology

ISBN 978-3-319-25338-1

DOI 10.1007/978-3-319-25340-4

ISSN 2197-7127 (electronic)

ISBN 978-3-319-25340-4 (eBook)

Library of Congress Control Number: 2015953262

Springer Cham Heidelberg New York Dordrecht London

© Springer International Publishing Switzerland 2016

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media
(www.springer.com)

*My work in this volume is dedicated to my
wife, Yeliz and my children, Melek,
Merve and Miray, in deep appreciation for
their support and encouragements.*

—Hilmi Ünlü

*My work in this volume is dedicated to my
wife, Jerilyn, in deep appreciation
for her enduring forbearance and love.*

—Norman J.M. Horing

Preface

Since the invention of the first transistor in 1949, there has been a continuous avalanche of growth and development in the science and technology of semiconductor materials and device applications. The thrust of the advances has involved the steady lowering of the dimensions in which confined charge carriers can move, ultimately reaching the nano-world. The term *nano*, a prefix derived from the Greek word for dwarf (νᾶνος), is used to indicate sizes in the range of about one to a hundred nanometers, or 10^{-9} to 10^{-7} m. By 1970, systems thin enough to be regarded as two dimensional (for example, semiconductor inversion layers, etc.) were being analyzed, and these were rapidly followed by one-dimensional quantum wires and zero-dimensional quantum dots. In conjunction with technology, “nano” was used for the first time by Norio Taniguchi in 1974. As he predicted, nanoscience and nanotechnology emerged strongly as research fields in the 1980s: the scanning tunneling microscope was invented by Gerd Binnig and Heinrich Rohrer in 1981 and the first carbon nanomaterial, C60 buckyballs, was discovered by Richard Smalley, Robert Curl, Harold Kroto, James Heath, and Sean O'Brien in 1985, followed by carbon nanotubes in 1991 found by Sumio Iijima. With the discovery of graphene in 2004 by Andre Geim and Kostya Novoselov, the world was presented with a two-dimensional device-friendly carbon material just a single atom thick, and this was followed by similar materials including silicene, germanene, stanene, metal dichalcogenides, and topological insulators.

The requisite support basis of nanotechnology, laden with promising developments tantamount to a new industrial revolution, has been a qualitatively reliable understanding of the underlying physics of semiconducting materials accompanied by quantitatively precise predictions of device performance. This has led to new concepts and techniques of semiconductor growth that have facilitated the emergence of a generation of advanced devices with more complex functionality and much higher densities for electronic, computational, and optical applications. Advances in the growth of semiconductor thin films of differing structural, electronic, and optical properties, and the diminution of layer thickness approaching atomic dimensions, have provided new opportunities for fundamental scientific

studies and technological applications of semiconductors in new devices. Moreover, contemporary fabrication technologies have made it possible to reduce device dimensions to the point where size effects must be properly described quantum mechanically in order to reliably predict the potential and performance of low-dimensional semiconductor systems for electronic and optical applications.

Advances in single and multilayer thin film growth technologies (e.g., molecular beam epitaxy (MBE), and metal organic chemical vapor deposition (MOCVD)), have made it possible to produce semiconductor heterostructure thin films based on group IV–IV, III–V and II–VI semiconductors in binary/binary and alloy/binary forms, having engineered electronic and optical properties that are not available in nature. One can now control the alloy composition and doping in ternary, quaternary and pentanary IV–IV, III–V and II–VI semiconductors over atomic distances. These growth techniques, taken jointly with advanced characterization and fabrication techniques have facilitated the development of a number of high performance devices for fast signal processing, as well as some novel structures that are of fundamental interest to solid state scientists and device engineers. The epitaxial layers are so thin that quantum mechanical effects govern the operation of such heterostructure devices. These materials, coupled with device design, have enabled the emergence of novel devices in which signals propagate faster, promising to replace silicon with compound semiconductors having much higher electron mobility and velocity. Low-dimensional bipolar and unipolar compound semiconductor devices (e.g., GaAs based heterojunction bipolar transistors (HBTs), modulation doped field effect transistors (MODFETs), light-emitting diodes, lasers, etc.) operate much faster than conventional homostructure silicon devices (e.g., bipolar junction transistors (BJTs), metal oxide field effect transistors (MOSFETs), etc.), leading to a many-thousand-fold increase in speed, which is of crucial importance to the electronic and optical communication and computer industries.

This book describes recent scientific and technological developments of low-dimensional nanomaterials in over 20 review and research articles. It begins with metrology and methodology, with simple carbon nanomaterials (nanotubes and graphene), and with metal oxide thin films, nanowires, and quantum dots. More complex issues associated with the environment and with energy production and storage follow. Furthermore, important achievements in materials pertinent to the fields of biology and medicine are also reviewed, exhibiting an outstanding confluence of basic physical science and vital human endeavor. Finally, after a brief excursion into device physics, new and interesting developments relating to quantum computing are addressed.

In recognition of the vastly important role that low-dimensional semiconductor systems are now playing in all fields of science and technology and the even greater role that they are poised to play in high-functionality devices for electronic, computational, optical, biological, and medical systems that will support much greater economic development and human well-being, much of this book has been prepared in a way that permits access to the various subjects by readers who are not

experts at the outset. This is to say that this book will be helpful to graduate students and young scientists who want to develop an understanding of the subject; and it will also be informative to seasoned scientists and engineers who are knowledgeable in some areas and wish to broaden their perspective in others in this multidisciplinary field.

Istanbul
Hoboken
Frankfurt

Hilmi Ünlü
Norman J.M. Horing
Jaroslaw Dabowski

Contents

1	Modelling of Heterostructures for Low Dimensional Devices.	1
	H. Hakan Gürel, Özden Akıncı and Hilmi Ünlü	
1.1	Introduction	2
1.2	Issues in Modelling of Electronic Structure in Heterostructures	5
1.2.1	Interface Strain Effects in Heterostructures.	6
1.2.2	Composition Effects in Heterostructures.	10
1.3	Semiempirical Tight Binding Modeling of Heterostructures . . .	12
1.3.1	Semiempirical Sp^3 Tight Binding Modeling	13
1.3.2	Semiempirical Sp^3s^* Tight Binding Modeling	21
1.3.3	Semiempirical $Sp^3d^5s^*$ Tight Binding Modeling	27
1.3.4	Semiempirical Sp^3d^5 Tight Binding Modeling	30
1.4	Density Functional Theory Modelling of Heterostructures	32
1.5	Modeling of Band Offsets in Heterostructures	40
1.6	Conclusion.	45
	References.	46
2	Aspects of the Modeling of Low Dimensional Quantum Systems.	49
	Norman J.M. Horing	
2.1	Introduction	49
2.2	3D Harmonic Oscillator with a Superposed Quantum Dot (Fig. 2.1).	51
2.3	Double Quantum Dot Systems in Low Dimensions.	53
2.4	Quantum Wire on a 2D Sheet in a Uniform Magnetic Field. . .	56
2.5	Inverse Dielectric Function of a Periodic Superlattice	59
2.6	An Impenetrable Barrier: Quantum Particle Dynamics in a Highly Singular 1D-Potential.	64
	References.	70

3	Wave Propagation and Diffraction Through a Subwavelength Nano-Hole in a 2D Plasmonic Screen	73
	Norman J.M. Horing and Désiré Miesse	
3.1	Introduction	73
3.2	Scalar Green's Function for a 2D Plasmonic Layer with a Nano-Hole Aperture Embedded in a 3D Bulk Host Medium	74
3.3	Scalar Field Response of a Perforated 2D Plasmonic Layer to an Incident Wave $U_0(\mathbf{r}, \omega)$: Transmission	79
3.4	Numerical Analysis of Helmholtz Scalar Wave Propagation in the Vicinity of a Perforated Plasmonic Layer with a Nano-Hole	86
3.5	Summary	103
	References.	104
4	The Challenge to Develop Metrology at the Nanoscale	105
	R. Ince	
4.1	Metrology	105
4.2	Nanotechnology Generations, Definitions, and Visions	108
4.3	Nanotechnology Research Drives	109
4.3.1	The Semiconductor Industry.	109
4.3.2	The Healthcare Industry.	110
4.4	Nanotechnology Artefacts	111
4.5	Nanomaterials Measurement Challenges.	112
4.6	Nanomaterials Instruments	113
4.6.1	Spectroscopic Ellipsometry (SE).	114
4.6.2	Metrological AFM	115
4.6.3	Scanning Electron Microscopy (SEM).	117
4.7	Other Methods for Dimensional Nanometrology	118
4.7.1	X-ray Interferometry (XRI)	119
4.7.2	Small Angle X-ray Scattering Diffractometer (SAXS)	120
4.7.3	Electron/X-ray Diffraction	121
4.7.4	Raman Spectroscopy as a Nanometrology Tool	121
4.8	The Challenge for Nanometrology	122
4.9	Conclusion.	123
	References.	124
5	Terahertz Devices and Systems for the Spectroscopic Analysis of Biomolecules—"Complexity Great and Small"	131
	Anthony John Vickers and David Crompton	
5.1	Introduction	131
5.2	THz Time Domain Spectroscopy	132
5.3	Small Complexity—Simple Polar Liquids	134

5.4	Great Complexity—Proteins	138
5.5	Conclusions	147
	References.	147
6	Recent Progress in XAFS Study for Semiconducting Thin Films	149
	Takafumi Miyanaga and Takashi Azuhata	
6.1	Introduction	149
6.2	XAFS Spectroscopy: Basic Theory and Measurements	150
6.2.1	Basic Theory of XAFS	150
6.2.2	Measurements of XAFS.	154
6.3	Application to Semiconducting Thin Films.	155
6.3.1	<i>c</i> -Plane InGaN SQWs	155
6.3.2	<i>c</i> -Plane InGaN MQW	158
6.3.3	<i>m</i> -Plane InGaN Thin Film	159
6.3.4	<i>m</i> -Plane AlGaIn Thin Films	161
6.3.5	<i>c</i> -Plane MgZnO Thin Film.	164
6.4	Summary	168
	References.	168
7	Pulsed-Laser Generation of Nanostructures	171
	Selcuk Akturk	
7.1	Introduction	171
7.2	Self-formation of Nanostructures Through Pulsed-Laser Ablation	172
7.3	Formation of Nanoscale Patterns by Femtosecond Laser Ablation	173
7.3.1	General Principles.	173
7.3.2	Femtosecond Laser Nanomachining on Thin Films with Bessel Beams	175
7.3.3	Experimental Results.	177
7.4	Conclusions	178
	References.	179
8	Graphene for Silicon Microelectronics: Ab Initio Modeling of Graphene Nucleation and Growth	181
	Jarek Dabrowski, Gunther Lippert and Grzegorz Lupina	
8.1	Introduction	181
8.2	Approach.	183
8.3	Carbon on Graphene	184
8.4	Silicon on Graphene	186
8.5	Carbon on h-BN.	188
8.6	Carbon on Mica	190
8.7	Graphene Base Transistor	194

8.8	Carbon on Germanium	196
8.9	Summary and Conclusions.	201
	References.	202
9	Recent Progress on Nonlocal Graphene/Surface Plasmons	205
	Norman J.M. Horing, A. Iurov, G. Gumbs, A. Politano and G. Chiarello	
9.1	Introduction	206
9.2	Nonlocal Dielectric Response of a Slab Plasma	206
9.3	2D Plasma Coulomb-Coupled with a Slab Plasma.	214
9.4	Numerical Results for Plasmon Dispersion: Graphene Layers Interacting with Semi-infinite Conductor	217
9.5	Experimental Studies on Epitaxial Graphene.	221
9.6	Influence of Adsorbed and Intercalated Atoms	227
9.7	Concluding Remarks	229
	Appendix 1: Dynamic Nonlocal Polarization Function for Free-Standing Graphene with no Bandgap; Brief Summary of the Results Derived in [19, 20]	229
	Appendix 2: Dynamic Nonlocal Polarization Function for Graphene with a Finite Energy Bandgap; Brief Summary of the Results Derived in [18]	231
	References.	234
10	Semiconducting Carbon Nanotubes: Properties, Characterization and Selected Applications	239
	Chiara Pintossi and Luigi Sangaletti	
10.1	Introduction to Carbon Nanotubes.	239
10.2	CNTs Synthesis	244
10.3	Carbon Nanotubes Applications	246
	10.3.1 Selected Applications of Semiconducting CNTs	246
	10.3.2 CNTs for Photovoltaic Applications	247
	10.3.3 CNT Interaction with Gases: From Surface Chemistry to Devices	250
	References.	257
11	Effects of Charging and Perpendicular Electric Field on Graphene Oxide.	261
	H. Hakan Gürel, M. Topsakal and S. Ciraci	
11.1	Introduction	262
11.2	Methodolgy	262
11.3	Theoretical Investigations of Charged Nanosystems.	263
11.4	Interaction of H ₂ O, OH, O and H with Graphene	267
	11.4.1 Binding of H ₂ O to Graphene	267
	11.4.2 Binding of OH to Graphene	267

11.4.3	Binding of O to Graphene	269
11.4.4	Binding of H and H ₂ to Graphene.	269
11.5	Effects of an Electric Field and Charging.	270
11.5.1	Effects of an Electric Field and Charging on Adsorbed O.	270
11.5.2	Effects of an Electric Field and Charging on Adsorbed OH	275
11.6	Desorption of Oxygen from GOX.	278
11.6.1	Formation of Oxygen Molecule	278
11.6.2	Interaction Between Adsorbed H and O.	280
11.6.3	Interaction Between Adsorbed H and OH	282
11.6.4	Interaction Between Two OH Co-Adsorbed in Close Proximity	284
11.7	Conclusion.	286
	References.	288
12	Structural and Optical Properties of Tungsten Oxide Based Thin Films and Nanofibers	291
	E.O. Zayim and A. Tabatabaei Mohseni	
12.1	Introduction	291
12.1.1	Amorphous and Crystalline Tungsten Oxide Based Nanomaterials	292
12.2	Tungsten Oxide Based Nanomaterials	295
12.2.1	Tungsten Oxide Based Thin Films and Mesoporous Thin Films.	296
12.2.2	Tungsten Oxide Based Nanofibers and Nanowires	298
12.3	Chromogenic Properties and Applications of Tungsten Oxide	303
12.3.1	Electrochromic Properties of Tungsten Oxide Films.	303
12.3.2	Coloration Phenomena in WO ₃	305
12.4	Conclusions	305
	References.	305
13	Electron Accumulation in InN Thin Films and Nanowires	309
	L. Colakerol Arslan and K.E. Smith	
13.1	Introduction	309
13.2	Fermi Level Pinning	311
13.3	Surface Electron Accumulation in InN Thin Films	312
13.4	Surface Charge Accumulation on InN NWs	317
13.5	Control of Surface Electron Accumulation in InN Nanowires	320
13.6	Conclusions	323
	References.	324

14	Optical and Structural Properties of Quantum Dots	327
	M.H. Yükseliçi, A. Aşıkoğlu Bozkurt, Ç. Allahverdi, Z. Nassar, D. Bulut, B. Can Ömür, M.K. Torun and A.T. İnce	
14.1	Introduction	328
14.2	CdSe _x S _{1-x} Nanocrystals.	328
14.3	Investigation of Raman Spectroscopy for CdTe Thin Film	329
14.3.1	Experimental Details	329
14.3.2	Modelling	330
14.3.3	Discussion	330
14.3.4	Importance of the Subject	331
14.3.5	Section Summary	331
14.4	Steady State Photoluminescence Spectroscopy	332
14.5	The Progression of Strain and Micro-electric Field Dependent Urbach Energy with Deposition Time in Chemical Bath Deposited CdS Thin Films	335
14.5.1	Experimental	335
14.5.2	Modeling of the Urbach Tail	341
14.5.3	The Progression of Strain with Deposition Time.	343
14.5.4	Section Conclusion	345
14.6	In Situ Low Temperature Optical Absorption Spectroscopy . . .	346
	References.	347
15	One-Dimensional Nano-structured Solar Cells	351
	H. Karaağaç, E. Peksu, E.U. Arici and M. Saif Islam	
15.1	Introduction	352
15.2	One-Dimensional Nanostructures Based Solar Cell Architectures	352
15.3	Synthesis of One-Dimensional Nanostructures	355
15.3.1	Solution-Based Synthesis of 1-D Nanostructures.	356
15.3.2	Vapor Phase-Based Synthesis of 1-D Nanostructures	359
15.4	Common Materials for 1-D Nano-Structured Solar Cells	361
15.4.1	Silicon.	362
15.4.2	Zinc-Oxide.	371
15.4.3	Titanium-Dioxide	376
15.4.4	Carbon	380
15.5	Summary and Future Outlook.	389
	References.	391
16	Computational Studies of Bismuth-Doped Zinc Oxide Nanowires	401
	Çetin Kılıç, Mehmet Aras and Sümeyra Güler-Kılıç	
16.1	Introduction	401

16.2	Computational Modeling	403
16.2.1	Supercell Approach.	403
16.2.2	Defect Calculations	404
16.2.3	Density- and Hybrid-Functional+ <i>U</i> Calculations	408
16.2.4	Computational Settings	411
16.3	Structure and Energetics of ZnO Nanowires	412
16.4	Defect Energetics and Transition Levels in ZnO:Bi Nanowire	414
	References.	419
17	Mixed-Phase TiO₂ Nanomaterials as Efficient Photocatalysts	423
	Juying Lei, Hong Li, Jinlong Zhang and Masakazu Anpo	
17.1	Introduction	424
17.2	Phases of TiO ₂	425
17.2.1	Structure Properties of Rutile, Anatase and Brookite	425
17.2.2	Stability and Phase Transformation	426
17.2.3	Photocatalytic Activity of Rutile, Anatase and Brookite	427
17.3	Synthesis of Mixed-Phase TiO ₂ Photocatalysts	429
17.3.1	Hydrothermal Method and Solvothermal Method	429
17.3.2	Microemulsion-mediated Solvothermal Method.	434
17.3.3	Sol-Gel Method	436
17.3.4	Solvent Mixing and Calcination Method	437
17.3.5	High-Temperature Calcination Method	438
17.4	Applications of Mixed-Phase TiO ₂ in Photocatalysis	441
17.4.1	Photocatalytic Hydrogen Production	441
17.4.2	Photocatalytic Reduction of CO ₂ with Water on Mixed-Phase TiO ₂	444
17.4.3	Photocatalytic Degradation of Organic Pollutants on Mixed-Phase TiO ₂	446
17.5	Mechanism of the Enhanced Photocatalytic Activities by the Mixed-Phase TiO ₂ Photocatalysis	451
17.6	Conclusion and Outlook	456
	References.	457
18	Electrochemical Impedance Study on Poly(Alkylenedioxy)Thiophene Nanostructures: Solvent and Potential Effect	461
	A. Sezai Sarac and Asli Gencturk	
18.1	Introduction	461

18.2	Experimental Details	464
18.2.1	Chemicals	464
18.2.2	Preparation of Carbon Fiber Microelectrode (CFME).	464
18.2.3	Apparatus and Procedure	464
18.3	Results and Discussion	465
18.3.1	Electropolymerization of ProDOT-Me ₂ on CFME.	465
18.3.2	FTIR-ATR Characterisation of PProDOT-Me ₂ Film on CFME.	466
18.3.3	Electrochemical Impedance Spectroscopy.	467
18.3.4	Potential Effect on EIS of PProDOT-Me ₂ /CFME Coated Electrode	468
18.3.5	Electrolyte and Solvent Effects on ProDOT-Me ₂ Coated CFMEs.	469
18.3.6	Electrical Equivalent Circuit Modeling	472
18.3.7	Morphology of Nanoporous and Compact Conductive Polymer Coatings on Carbon Fiber.	474
18.4	Conclusion.	474
	References.	474
19	Application of Nanoporous Zeolites for the Removal of Ammonium from Wastewaters: A Review	477
	Mustafa Turan	
19.1	Introduction	477
19.2	Methodology	480
19.2.1	Batch Adsorption Studies.	480
19.2.2	Column Adsorption Studies	482
19.2.3	Modification of Zeolite	484
19.2.4	Synthesized Zeolites	485
19.2.5	Regeneration	487
19.3	Nanoporous Zeolites for Ammonium Removal from Wastewaters	488
19.3.1	Ammonium Adsorption Capacities and Other Parameters for Zeolites in the Batch Systems.	488
19.3.2	Ammonium Removal from Wastewaters in the Batch and Column Systems.	495
19.3.3	Ammonium Removal from Wastewaters in the Fixed Bed Systems	497
19.3.4	Ammonium Removal for Wastewater Treatment Systems Combined with Zeolites	498
19.4	Conclusions	501
	References.	501

20 Synthesis and Biological Applications of Quantum Dots	505
Manolya Kukut Hatipoglu, Seda Kelestemur and Mustafa Culha	
20.1 Introduction	505
20.2 Synthesis of CdSe and CdSe/ZnS QDs	506
20.3 Design of CdSe and CdSe/ZnS QDs for Biological Applications	508
20.3.1 Ligand Exchange Process	509
20.3.2 Silanization Process	509
20.3.3 Amphiphilic Molecules	510
20.3.4 PEG and Phospholipid Micelle	511
20.3.5 Biomacromolecules	512
20.4 Biological Applications of QDs	514
20.4.1 In Vitro Targeting with Antibody Conjugation	515
20.4.2 In Vitro Targeting with Peptide Conjugation	517
20.4.3 In Vitro Targeting with Small Molecule Conjugation	518
20.4.4 In Vivo Applications of QDs	520
20.4.5 In Vivo Vascular Imaging	522
20.4.6 Förster Resonance Energy Transfer (FRET) Based Applications	523
20.5 Conclusions	525
References	526
 21 Bionanotechnology: Lessons from Nature for Better Material Properties	 535
F.N. Kök	
21.1 Introduction	535
21.2 Biomineralization	536
21.3 Biomimetic Proteins: Receptors, Catalysts, Channels	538
21.4 Optics/Biophotonics	539
21.5 Natural Adhesives	540
21.6 Biointerfaces	541
21.6.1 Self-cleaning Surfaces	542
21.6.2 Bioinspired Interfaces for Better Biocompatibility	543
21.7 Biomimetic Membranes	544
21.7.1 Cell Membrane Mimics	544
21.7.2 Membranes for Water Treatment	546
21.8 Hints from Nature for Endurance	546
21.9 Conclusions	548
References	549

22	Quantum Dots in Bionanotechnology and Medical Sciences: Power of the Small	555
	İ. Ergal and A.T. Akarsubasi	
22.1	Introduction	555
22.2	Types and Characteristics of QDs	556
22.3	Advantages and Disadvantages of QDs	557
22.4	Synthesis of QDs	558
22.5	Toxicity	559
22.6	Surface Modification and Functionalization	560
22.6.1	Surface Coatings to Minimize Hydrodynamic Size	561
22.7	Biocompatibility in QDs (Bioconjugation)	562
22.8	Next Generation QDs (Silica, Carbon, Metal Nanocluster)	564
22.8.1	Metal Nanoclusters	564
22.8.2	Carbon Dots (C-Dots/GQDs)	564
22.8.3	Silicon Dots (Si QDs)	565
22.9	Applications of QDs in Bionanotechnology and Nanomedicine	565
22.9.1	Cell Labeling	565
22.9.2	In Vitro Imaging	566
22.9.3	In Vivo Imaging	568
22.10	Dual-Modality Imaging with QDs	571
22.10.1	Fluorescence/MRI	571
22.10.2	Fluorescence/CT	571
22.10.3	Fluorescence/PET	572
22.11	Conclusions	572
	References	573
23	Nanomedicine	579
	Eda Tahir Turanlı and Elif Everest	
23.1	Introduction	579
23.2	Nanopharmaceuticals	581
23.3	Diagnostic and Theranostic Nanomedicine	583
23.4	Ethics and Regulation	584
23.5	Conclusion	585
	References	586
24	Microfluidics and Its Applications in Bionanotechnology	589
	Z.P. Çakar and B. Sönmez	
24.1	Introduction to Bionanotechnology and Microfluidics	589
24.2	Microfluidic PCR Applications	592
24.3	Microfluidic DNA Microarray Systems	594
24.4	Microfluidic Applications in Electrophoresis	595
24.5	Microfluidic Bioreactors	597
24.6	Monitoring Microbial Behaviour by Microfluidics	599

24.7	The Use and Potential of Microfluidics in Microbial Strain Development.	600
24.8	Microfluidic Applications in Single Cell Studies.	602
24.9	Conclusions	602
	References.	603
25	Non-Markovian Dynamics of Qubit Systems: Quantum-State Diffusion Equations Versus Master Equations	609
	Yusui Chen and Ting Yu	
25.1	Introduction	609
25.2	Non-Markovian Quantum-State Diffusion Approach	610
25.3	Non-Markovian Master Equation Approach	615
25.4	Multiple-Qubit Systems	617
25.4.1	Two-Qubit Systems.	619
25.4.2	Three-Qubit Systems.	623
25.4.3	A Note on General N-Qubit Systems.	627
25.5	Conclusion.	628
	Appendix 1	628
	Appendix 2	630
	Appendix 3	632
	References.	633
26	Computing with Emerging Nanotechnologies	635
	M. Altun	
26.1	Introduction	635
26.2	Computing with Nano-crossbar Arrays	636
26.2.1	Implementing Boolean Logic Functions.	638
26.2.2	Defect Tolerance.	643
26.2.3	Simulation Results	649
26.3	Stochastic Computing	652
26.3.1	Reducing Error Rates	654
26.3.2	Error Free Stochastic Computing	657
26.4	Conclusions	659
	References.	659
	Index	661

Contributors

A.T. Akarsubasi Molecular Biology and Genetics Department, Istanbul Technical University, Maslak, Istanbul, Turkey

Özden Akıncı Computational Science and Engineering Programme , Informatics Institute, Istanbul Technical University, Maslak, Istanbul, Turkey

Selcuk Akturk Department of Physics, Istanbul Technical University, Maslak, Istanbul, Turkey

Ç. Allahverdi Faculty of Engineering, Electrical-Electronics Engineering Department, Toros University, Mersin, Turkey

M. Altun Electronics and Communication Engineering Department, Istanbul Technical University, Maslak, Istanbul, Turkey

Masakazu Anpo Department of Applied Chemistry, Graduate School of Engineering, Osaka Prefecture University, Sakai, Osaka, Japan

Mehmet Aras Department of Physics, Gebze Technical University, Gebze, Kocaeli, Turkey

E.U. Arici Energy Institute, Istanbul Technical University, Maslak, Istanbul, Turkey

A. Aşıkoğlu Bozkurt Faculty of Science and Letters, Department of Physics, Yıldız Technical University, Istanbul, Turkey

Takashi Azuhata Department of Advanced Physics, Hirosaki University, Hirosaki, Aomori, Japan

D. Bulut Department of Physics, İnönü Mah. Kayışdağı Cad., Yeditepe University, Ataşehir, İstanbul, Turkey

Z.P. Çakar Faculty of Science and Letters, Department of Molecular Biology and Genetics, Istanbul Technical University, Maslak, Istanbul, Turkey; Dr. Orhan Öcalgiray Molecular Biology, Biotechnology and Genetics Research Center (ITU-MOBGAM), Istanbul Technical University, Maslak, Istanbul, Turkey

B. Can Ömür Faculty of Science and Letters, Department of Physics, Yıldız Technical University, Istanbul, Turkey

Yusui Chen Department of Physics and Engineering Physics, Stevens Institute of Technology, Hoboken, NJ, USA

G. Chiarello Dipartimento di Fisica, Università degli Studi della Calabria, Rende, Italy

S. Ciraci Department of Physics, Bilkent University, Ankara, Turkey

L. Colakerol Arslan Department of Physics, Gebze Technical University, Kocaeli, Turkey

David Crompton School of Computer Science and Electronic Engineering, University of Essex, Colchester, Essex, UK

Mustafa Culha Genetics and Bioengineering Department, Yeditepe University, Faculty of Engineering and Architecture, Istanbul, Turkey

Jarek Dabrowski IHP, Frankfurt(Oder), Germany

İ. Ergal Molecular Biology and Genetics Department, İstanbul Technical University, Maslak, İstanbul, Turkey

Elif Everest Graduate School of Science Engineering and Technology, Molecular Biology-Genetics and Biotechnology Programme, Dr. Orhan Öcalgiray Molecular Biology and Genetics Research Centre, Istanbul Technical University, Istanbul, Turkey

Ash Gencturk Department of Chemistry, Polymer Science and Technology, Istanbul Technical University, Istanbul, Turkey

Sümeysra Güler-Kılıç Department of Physics, Gebze Technical University, Gebze, Kocaeli, Turkey

G. Gumbs Department of Physics, Hunter College, Cuny, NY, USA; Donostia International Physics Center (DIPC), San Sebastian, Basque Country, Spain

H. Hakan Gürel Technology Faculty, Information Systems Engineering Department, Computational Science and Engineering Programme, Kocaeli University, Kocaeli, Turkey

Manolya Kukut Hatipoglu Faculty of Engineering and Architecture, Genetics and Bioengineering Department, Yeditepe University, Istanbul, Turkey

Norman J.M. Horing Department of Physics and Engineering Physics, Stevens Institute of Technology, Hoboken, NJ, USA

R. Ince Department of Physics, Yeditepe University, Istanbul, Turkey

A.T. İnce Department of Physics, İnönüMah. Kayışdağı Cad., Yeditepe University, Ataşehir, Istanbul, Turkey

A. Iurov Center for High Technology Materials, University of New Mexico, Albuquerque, NM, USA

H. Karaağaç Department of Physics, Istanbul Technical University, Maslak, Istanbul, Turkey

Seda Kelestemur Faculty of Engineering and Architecture, Genetics and Bioengineering Department, Yeditepe University, Istanbul, Turkey; Istanbul Medipol University Regenerative and Restorative Medicine Research Center, Istanbul, Turkey

Çetin Kılıç Department of Physics, Gebze Technical University, Gebze, Kocaeli, Turkey

F.N. Kök Molecular Biology and Genetics Department, Istanbul Technical University, Maslak, Istanbul, Turkey

Juying Lei State Environmental Protection Key Laboratory of Environmental Risk Assessment and Control on Chemical Process, East China University of Science and Technology, Shanghai, People's Republic of China

Hong Li Key Lab for Advanced Materials and Institute of Fine Chemicals, East China University of Science and Technology, Shanghai, People's Republic of China

Gunther Lippert IHP, Frankfurt(Oder), Germany

Grzegorz Lupina IHP, Frankfurt(Oder), Germany

Désiré Miessein Department of Physics and Engineering Physics, Stevens Institute of Technology, Hoboken, NJ, USA

Takafumi Miyanaga Department of Advanced Physics, Hirosaki University, Hirosaki, Aomori, Japan

Z. Nassar Faculty of Science and Letters, Department of Physics, Yıldız Technical University, Istanbul, Turkey

E. Peksu Department of Physics, Istanbul Technical University, Maslak, Istanbul, Turkey

Chiara Pintossi I-LAMP, Interdisciplinary Laboratory for Advanced Materials Physics, Università Cattolica del Sacro Cuore, Brescia, Italy

A. Politano Dipartimento di Fisica, Università degli Studi della Calabria, Rende, Italy

M. Saif Islam Department of Electrical and Computer Engineering, University of California at Davis, Davis, CA, USA

Luigi Sangaletti I-LAMP, Interdisciplinary Laboratory for Advanced Materials Physics, Università Cattolica del Sacro Cuore, Brescia, Italy

A. Sezai Sarac Department of Chemistry, Polymer Science and Technology, Istanbul Technical University, Istanbul, Turkey

K.E. Smith Department of Physics, Boston University, Boston, MA, USA; School of Chemical Sciences, The MacDiarmid Institute for Advanced Materials and Nanotechnology, The University of Auckland, Auckland, New Zealand

B. Sönmez Faculty of Science and Letters, Department of Molecular Biology and Genetics, Istanbul Technical University, Maslak, Istanbul, Turkey; Dr. Orhan Öcalgiray Molecular Biology, Biotechnology and Genetics Research Center (ITU-MOBGAM), Istanbul Technical University, Maslak, Istanbul, Turkey

A. Tabatabaei Mohseni Nanoscience and Nanotechnology Program, Istanbul Technical University, Istanbul, Turkey

M. Topsakal UNAM-National Nanotechnology Research Center, Bilkent University, Ankara, Turkey

M.K. Torun Department of Physics, İnönüMah. Kayışdağı Cad., Yeditepe University, Ataşehir, İstanbul, Turkey

Mustafa Turan Department of Environmental Engineering, İstanbul Technical University, İstanbul, Turkey

Eda Tahir Turanlı Molecular Biology and Genetics Department, Istanbul Technical University, Istanbul, Turkey; Graduate School of Science Engineering and Technology, Molecular Biology-Genetics and Biotechnology Programme, Dr. Orhan Öcalgiray Molecular Biology and Genetics Research Centre, Istanbul Technical University, Istanbul, Turkey

Hilmi Ünlü Computational Science and Engineering Programme, Informatics Institute, İstanbul Technical University, Maslak, Istanbul, Turkey; Faculty of Science and Letters, Department of Physics, İstanbul Technical University, Maslak, Istanbul, Turkey

Anthony John Vickers School of Computer Science and Electronic Engineering, University of Essex, Colchester, Essex, UK

Ting Yu Department of Physics and Engineering Physics, Stevens Institute of Technology, Hoboken, NJ, USA

M.H. Yükselici Faculty of Science and Letters, Department of Physics, Yıldız Technical University, Istanbul, Turkey

E.O. Zayim Physics Department, Istanbul Technical University, Istanbul, Turkey

Jinlong Zhang Key Lab for Advanced Materials and Institute of Fine Chemicals, East China University of Science and Technology, Shanghai, People's Republic of China

Chapter 1

Modelling of Heterostructures for Low Dimensional Devices

H. Hakan Gürel, Özden Akıncı and Hilmi Ünlü

Abstract Advancement in the theoretical understanding and experimental development of the science and technology of low dimensional electronic and optical devices requires qualitatively reliable and quantitatively precise theoretical modelling of the structural, electronic and optical properties of semiconducting materials and their heterostructures to predict their potential profiles. In this chapter, we review the calculation techniques of electronic band structures of III–V and II–VI compounds and their heterostructures. We focus on the semiempirical tight binding theory (with sp^3 , sp^3s^* , $sp^3d^5s^*$ and sp^3d^5 orbital sets) and density functional theory (DFT), which, in turn, employs the modified Becke-Johnson exchange-correlation potential with a local density approximation (DFT-MBJLDA). We conclude that the density functional theory and semiempirical tight binding theory can easily be employed in relation to charge transport in heterostructure devices as well as in the accurate design and simulation of low dimensional semiconductor electronic and optical devices.

H. Hakan Gürel

Computational Science and Engineering Programme, Information Systems Engineering
Department, Technology Faculty, Kocaeli University, 41380 Kocaeli, Turkey
e-mail: hhakan.gurel@kocaeli.edu.tr

Ö. Akıncı · H. Ünlü (✉)

Computational Science and Engineering Programme, Informatics Institute, İstanbul
Technical University, 34469 Maslak, İstanbul, Turkey
e-mail: hunlu@itu.edu.tr

Ö. Akıncı

e-mail: akincio@gmail.com

H. Ünlü

Department of Physics, Faculty of Science and Letters, İstanbul Technical University, 34469
Maslak, İstanbul, Turkey

1.1 Introduction

Progress in the theoretical understanding and experimental development of the science and technology of semiconductor devices has grown rapidly ever since the invention of first transistor in 1949 [1–6]. Advances in growing semiconductor thin films having differing structural, electronic and optical properties with varying composition and layer thickness (even approaching atomic dimensions) have provided new opportunities and challenges in basic scientific studies and in their device applications in the electronics industry. The combination of advanced growth, characterization and fabrication technologies has led to the production of an impressive number of high performance semiconductor devices for fast signal processing and the discovery of some novel structures that are of special interest to solid state scientists and device engineers. One can now control the alloy composition and doping in ternary and quaternary alloys based on group IV–IV, III–V and II–VI semiconductor compounds over atomic distances, as low as tens of angstroms. These man-made heterostructures, are easily grown as lattice mismatched substrates (e.g. growth of GaAs on Si substrates [5, 6]) by modern crystal growth techniques including molecular beam epitaxy (MBE) and metal organic chemical vapor deposition (MOCVD). The epitaxial layers are so thin that quantum mechanical effects are prominently realized. The resulting contemporary low dimensional semiconductor devices (e.g., heterostructure bipolar transistors (HBTs) and modulation doped field effect transistors (MODFETs), nanowires, quantum dots, etc.) are known to operate much faster than conventional silicon devices (e.g., silicon bipolar junction transistors (BJTs) and metal oxide semiconductor field effect transistors (MOSFETs)); this is crucial for the electronic and optical communication and computer industries.

When the semiconductor composition changes abruptly across the interface between two constituents, the difference in their energy bandgaps is accommodated by the discontinuities in the conduction and valence bands [7, 8]. There are three types of interface formation between a wide bandgap semiconductor and a narrow bandgap semiconductor shown in Figs. 1.1, 1.2 and 1.3;

- (i) **Type I heterostructure:** The bandgap of the barrier semiconductor overlaps (straddling lineup) that of the well and equilibrium Fermi level is near the middle of the bandgap on both sides, as shown in Fig. 1.1. Both an electron and a hole tend to localize in the narrow bandgap quantum well. Such semiconductor heterostructure is useful in optoelectronic applications such as lasers, with both electrons and holes participate to device operation.
- (ii) **Type II heterostructure:** The bandgap of the barrier semiconductor partially overlaps (staggered lineup) that of the well and the equilibrium Fermi level is close to the conduction band (or valence band) on one side while it is near the middle of the bandgap on other side, as shown in Fig. 1.2. The potential energy gradient tends to spatially separate electron and hole on different sides of heterointerface. Large potential barrier in conduction (or valence) band at heterointerface leads to a better electron (hole) confinement in field effect transistors with higher electron (hole) concentration. Position of equilibrium

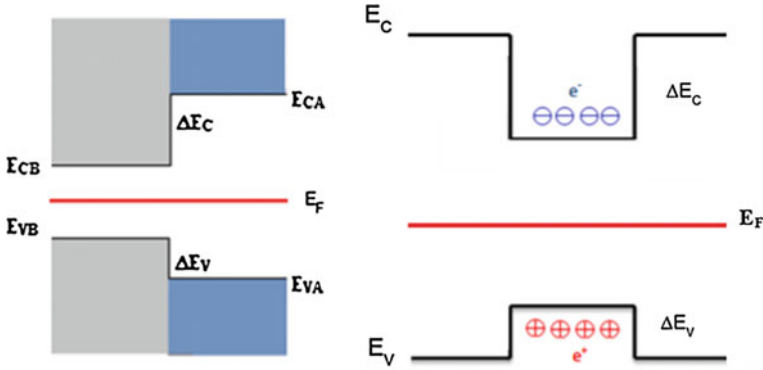


Fig. 1.1 The schematic energy band diagram of Type I heterostructure: electrons and holes localize in the narrow gap quantum well

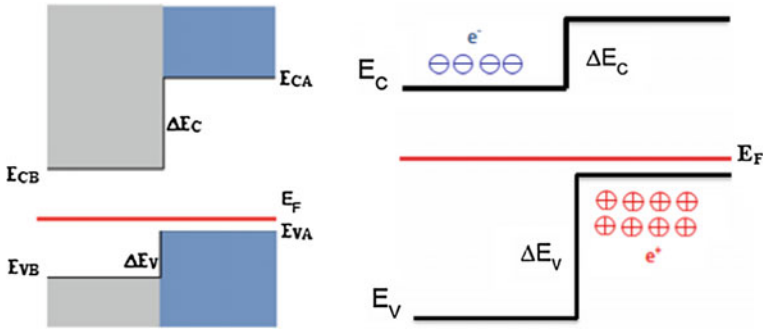


Fig. 1.2 Schematic energy band diagram of Type II heterostructure: the potential energy gradient tends to spatially separate electrons and holes on different sides of heterointerface

Fermi level determines the density of carrier confinement on either side of the heterointerface. When equilibrium Fermi level is close to conduction band of wide bandgap constituent (left) the density of electrons in its conduction band is greater but when it is closer to the valence band of narrow bandgap constituent (right) the number of holes is greater there.

- (iii) **Type III heterostructure:** This is the extreme case of Type II band alignment. The bandgap of the barrier does not overlap at all (broken gap lineup), as shown in Fig. 1.3. The equilibrium Fermi level is well above the conduction band minimum (or well below the valence band maximum) on one side while it is near the middle of the bandgap on the other side. Position of equilibrium Fermi level determines the carrier confinement on either side of the heterointerface. When equilibrium Fermi level is close to conduction band of wide bandgap constituent (left) the number of electrons in its conduction band is greater but when it is closer to the valence band of narrow bandgap constituent (right) the number of holes is greater there.

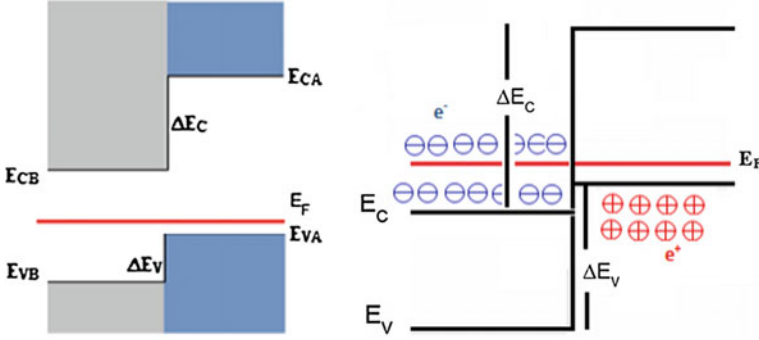


Fig. 1.3 Schematic energy band diagram of Type II heterostructure: the potential energy gradient tends to spatially separate electrons and holes on different sides of heterointerface

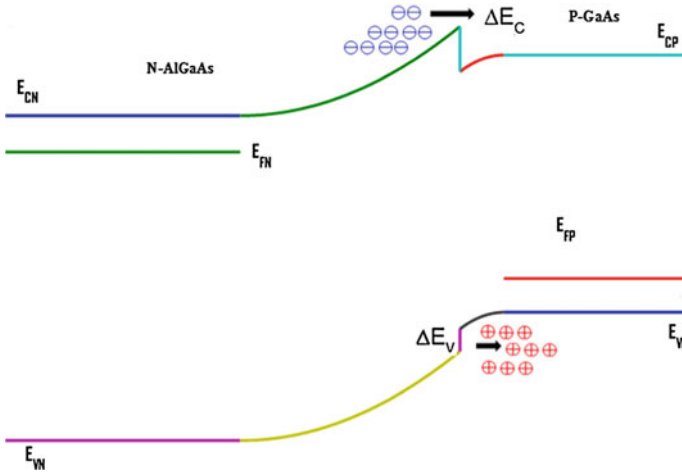
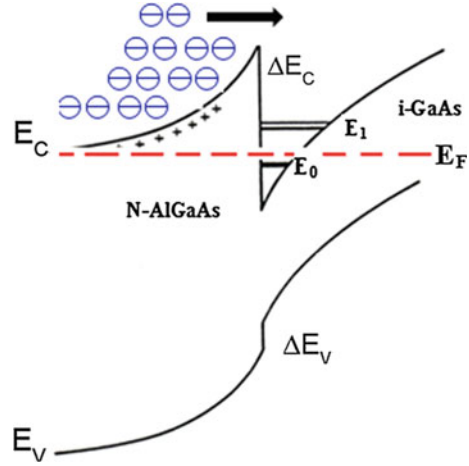


Fig. 1.4 The schematic energy band diagram of a forward biased N-AlGaAs/p-GaAs emitter/base part a of Npn heterojunction bipolar transistor (HBT) in forward active mode

The impact of heterostructures on semiconductor device physics and technology is twofold [1, 2]. The first impact of heterostructures is a high charge carrier injection efficiency that can be obtained in an anisotype heterojunction bipolar transistor (HBT), in which the charge carriers are flowing from a wide bandgap emitter (e.g., AlGaAs) to narrow bandgap base (e.g., GaAs) as shown in Fig. 1.4. When a Np hetero-emitter is forward biased, the potential barrier blocking electron emission from the emitter to the base is lowered and electrons injected from the emitter diffuse across the base and are collected in the collector. Furthermore, holes injected from the base into the emitter are blocked by the valence band offset as $\Delta E_v/kT$.

The second impact of heterostructures involves the confinement of charge carriers in a narrow bandgap two dimensional electron gas (2DEG) quantum well structure to reduce their scattering by parent impurities in a doped wide bandgap

Fig. 1.5 Schematic energy band diagram of an AlGaAs/GaAs 2DEG. The triangular well in a narrow bandgap GaAs well leads to quantization of the allowed energy levels E_0 and E_1 near the heterointerface



barrier layer, as shown in Fig. 1.5. The charge carrier scattering that limits the high speed character of MOSFETs can be minimized by replacing an oxide/semiconductor junction with a widegap/narrowgap semiconductor heterojunction to separate the free charge carriers in a nominally undoped narrow bandgap quantum well (e.g., i-GaAs) from their parent impurities in the doped wide bandgap (e.g., N-AlGaAs) barrier layer, leading to a much higher mobility and a high current from the drain to the source of a MODFET [2, 4].

Conduction and valence band offsets at the AlGaAs/GaAs interface lead to a depletion of electrons from the wide-bandgap AlGaAs barrier and an accumulation of electrons in the narrow-bandgap GaAs channel where they are quantized due to the potential ‘notch’ at the interface. Quantized energy levels are formed in this well, with the lowest level filled and the second level partially or fully filled by electrons, depending on the electron density; the Fermi level lies between the first and the second sub-bands. Free electrons confined in the channel conduct current parallel to the hetero-interface between the source and the drain of a FET when an electric field is applied parallel to the hetero-interface. As the wide bandgap and narrow bandgap semiconductors reach thermal equilibrium, their Fermi levels align at the heterointerface and the resultant energy band structure is as shown in Fig. 1.4 for a Np heterostructure bipolar device and in Fig. 1.5 for a unipolar heterostructure device, respectively.

1.2 Issues in Modelling of Electronic Structure in Heterostructures

Ever since the invention of the first transistor in 1949, device scientists and engineers have witnessed an amazing growth in semiconductor science and technology. Advances in the growth of semiconductor thin films of different structural,

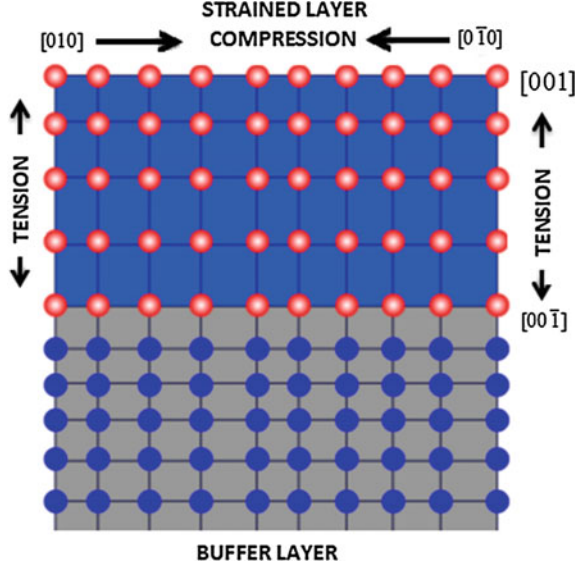
electronic and optical properties and with layer thickness approaching atomic dimensions has provided new opportunities in fundamental science and technology of semiconductors for device applications. Furthermore, contemporary fabrication technologies have further made it possible to reduce the device dimensions to the point where quantum size effects must be described in order to realistically describe the operation and reliably predict the potential and performance of low dimensional semiconductor devices for electronic and optical applications. Such advancement could not have been possible without a qualitatively reliable understanding of the basic physics of semiconducting materials and quantitatively precise potential predictions and performance of devices, leading to new concepts in the semiconductor growth that allowed previously many unknown devices with more complex functionality and much higher densities for electronic and optical applications.

In order to emphasize the importance of heterostructure in low dimensional semiconductor systems it is necessary to understand the interface formation and modelling of the electronic structure of heterostructure constituents. The composition variation in alloy constituent and lattice mismatch between two semiconductors and thermal expansion of lattice constant with the growth temperature will cause strain across the interface that will modify the electronic properties of both materials and consequently, the energy of the moving charge carriers across the interface that influence the device performance [7, 8]. Qualitatively reliable and quantitatively precise modelling and simulation of electronic properties of semiconductor constituents is important for a better prediction of their potential in making low dimensional electronic and optical devices. In the following we will give a brief discussion about the critical structural properties and how they can be implemented in theoretical models for the calculations of electronic properties of heterostructures.

1.2.1 Interface Strain Effects in Heterostructures

When two semiconductors with different lattice constants are grown upon each other, strains will develop across the heterointerface increasing with layer thickness. As long as the thickness of epilayer is kept under a critical thickness, the lattice mismatch will be accommodated by uniform elastic strain, as shown in Fig. 1.6. The interface strain will modify the structural and electronic properties of the constituent semiconductors in directions parallel and perpendicular to the growth direction. Therefore, the lattice matching is known to be important to the electronic properties of semiconductor layers in heteroepitaxy. Until the early 1980s, the studies of heterostructures had focused on semiconducting materials with close lattice matching with a substrate, such as an AlGaAs/GaAs heterostructure quantum well. The constraint on lattice matching in crystal growth is essential for relatively thick epitaxial layers since a large lattice mismatch can cause the occurrence of misfit dislocations with large densities. Such lattice-matched growth will prevent the generation of misfit dislocations that would degrade charge transport and lower device performance.

Fig. 1.6 Strained heteroepitaxy is a result of growth of materials with dissimilar lattice constants



The lattice constant of strained epilayer along the growth direction is equal to that of the substrate and is expanded or compressed in the direction parallel to the interface, shown in Fig. 1.6. Considering the (001) heteroepitaxial growth, elastic theory allows one to decompose the biaxial strain tensor into the sum of hydrostatic and uniaxial strains along the growth direction.

$$\varepsilon_{xx} = \varepsilon_{yy} = \varepsilon_{f||} = \left(\frac{a_{f||}}{a_{f0}} - 1 \right), \quad \varepsilon_{zz} = \varepsilon_{f\perp} = \left(\frac{a_{f\perp}}{a_{f0}} - 1 \right) = 2 \left(\frac{C_{12}}{C_{11}} \right)_f \varepsilon_{f||} \quad (1.1)$$

where $\varepsilon_{f\perp}$ and $\varepsilon_{f||}$ are the strain components perpendicular and parallel to growth direction. The epilayer lattice constant will be equal to that of the substrate along the growth direction: $a_{f||} = a_s$ and is expanded by the bulk value of its Poisson ratio perpendicular to the growth direction:

$$a_{f\perp} = a_{f0} \left[1 - 2 \left(\frac{C_{12}}{C_{11}} \right)_f \left(\frac{a_{f||}}{a_{f0}} - 1 \right) \right], \quad (1.2)$$

where, a_{f0} is the bulk lattice constant and C_{11} and C_{12} are the bulk elastic stiffness constants of the strained layer. a_s is the bulk lattice constant of the substrate. Equations (1.1) and (1.2), suggest that when a thin layer is deposited on a buffer (or substrate) with large mismatch, the epilayer will be under strain. The lattice constant of the epilayer along and perpendicular to the interface will change to minimize its elastic energy. The interface strains due to lattice mismatch and thermal expansion gradients over the crystal growth temperature can co-exist in heterolayers in low dimensional heterostructures; it is difficult to assign the observed stress to

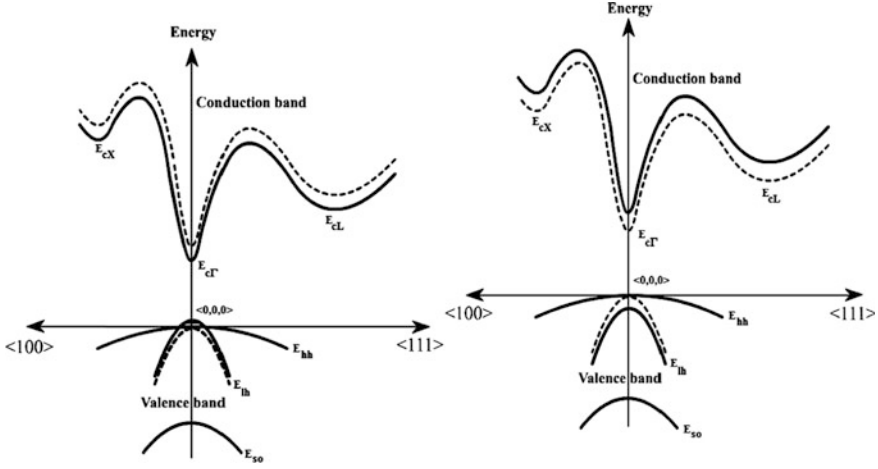


Fig. 1.7 The schematic view of energy band diagram of a semiconductor under tensile strain bandgap decreases (*left*) and compressive strain bandgap increases (*right*). Heavy hole (HH), light hole (LH) and split-off (SO) bands are also shifted with respect to top of the valence band energy

either one or the other. The thermal expansion of lattice constants of epilayer and the substrate with the growth temperature cause the lattice constant parallel and perpendicular to the hetero-interface to change with temperature. High resolution x-ray scattering measurements of GaAs grown on a Si substrate [9] show that the thermal expansion of GaAs perpendicular to the growth direction follows the thermal expansion of a Si substrate $\alpha_{f\parallel}(T) = \alpha_s(T)$ and is therefore smaller than that of bulk GaAs. However, the thermal expansion along the growth direction exceeds the bulk GaAs value by the Poisson ratio as a result of the in-plane constraint;

$$\alpha_{f\perp}(T) = \alpha_f(T) + 2 \frac{C_{12f}}{C_{11f}} [\alpha_f(T) - \alpha_s(T)], \quad (1.3)$$

where $\alpha_s(T)$ and $\alpha_f(T)$ are the bulk linear thermal expansion coefficients of Si and GaAs. Strain across the interface that will modify the electronic properties of both materials, including the band offsets. Compressive (or tensile) strain in epilayer results in an increase (or decrease) in conduction and valence band energy levels [2], as shown in Fig. 1.7.

The uniaxial component of the biaxial strain tensor splits the heavy-hole, light-hole and split-off valence band edges relative to the average valence band edge. The heavy-hole, light-hole and split-off band energies relative to the average valence band edge E_v are

$$E_{vh}(\varepsilon) = E_v(\varepsilon) + \frac{1}{3}\Delta - \frac{1}{2}\delta E, \quad (1.4a)$$

$$E_{vl}(\varepsilon) = E_v(\varepsilon) - \frac{1}{6}\Delta + \frac{1}{4}\delta E + \frac{1}{2}\sqrt{\Delta^2 + \Delta\delta E + \frac{9}{4}\delta E^2}, \quad (1.4b)$$

$$E_{vs}(\varepsilon) = E_v(\varepsilon) - \frac{1}{6}\Delta + \frac{1}{4}\delta E - \frac{1}{2}\sqrt{\Delta^2 + \Delta\delta E + \frac{9}{4}\delta E^2}, \quad (1.4c)$$

where $\delta E = 2b(\varepsilon_{zz} - \varepsilon_{xx}) = 2b(\varepsilon_{\perp} - \varepsilon_{\parallel})$ and b is the shear deformation potential which describes splitting in the valence band energy due to the [001] uniaxial strain. $E_v(\varepsilon)$ is the average valence band maximum under hydrostatic strain. The hydrostatic component of biaxial strain tensor corresponds to the relative volume change of the strained epilayer; $Tr(\varepsilon) = 2\varepsilon_{\perp} + \varepsilon_{\parallel} = \Delta V/V_0 = (V - V_0)/V_0$, which leads to shifts of the conduction band minimums relative to the average valence band maximum at the Γ point.

According to a statistical thermodynamic model in which the conduction electrons and valence holes are treated as charged chemical particles, the conduction and valence band energy levels energy are expressed as a function of pressure at any temperature as [10, 11]

$$E_i(T, P) = E_i(0, P_0) + C_{iP}^0 T(1 - \ln T) - \frac{a_i}{B} \left[P - \frac{P^2}{2B} - \frac{(1 + B')}{6B^2} P^3 \right], \quad (1.5)$$

where i represents the conduction (c) or valence (v) band energy levels, P is the applied pressure, T is the temperature, $E_i(0, P_0)$ are the conduction or valence band edges at Γ , L and X high symmetry points. E_{ci} and E_v are the conduction and valence band edges at Γ , L and X points with deformation potentials $a_{ci} = -B(\partial E_{ci}/\partial P)$ and $a_v = -B(\partial E_v/\partial P)$. B is the bulk modulus; $B' = \partial B/\partial P$ is its derivative. C_{iP}^0 is the standard heat capacity of conduction electrons and valence holes. $C_{cP}^0 = C_{nP}^0 - C_{0P}^0 = C_{pP}^0 + \Delta C_P^0$ and $C_{vP}^0 = C_{pP}^0$ are the standard heat capacities of conduction electrons and valence holes; $C_{nP}^0 = C_{pP}^0 = (5/2)k$, where k is the Boltzmann's constant. $\Delta C_P^0 = C_{nP}^0 + C_{pP}^0 - C_{0P}^0$ is the heat capacity of reaction of free electron and hole formation obtained by fitting (1.6) to the experimental data [12, 13] and empirical pseudopotential bandgap energy [14] and at high symmetry points in the first Brillouin zone of semiconductors:

$$E_{gl}(T, P) = E_{gi} + \Delta C_P^0 T(1 - \ln T) - \frac{a_{gl}}{B} \left[P - \frac{P^2}{2B} - \frac{(1 + B')P^3}{6B^2} \right], \quad (1.6)$$

where $P = -2B_f C_f \varepsilon_{f\parallel}$ and $P = -3B_s \varepsilon_{s\parallel}$ for the epilayer and the substrate, respectively. In the case of (001) pseudomorphic growth, $\varepsilon = \varepsilon_{f\parallel} = (a_{f\parallel} - a_f)/a_f$ is the strain in the epilayer along the growth direction and $C_f = (C_{11} - C_{12})/C_{11}$.

Any interface strain will modify the energy band structure and charge transport in heterobipolar and unipolar devices. It should be noted that as the device dimension is reduced to nanoscale of the order of Bohr radius of charged carriers, the use of continuum elastic theory becomes questionable [15]. In such a case one should use the atomic elasticity theory known as valence force field (VFF) approach [16], especially as the lattice mismatch at heterointerface becomes large.

1.2.2 Composition Effects in Heterostructures

Semiconductor alloys based on IV–VI, III–V and II–VI compounds are important in fabricating low dimensional bipolar and unipolar heterostructure devices since their structural and electronic properties (e.g., lattice constants and bandgaps) can be tailored independently. Therefore, reliable and accurate determination of composition variation of lattice constant and bandgap energies are very important. In the theoretical determination of composition effects on the structural properties such as lattice constants, a virtual crystal approximation (VCA) is often used [17, 18] in which the compositional disorder effect is neglected. Since in VCA the alloy potential is taken as the concentration weighted average of the constituent potentials the bandgap energy is linear function of alloy composition. However, many experimental studies report that the bandgap energies of semiconductor alloys is nonlinear function of composition. Furthermore, it is believed that the compositional disorder, which is related to the differences of electronegativity of atoms forming a ternary semiconductor, plays a major role in determining its bowing of the bandgap energy when lattice mismatch induced strain plays a crucial role in heterostructure electronic properties.

In our recent studies [19–28] the effects of composition and strain on the electronic properties of semiconductor alloy constituents in heterostructures are implemented in the semiempirical tight binding models and first principles DFT with MBJLDA functional in terms of host bond length and distorted bond length by the substitutional impurity without any adjustable parameter. In determining the composition effects on band structures, we do not follow the common practice which employs the so called virtual crystal approximation (VCA) to the TB Hamiltonian matrix elements and bond length, with or without the compositional disorder of the semiconductor alloys, since the detailed treatment requires inclusion of the compositional disorder effect on the electronic properties. Instead, we employ the method of the modified virtual crystal approximation (MVCA) in which one formulates the composition dependence of the bond length of ternary material [19–28]. The MVCA allows one to accurately take into account the effect of disorder-induced nonlinear variation of the lattice constant on the TB parameters used in calculating the band structure properties. The composition dependence of bond length (or lattice constant) of the ABC ternary is written as the sum of

undistorted bond length ($d_{VCA} = (1-x)d_{AC}^0 + xd_{BC}^0$) due to the virtual crystal approximation (VCA) and the distorted bond length ($d_{relax} = x(1-x)\delta_c(d_{BC}^0 - d_{AC}^0)$) due to cation-anion relaxation of binary in ternary [19–28]:

$$d(x) = (1-x)d_{AC}(x) + xd_{BC}(x) = d_{VCA}(x) + x(1-x)\delta_c(d_{AC}^0 - d_{BC}^0), \quad (1.7a)$$

$$d_{BC}(x) = d_{BC}^0 - (1-x)\xi_{AC:B}(d_{BC}^0 - d_{AC}^0), \quad (1.7b)$$

$$d_{AC}(x) = d_{AC}^0 - x\xi_{BC:A}(d_{AC}^0 - d_{BC}^0), \quad (1.7c)$$

where $d_{AC}(x)$ and $d_{BC}(x)$ are the bond lengths of AC and BC binaries in an ABC ternary, d_{AC}^0 and d_{BC}^0 are the undistorted bond lengths of the host materials AC and BC and $\xi_{AC:B}$ and $\xi_{BC:A}$ are two dimensionless relaxation parameters [29],

$$\xi_{AC:B} = \frac{1}{1 + \frac{\alpha_{AC}}{6\alpha_{BC}}(1 + 10\frac{\beta_{AC}}{\alpha_{AC}})}, \quad \xi_{BC:A} = \frac{1}{1 + \frac{\alpha_{BC}}{6\alpha_{AC}}(1 + 10\frac{\beta_{BC}}{\alpha_{BC}})}, \quad (1.7d)$$

δ_c is the difference between dimensionless relaxation parameters: $\delta_c = \xi_{AC:B} - \xi_{BC:A}$. As an example, Fig. 1.8 shows the composition variation of the lattice constants and interface strain in AlGaIn and InGaIn ternaries in AlGaIn/GaN and InGaIn/GaN heterostructures, respectively. Compressive strain at the InGaIn/GaN heterointerface decreases the lattice constant of the InGaIn ternary constituent as composition increases. Whereas tensile strain at the AlGaIn/GaN heterointerface increases the lattice constant of the AlGaIn ternary constituent as composition increases, such an increase or decrease in lattice constant due to interface strain will change the electronic properties of the heterostructure ternary constituents. Therefore, it is essential that we include the compositional disorder effect on the structural and electronic properties of constituents in lattice mismatched heterostructures. Equation (1.7a–1.7d) can then be used to take into account the composition effects on the band structures of the ternary constituent of heterostructures in the theoretical models such as first principles WIEN2K simulations package with the MBJLDA functional embedded in DFT and semiempirical sp^3 , sp^3s^* , $sp^3d^5s^*$ and sp^3d^5 tight binding models.

The diagonal or off-diagonal matrix elements in the NN sp^3d^5 and 2NN sp^3s^* TB Hamiltonian matrix elements for an ABC ternary semiconductor are expressed as [19–28]

$$E_{\alpha/\beta}(x) = (1-x)E_{\alpha/\beta}(AC) + xE_{\alpha/\beta}(BC) + x(1-x)[E_{\alpha/\beta}(AC) - E_{\alpha/\beta}(BC)], \quad (1.8)$$

where $E_{\alpha/\beta}(AC)$ and $E_{\alpha/\beta}(BC)$ represent the fitted energies of the s, p and d states of anion and cation atoms forming the AC and BC binary compounds. This allows

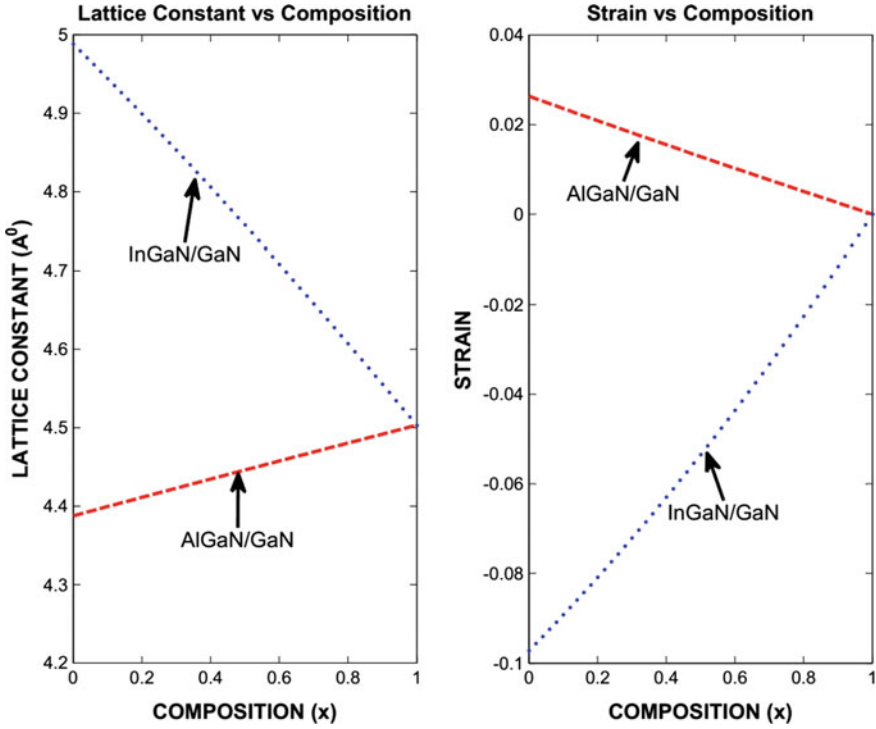


Fig. 1.8 Lattice constant versus alloy composition (*left*) and interface strain versus alloy composition (*right*) for AlGaIn/GaN and InGaIn/GaN GaN heterostructures

one to take into account the composition variations of the diagonal and off-diagonal elements of the TB Hamiltonian matrix for a ternary ABC semiconductor can be taken as a nonlinear function of alloy composition.

1.3 Semiempirical Tight Binding Modeling of Heterostructures

Advances in growing multilayered structures with layer thicknesses approaching atomic dimensions have provided new opportunities and new challenges for the theoretical modelling of electronic structures of low dimensional heterostructures based on the universally accepted fundamental principles of solid state physics and quantum mechanics. The electronic properties of low dimensional bipolar and unipolar semiconductor device structures are often calculated based on the following theoretical models: (i) First principle *ab initio* methods, such as density functional theory (DFT) [29, 30], (ii) *methods*, such as local/empirical pseudopotential method [14], or (iii) *tight binding* method (also known as the linear

combination of atomic orbitals (LCAO) method) [19–28, 31–38], and *k.p method* [39, 40]. The first principles ab initio methods are computationally expensive and require heavy parallel computations. Since the ab initio models are based on the calculations of ground state properties, they can only give limited physical insight about the energy band structure at high temperatures and pressures.

On the other hand, the semiempirical tight binding models are less expensive than the ab initio methods and can easily be implemented in calculating electronic properties of low dimensional heterostructures such as nanowires or quantum dots. The tight binding model has recently become popular, because of its simplicity and ability to realistically describe the structural and dielectric properties of semiconductors in terms of chemical bonds. The tight binding model is an atomistic approach and is well suited for calculating the electronic band structure of semiconductor heterostructures such as two dimensional quantum wells and zero dimensional quantum dots. In the following Sects. (1.3.1–1.3.4), we will give a detailed discussion of the semiempirical tight binding theory with first and second nearest neighbor (NN and 2NN) sp^3 and sp^3s^* atomic orbitals sets and first nearest neighbor (NN) sp^3d^5 orbitals set with spin-orbit coupling of cation (Al, Ga; In, Cd, Zn) and anion (P, As, Sb, S, Se, Te) atoms for calculating the electronic structure of III–V and II–VI compounds and their heterostructures. We shall later on (Sect. 1.4) compare the semiempirical tight binding models with the density functional theory (DFT) that uses the modified Becke-Johnson exchange-correlation potential with the local density approximation (LDA), called MBJLDA functional, for calculating the band structure of group III–V and II–VI compounds and their alloys.

1.3.1 Semiempirical Sp^3 Tight Binding Modeling

The elementary tight-binding method, also known as the linear combination of atomic-orbitals (LCAO), provides a physically insightful formulation of energy band theory in a crystalline solid, which permits tracing trends from system to system entirely in terms of the fundamental physical features [31–34]. In the semiempirical tight binding theory of semiconductors one assumes that the valence electrons are tightly bound to their nuclei as in the free atom. Anion and cation atoms are brought together until their separations become comparable to the lattice constant of semiconductors, at which point their wave functions will overlap. In the Slater-Koster formalism, the crystal potential of a binary semiconductor is defined as a sum of the symmetrical spherical potentials around each atom. This allows the electronic wave function $\psi_{\alpha k}$ to be written as linear combination of atomic orbitals $\varphi_{\alpha}(k)$ (called the Löwdin orbitals):

$$|\psi_{\alpha k}\rangle = \sum_{\alpha} u_{\alpha} |\varphi_{\alpha}(k)\rangle = \sum_i \frac{e^{i \cdot k \cdot r_i}}{\sqrt{N}} |\varphi_{\alpha}(r - r_i)\rangle, \quad (1.9)$$

where $\mathbf{r}$ is the position of electron with respect to origin of coordinate system is real space, $\mathbf{k}$ is the wave vector, N is the number of atoms, and $\mathbf{r}_i$ is the crystal lattice site. The energy state is given by the eigenvalue of the linear equation written in matrix form [31–34];

$$\sum_{\beta} [H_{\alpha\beta}(k) - S_{\alpha\beta}(k)E] u_{\beta} = 0, \quad (1.10)$$

where E is the eigenvalue and $H_{\alpha\beta} = \langle \varphi_{\alpha}(k) | H | \varphi_{\beta}(k) \rangle$ and $S_{\alpha\beta} = \langle \varphi_{\alpha}(k) | \varphi_{\beta}(k) \rangle$ are the Hamiltonian and overlap matrices, respectively.

$$H_{\alpha\beta} = \begin{bmatrix} H_{cc} & H_{ca} \\ H_{ac} & H_{aa} \end{bmatrix}, \quad S_{\alpha\beta} = \begin{bmatrix} 1 & S_{ca} \\ S_{ac} & 1 \end{bmatrix}, \quad (1.11)$$

Here α correspond to a cation s (p) orbital and β corresponds to an anion s (p) orbital and u_{β} is the wave function coefficient. In the orthogonal sp^3 tight binding formalism in which only the first nearest neighbor interactions are taken into account, the Bloch functions in (1.8) are taken to be orthogonal so that the overlap matrix $S_{\alpha\beta}(k)$ is an identity matrix. There are nine independent matrix elements, namely the four on-site atomic energies, E_{sa}, E_{sc}, E_{pa} and E_{pc} (with a = anion and c = cation) and five hopping terms, $E_{ss}, E_{xx}, E_{sapc}, E_{scpa}$ and E_{xy} . The diagonal elements (E_{sa}, E_{sc}, E_{pa} and E_{pc}) are expressed as the (4×4) matrix

$$H_{aa} = \begin{bmatrix} E_s^a & 0 & 0 & 0 \\ 0 & E_p^a & 0 & 0 \\ 0 & 0 & E_p^a & 0 \\ 0 & 0 & 0 & E_p^a \end{bmatrix}, \quad H_{cc} = \begin{bmatrix} E_s^c & 0 & 0 & 0 \\ 0 & E_p^c & 0 & 0 \\ 0 & 0 & E_p^c & 0 \\ 0 & 0 & 0 & E_p^c \end{bmatrix}, \quad (1.12a)$$

The off-diagonal elements ($E_{ss}, E_{xx}, E_{sapc}, E_{scpa}$ and E_{xy}) are expressed as the (4×4) matrix

$$H_{ac} = \begin{bmatrix} B_0 E_{ss} & B_1 E_{sp} & B_2 E_{ss} & B_3 E_{sp} \\ -B_0 E_{sp} & B_0 E_{xx} & B_3 E_{xy} & B_2 E_{xy} \\ -B_2 E_{sp} & B_3 E_{xy} & B_0 E_{xx} & B_1 E_{xy} \\ -B_3 E_{sp} & B_2 E_{xy} & B_1 E_{xy} & B_0 E_{xx} \end{bmatrix}, \quad (1.12b)$$

$$H_{ca} = \begin{bmatrix} B_0^* E_{ss} & -B_1^* E_{sp} & B_2^* E_{ss} & B_3^* E_{sp} \\ B_1^* E_{sp} & B_0^* E_{xx} & B_3^* E_{xy} & B_2^* E_{xy} \\ B_2^* E_{sp} & B_3^* E_{xy} & B_0^* E_{xx} & B_1^* E_{xy} \\ B_3^* E_{sp} & B_2^* E_{xy} & B_1^* E_{xy} & B_0^* E_{xx} \end{bmatrix},$$

where B_i^* is the complex conjugate of matrix element B_i , defined in (1.15a). The diagonal matrix elements, representing the interactions between the same p-orbitals of cation and anion atoms yield symmetric sums; off-diagonal matrix elements

representing the interactions between different p-orbitals of cation and anion atoms yield asymmetric sums, given as:

$$\langle p_x^c | H | p_x^a \rangle = \langle p_y^c | H | p_y^a \rangle = \langle p_z^c | H | p_z^a \rangle = E_{xx} B_0(k), \quad (1.13a)$$

$$\langle p_x^c | H | p_y^a \rangle = \langle p_y^c | H | p_x^a \rangle = E_{xy} B_3(k), \quad \langle p_y^c | H | p_z^a \rangle = \langle p_z^c | H | p_y^a \rangle = E_{xy} B_1(k), \quad (1.13b)$$

$$\langle p_y^c | H | p_z^a \rangle = \langle p_z^c | H | p_y^a \rangle = E_{xy} B_1(k), \quad \langle p_x^c | H | p_z^a \rangle = \langle p_z^c | H | p_x^a \rangle = E_{xy} B_2(k), \quad (1.13c)$$

Interactions between an s orbital of atom A and the p_x, p_y, p_z orbitals of neighboring atoms B are

$$\langle s^c | H | s^a \rangle = E_{ss} \sum_{n=1}^4 e^{ik \cdot r_n} = E_{ss} B_0(k), \quad \langle s^c | H | p_x^a \rangle = E_{sp} \sum_{n=1}^4 e^{ik \cdot r_n} = E_{sp} B_1(k), \quad (1.14a)$$

$$\langle s^c | H | p_y^a \rangle = E_{sp} \sum_{n=1}^4 e^{ik \cdot r_n} = E_{sp} B_2(k), \quad \langle s^c | H | p_z^a \rangle = E_{sp} \sum_{n=1}^4 e^{ik \cdot r_n} = E_{sp} B_3(k), \quad (1.14b)$$

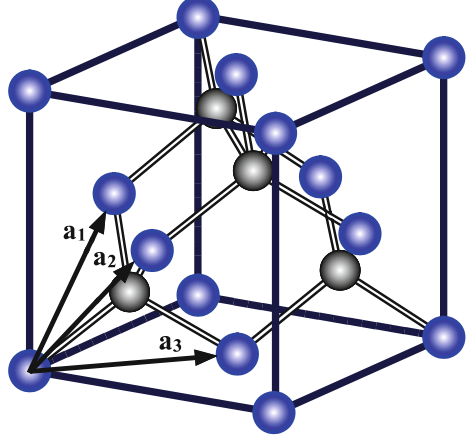
$$B_0(k) = 4 \cos\left(\frac{k_x a}{2}\right) \cos\left(\frac{k_y a}{2}\right) \cos\left(\frac{k_z a}{2}\right) - 4i \sin\left(\frac{k_x a}{2}\right) \sin\left(\frac{k_y a}{2}\right) \sin\left(\frac{k_z a}{2}\right), \quad (1.15a)$$

$$B_1(k) = 4 \cos\left(\frac{k_x a}{2}\right) \sin\left(\frac{k_y a}{2}\right) \sin\left(\frac{k_z a}{2}\right) + 4i \sin\left(\frac{k_x a}{2}\right) \cos\left(\frac{k_y a}{2}\right) \cos\left(\frac{k_z a}{2}\right), \quad (1.15b)$$

$$B_2(k) = -4 \sin\left(\frac{k_x a}{2}\right) \cos\left(\frac{k_y a}{2}\right) \sin\left(\frac{k_z a}{2}\right) + 4i \sin\left(\frac{k_x a}{2}\right) \sin\left(\frac{k_y a}{2}\right) \cos\left(\frac{k_z a}{2}\right), \quad (1.15c)$$

$$B_3(k) = -4 \sin\left(\frac{k_x a}{2}\right) \sin\left(\frac{k_y a}{2}\right) \cos\left(\frac{k_z a}{2}\right) - 4i \cos\left(\frac{k_x a}{2}\right) \cos\left(\frac{k_y a}{2}\right) \sin\left(\frac{k_z a}{2}\right), \quad (1.15d)$$

Fig. 1.9 Crystal unit cell group III–V and II–VI compounds with zinc-blende structure



where $i = \sqrt{-1}$ and $r_1 = (a/2)(1, 1, 1)$, $r_2 = (a/2)(1, -1, -1)$, $r_3 = (a/2)(-1, 1, -1)$ and $r_4 = (a/2)(-1, -1, 1)$ are the displacement vectors of nearest neighbors. Figure 1.9 shows the schematic view of unit cells for tetrahedral semiconductors.

In the first nearest neighbor (NN) sp^3 tight binding theory, the acceptable features of electronic band structure of semiconductors are dominated by the diagonal terms (on-site atomic energies: E_{sa} , E_{sc} , E_{pa} and E_{pc}) and off-diagonal terms (E_{ss} , E_{xx} , E_{sapc} , E_{scpa} and E_{xy}) representing the nearest neighbor interactions. The first two off-diagonal elements E_{ss} and E_{xx} can be obtained from the bandgap at the Γ point ($k_x = k_y = k_z = 0$) for which B_1, B_2 and B_3 are all zero, but B_0 is nonzero. The solution of the matrix (1.8) at the Γ point ($k = 0$) then yields

$$E(\Gamma_{1c/v}) = \frac{1}{2}(E_s^a + E_s^c) \pm \sqrt{[(E_s^a - E_s^c)/2]^2 + 16E_{ss}^2}, \quad (1.16)$$

$$E(\Gamma_{15c/v}) = \frac{1}{2}(E_p^a + E_p^c) \pm \sqrt{[(E_p^a - E_p^c)/2]^2 + 16E_{xx}^2}, \quad (1.17)$$

where $E(\Gamma_{15v})$ are $E(\Gamma_{1c})$ the top of the triply degenerate valence band and bottom of the conduction band, respectively. The two unknown parameters E_{ss} and E_{xx} can be determined by inverting (1.7) provided we have accurate s and p energies E_s and E_p for the anion and cation atoms, the bandgap at Γ point and the width of the valence band. The next three off-diagonal matrix elements E_{sapc} , E_{scpa} and E_{xy} are obtained from the bandgap data at the X and L high symmetry points. At the X high symmetry point ($k_x = 1, k_y = k_z = 0$) the Block sums are $B_0 = 0, B_1 = 4i, B_2 = 0$ and $B_3 = 0$, so that conduction and valence band energies are

$$E(X_{3c/v}) = \frac{1}{2}(E_s^a + E_p^c) \pm \sqrt{[(E_s^a - E_p^c)/2]^2 + 16E_{sapc}^2}, \quad (1.18a)$$

$$E(X_{1c/v}) = \frac{1}{2}(E_s^a + E_p^c) \pm \sqrt{[(E_s^a - E_p^c)/2]^2 + 16E_{scpa}^2}, \quad (1.18b)$$

$$E(X_{5c/v}) = \frac{1}{2}(E_p^a + E_p^c) \pm \sqrt{[(E_p^a - E_p^c)/2]^2 + 16E_{xy}^2}, \quad (1.18c)$$

Equations (1.18a) can be inverted for E_{saPc} , E_{scpa} and E_{xy} in terms of the bandgap data at the X symmetry point with appropriate s and p energies. Finally at the L symmetry point ($k_x = k_y = 1, k_z = 0$) the conduction and valence band energies are

$$E(L_{3c/v}) = \frac{1}{2}(E_p^a + E_p^c) \pm \sqrt{[(E_p^a - E_p^c)/2]^2 + 4(E_{xx} + E_{yy})^2}, \quad (1.19)$$

Likewise L energy levels can be improved. It should be noted that fitting of (1.16)–(1.19) to the bandgap data [12, 13] and empirical pseudopotential bandgaps [14] at high symmetry points is reliable only for the valence band but not to the conduction bands; this is vital in studying the electronic structures of low dimensional heterostructures. In this respect we find it useful to use the semiempirical second nearest neighbor (2NN) sp^3 tight binding models [35, 36]. Talwar and Ting [35] incorporate the second nearest neighbor (2NN) interactions of cation and anion atoms in the semiempirical sp^3 tight binding formalism by constructing a nonzero symmetrized (8×8) Hamiltonian matrix with 23 two-center 2NN integrals. The (2NN) sp^3 TB Hamiltonian matrix elements are given as

$$H_{\alpha\beta} = \langle \varphi_\alpha(k) | H | \varphi_\beta(k) \rangle = E_{\alpha\beta} + \sum_{i \neq 0} I_{\alpha\beta}(0, i) e^{ik \cdot r_i} + H''_{\alpha\beta} + H_{so}, \quad (1.20)$$

where $E_{\alpha\beta}$ represents the intra-atomic integrals, coupling atomic orbitals located in the same cell. The second term $I_{\alpha\beta}(0, i)$ is the integral that represents the NN interactions, coupling atomic orbitals located in different cells. Third term represents 2NN interaction integrals. Finally, the last term represents the spin-orbit interactions. In this 2NN sp^3 TB approach there are 23 TB parameters to be determined from the fitting of the TB model energy bands to experimental data [12, 13] and empirical pseudopotential bands at symmetry points [14].

The modified virtual crystal approximation (MVCA) can be implemented in the semiempirical tight binding model [19–28] by using (1.7a–1.7d) and (1.8) for the composition variation of bond length and diagonal and off-diagonal elements in the TB Hamiltonian matrix for ternary constituent of heterostructures. This allows one to take into account the composition variations of the off-diagonal terms in the TB Hamiltonian matrix elements. Likewise, the TB parameters representing the diagonal terms in the Hamiltonian matrix for ABC ternary semiconductor can be taken nonlinear function of composition using (1.8). We can then use the semiempirical NN or 2NN sp^3 TB model calculate the composition effects on electronic properties of ternary/binary heterostructures by taking account of lattice mismatch induced

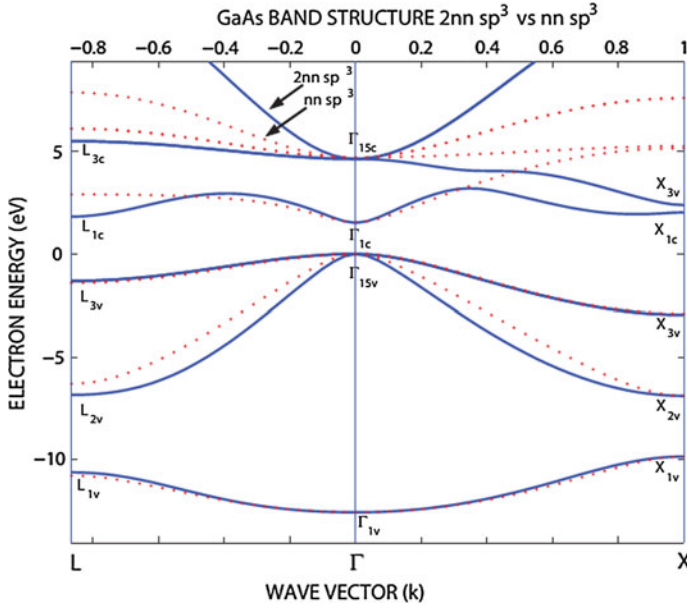


Fig. 1.10 Band structure of GaAs obtained using 2NN sp^3 TB model of Talwar and Ting [35] (dashed lines) and 2NN sp^3 TB model of Loehr and Talwar [36] (solid lines)

interface strain that causes modification of the energy levels of the heterostructure constituents.

Figure 1.10 shows the band diagram of GaAs obtained using the NN and 2NN sp^3 TB model due to the parametrization of Talwar and Ting [35] and Loehr and Talwar [36]. As shown in Fig. 1.10, improvements made in the band structure calculations by the 2NN sp^3 TB model of Loehr and Talwar is visible when it is compared with that of Talwar and Ting at varying values of wave vector over the entire first Brillouin zone, although both produce the experimental bandgap data and empirical pseudopotential bands at symmetry points. However, the low dimensional device performance modeling requires accurate numerical values for the electron and hole masses since they determine the charge transport and device performance. In aiming for this goal, Loehr and Talwar inverted the expressions for energy levels at the Γ , L and X high symmetry point, fitting the electron and hole effective masses, to reduce the number of free tight binding parameters from 23 to 8. The 2NN sp^3 TB parametrization of Loehr and Talwar yields much better conduction band dispersion curves than that of Talwar and Ting. The optimized 2NN sp^3 TB parameters of Loehr and Talwar reproduce the electron mass at the lowest conduction valley and heavy hole effective mass of the valence band at Γ point, in addition to its ability to reproduce the correct values of the critical point energies at the Γ , L and X high symmetry points.

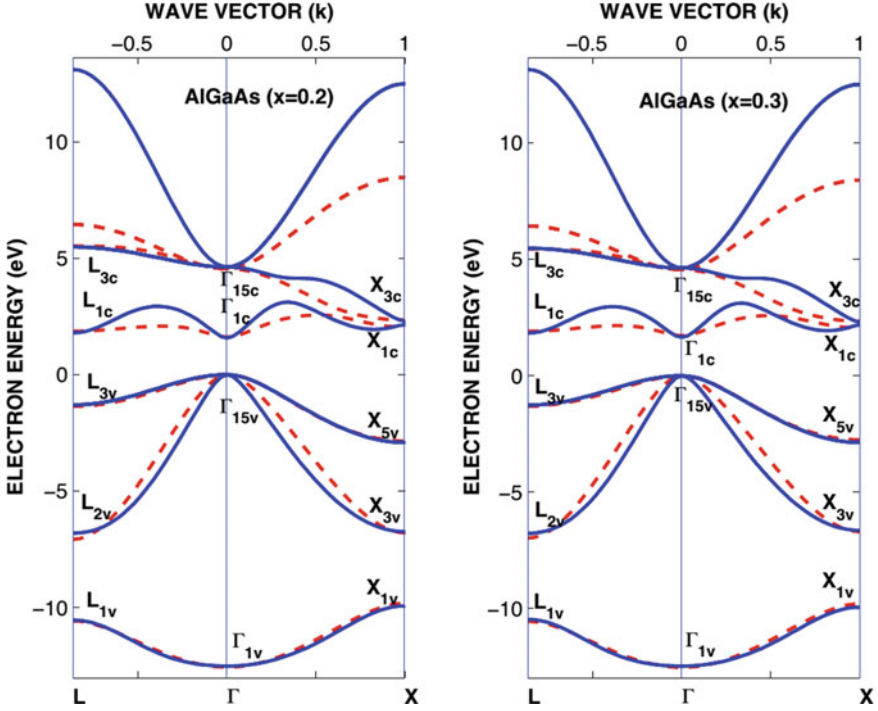


Fig. 1.11 Band structure of AlGaAs with $x = 0.20$ (left) and $x = 0.30$ (right) calculated with the 2NN sp^3 TB parameters of Talwar and Ting [23] (dashed lines) and Loehr and Talwar [24] (solid lines)

Using the tight binding parameters of Talwar and Ting [35] and Loehr and Talwar [36] we calculated the composition effects on band structure of AlGaAs ternary semiconductor and the results are displayed in Fig. 1.11 for $x = 20$ and 30% of aluminum fraction. Furthermore, the composition variation of band gaps of AlGaAs ternary in lattice matched AlGaAs/GaAs heterostructure is shown in Fig. 1.12 (left) and of InGaAs ternary in pseudomorphic InGaAs/GaAs heterostructures is shown in Fig. 1.12 (right), for which band gaps at 0 K are computed using the 2NN sp^3 TB model with the TB parameters of Talwar and Ting and Loehr and Talwar compared with experimental data and empirical pseudopotential model results at symmetry points. As can be seen from Fig. 1.12 (left), there is an excellent agreement between theory and experiment for lattice matched AlGaAs/GaAs heterostructures. However, it should be noted that there is a considerable lattice mismatch across many of the III–V and II–VI ternary/binary heterointerfaces (e.g., InGaAs/GaAs). Interface strain due to lattice mismatch causes a shift in the lattice constant of an epilayer: $a = (1 + \varepsilon)a_0$, where ε is the symmetric strain tensor. Therefore, the bond lengths, defined in (1.7a–1.7d), and TB matrix elements, defined in (1.8), will be modified with strain. The effect of

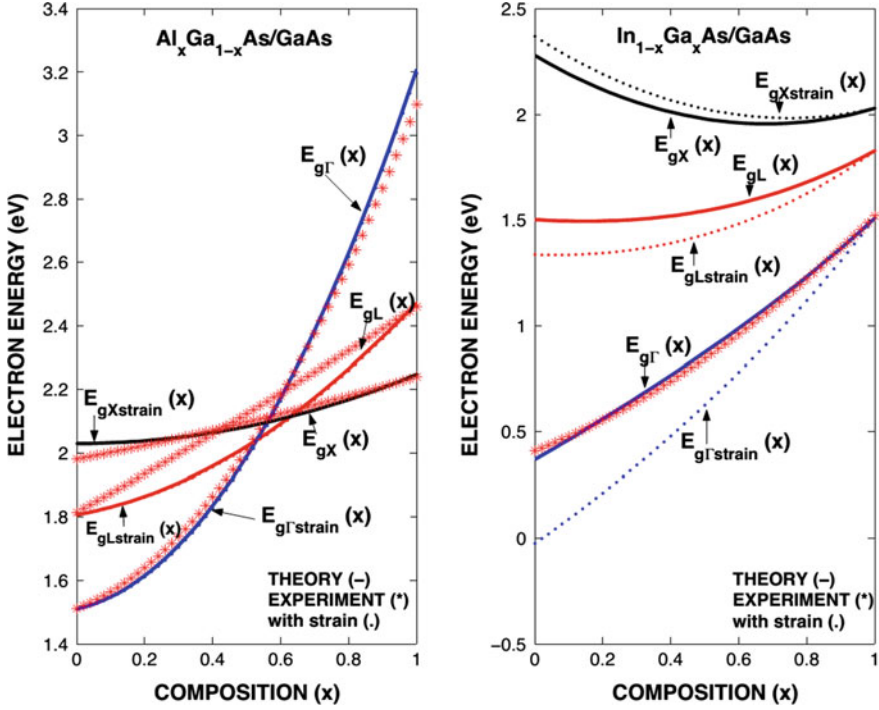


Fig. 1.12 Predicted strain effects on bandgaps of AlGaAs in lattice matched AlGaAs/GaAs (*left*) and of InGaAs in pseudomorphic InGaAs/GaAs (*right*) heterostructures

interface strain is evident in the case of InGaAs/GaAs heterostructure, as seen in Fig. 1.12 (right). There is a large strain shift in the band gaps of the InGaAs ternary constituent in an InGaAs/GaAs heterostructure at symmetry points due to large lattice mismatch that increases with indium mole fraction. There is good agreement between theory and experiment for the InGaAs ternary as bulk at around 0 K.

Qualitatively reliable and quantitatively accurate determination of the strain effects on the energy levels by fitting the off-site tight binding matrix elements to a set of observables is difficult. The off-site TB matrix elements representing nearest neighbor interactions, known as the hopping strength, will be modified with respect to their unstrained values and are often determined by assuming that they obey the Harrison Scaling Law [34]: $V_{ll'm}(\varepsilon) = V_{ll'm}(a/a_0)^{-\eta_{llm}}$, where $V_{ll'm}(\varepsilon)$ is the strained and $V_{ll'm}$ the unstrained value of interaction potential for anion and cation atoms. The exponents η_{llm} are determined to reproduce the strain variations of the band structure of relevant semiconductors under hydrostatic pressure, namely the volume deformation potential $a_{gl} = -B(\partial E_{gl}/\partial P)$ for the corresponding band gap energies $E_{g\Gamma}$, E_{gL} and E_{gX} at symmetry points. Since these bandgaps depend on experimental data, in order to obtain reliable and accurate TB parameters, one must go through a fitting process that depends on the mapping of a large number of orbital coupling

parameters on the set of observables; in many cases there are not many analytical expressions available. One can overcome this difficulty by using a so-called statistical thermodynamic model [10, 11], which considers the conduction electrons and valence holes as distinct electrically chemical species, to study the interface strain effects on the electronic structure of heterostructures at symmetry points, defined by (1.5) and (1.6). In this model, one first expresses the shifts in the conduction and valence band edges at Γ , L and X high symmetry points obtained by using the 2NN sp^3 TB orbitals basis sets as a function of pressure at any lattice temperature and then obtains the refined tight binding matrix elements to find the band structure.

Although the Slater-Koster type semiempirical 2NN sp^3 s tight binding treatment of electronic structures yields a good description of valence band dispersion curves, the conduction band dispersion curves are inaccurately given, especially the indirect band gap at the X symmetry point is not well reproduced. Vogl et al. [34] introduced a nearest neighbour sp^3s^* tight binding model in order to include the influence of excited d-states, which will be discussed next.

1.3.2 Semiempirical Sp^3s^* Tight Binding Modeling

In the semiempirical nearest neighbor sp^3s^* tight binding model of Vogl et al. [41], each atom is described by its outer valence s orbital, three p orbitals and fictitious excited s^* orbital is added to mimic the effects of higher lying d-states. This results in a semiempirical NN sp^3s^* tight binding model with a total of 13 parameters, which are determined by comparing the predicted bandgaps with those produced by empirical pseudopotential model [14] at high symmetry points. Adding the excited s^* state and spin-orbit coupling to sp^3 orbital basis set makes it possible to accurately calculate band dispersion curves at the X high symmetry point. Furthermore, the inclusion of 2NN interactions of cation and anion atoms in the sp^3s^* TB model yields better fit of conduction band dispersion curve at the L symmetry point.

The sub-matrices (diagonal (H_{cc} and H_{aa}) and off-diagonal (H_{ca} and $H_{ac} = H_{ca}^*$) elements of the 2NN sp^3s^* Hamiltonian matrix $H_{\alpha\beta}$ are written as [41]

$$H_{cc} = \begin{bmatrix} E_s^c & -\varepsilon(s, p_x)B_6 & -\varepsilon(s, p_x)B_5 & -\varepsilon(s, p_x)B_4 & 0 \\ -\varepsilon(s, p_x)B_6 & E_p^c & -\varepsilon(p_x, p_y)B_4 & -\varepsilon(p_x, p_y)B_5 & 0 \\ -\varepsilon(s, p_x)B_5 & -\varepsilon(p_x, p_y)B_4 & E_p^c & -\varepsilon(p_x, p_y)B_6 & 0 \\ -\varepsilon(s, p_x)B_4 & -\varepsilon(p_x, p_y)B_5 & -\varepsilon(p_x, p_y)B_6 & E_p^c & 0 \\ 0 & 0 & 0 & 0 & E_{s^*}^c \end{bmatrix}, \quad (1.21a)$$

$$H_{aa} = \begin{bmatrix} E_s^a & \varepsilon(s, p_x)B_6 & \varepsilon(s, p_x)B_5 & \varepsilon(s, p_x)B_4 & 0 \\ \varepsilon(s, p_x)B_6 & E_p^a & \varepsilon(p_x, p_y)B_4 & \varepsilon(p_x, p_y)B_5 & 0 \\ \varepsilon(s, p_x)B_5 & \varepsilon(p_x, p_y)B_4 & E_p^a & \varepsilon(p_x, p_y)B_6 & 0 \\ \varepsilon(s, p_x)B_4 & \varepsilon(p_x, p_y)B_5 & \varepsilon(p_x, p_y)B_6 & E_p^a & 0 \\ 0 & 0 & 0 & 0 & E_{s^*}^a \end{bmatrix}, \quad (1.21b)$$

$$H_{ca} = \begin{bmatrix} B_0E_{ss} & B_1E_{sp} & B_2E_{sp} & B_3E_{sp} & 0 \\ -B_1E_{sp} & B_0E_{xx} & B_3E_{xy} & B_2E_{xy} & -B_1E_{ps^*} \\ -B_2E_{sp} & B_3E_{xy} & B_0E_{xx} & B_1E_{xy} & -B_2E_{ps^*} \\ -B_3E_{sp} & B_2E_{xy} & B_1E_{xy} & B_0E_{xx} & -B_3E_{ps^*} \\ 0 & B_1E_{s^*p} & B_2E_{s^*p} & B_3E_{s^*p} & 0 \end{bmatrix}, H_{ac} = (H_{ca})^*, \quad (1.21c)$$

where $E_s^a, E_s^c, E_p^a, E_p^c, E_{s^*}^a$, and $E_{s^*}^c$ are diagonal elements of $H_{\alpha\beta}$ matrix and represent the on-site atomic energies of cation and anion atoms. $E_{ss}, E_{xx}, E_{s_a p_c}, E_{s_c p_a}, E_{xy}, E_{s^* p}$ and E_{ps^*} are the off-diagonal elements of $H_{\alpha\beta}$ matrix and represent the hopping terms (transfer matrix elements). Finally, $\varepsilon_{sx} = \varepsilon(sc(a), p_x c(a))$ and $\varepsilon_{xy} = \varepsilon(p_x c(a), p_y c(a))$ are the two 2NN transfer matrix elements for the cation and anion atoms. Here s and p refer to the basis states and a and c refer to anion (e.g., As, Sb, N) and cation (e.g., Al, Ga and In) atoms, respectively. In the 2NN sp^3s^* TB Hamiltonian matrix, B_i^* is the complex conjugate of the matrix element B_i which gives the $\mathbf{k}$ wave vector dependence and the first four elements are the same as in (1.15a) and other three are

$$B_4(\mathbf{k}) = 4\text{Sin}(\mathbf{k}_x a)\text{Sin}(\mathbf{k}_y a), \quad (1.22a)$$

$$B_5(\mathbf{k}) = 4\text{Sin}(\mathbf{k}_x a)\text{Sin}(\mathbf{k}_z a), \quad (1.22b)$$

$$B_6(\mathbf{k}) = 4\text{Sin}(\mathbf{k}_y a)\text{Sin}(\mathbf{k}_z a), \quad (1.22c)$$

where $i = \sqrt{-1}$ and displacement vectors of nearest neighbors. Including the spin-orbit coupling the size of the (10×10) 2NN sp^3s^* TB Hamiltonian matrix increases to (20×20) matrix which is diagonalized for each $\mathbf{k}$ vector to obtain the band structure [42]. The spin-orbit effects are included with the 2NN sp^3s^* orbitals set by coupling different spin states of different on-site p orbitals through the spin-orbit interaction. Tight binding parameters of 2NN sp^3s^* orbitals set is obtained by fitting the obtained band gaps given in Table 1.1 to those produced by empirical pseudopotential theory [14]. Having reliable diagonal matrix elements one can make realistic tight binding parametrization of the off-diagonal matrix elements representing the first nearest neighbor (NN) and/or 2NN interactions, which are obtained by focusing on the reproduction of the band gap energies at symmetry points in the energy dispersion curve. As examples of 2NN sp^3s^* tight binding parametrization, the bandgaps given in Table 1.1 [21, 23] for several group III–V compounds may be used to calculate electronic band structure of their heterostructures.

Table 1.1 Bandgaps at symmetry points of AlAs, GaAs, InAs, GaP, AlN, GaN and InN compounds obtained using the 2NN sp^3s^* TB model [21, 23] and experimental data [12] are given in parenthesis

(eV)	GaAs	InAs	GaP	AlN	GaN	InN
$E_{g\Gamma}$	1.52 (1.52)	0.43 (0.42)	2.88 (2.88)	5.99 (6.00)	3.30 (3.27)	0.90 (0.90)
E_{gX}	1.98 (2.03)	2.28 (2.50)	2.33 (2.16)	4.90 (4.80)	4.70 (4.70)	2.83 (3.00)
E_{gL}	1.82 (1.85)	1.61 (1.43)	2.56 (2.79)	8.63	6.10	3.81

Table 1.2 2NN sp^3s^* TB parameters for GaAs, InAs, GaP, AlN, GaN and InN obtained using Table 1.1

(eV)	GaAs	InAs	GaP	AlN	GaN	InN
$E_{s,a}$	-8.4399	-9.5381	-8.1124	-11.505	-12.915	-12.860
$E_{p,a}$	0.9252	0.7733	1.0952	4.3815	3.1697	1.9800
$E_{s,c}$	-2.6569	-2.7219	-2.1976	0.5047	-1.5844	-0.3994
$E_{p,c}$	3.5523	3.5834	4.0851	10.2184	9.0302	8.0200
$E_{s^*,a}$	6.6235	7.2730	8.4796	12.0400	12.2000	10.6300
$E_{s^*,c}$	7.4249	6.6095	7.1563	13.7400	12.2000	13.0000
$4V_{s,s}$	-6.4210	-5.6052	-7.4909	-9.8077	-8.8996	-4.2285
$4V_{x,x}$	1.9850	1.8398	2.1516	6.6900	5.3500	3.9800
$4V_{x,y}$	4.9100	4.3977	5.1213	8.9400	8.6200	7.4100
$4V_{sa,pc}$	4.2390	3.0205	4.2724	7.8500	6.4000	3.8100
$4V_{pa,sc}$	5.15358	5.3894	6.3075	7.6800	7.2400	6.1900
$4V_{s^*a,pc}$	3.80624	3.2191	4.8184	8.0300	7.0600	6.8800
$4V_{pa,s^*c}$	4.7009	3.7234	50534	2.4700	1.8200	3.3600
ε_{xx}	0.2459	0.1441	0.2325	-1.4000	0.9500	0.6150
ε_{xy}	-0.1050	0.0249	-0.2200	6.9000	1.0100	0.7100
λ_a	0.0553	0.1385	0.0578	0.0035	0.0035	0.0035
λ_c	0.1338	0.1290	0.0222	0.0070	0.0410	0.1100

As an example of 2NN sp^3s^* TB parametrization, the diagonal and off-diagonal tight binding parameters for AlAs, GaAs, InAs, GaP, AlN, GaN and InN are given in Table 1.2 [21, 23].

Using the 2NN sp^3s^* TB model the electronic band structures of GaAs, GaN, InN and AlN are calculated [21, 23] and are shown in Fig. 1.13, which reproduce the conduction and valence band structures, including the heavy hole and light hole bands as well as spin-orbit splitting bands. Tight binding interaction parameters $4V_{s^*,p}$ and $4V_{p,s^*}$ were adjusted to fit to the X bands and 2NN interaction parameters ε_{xx} and ε_{xy} to get a good fit to the L bands in reproducing the empirical pseudopotential energy bands [14]. As shown in Fig. 1.13, adding the excited s^* state to the sp^3 orbitals set on the cation and anion atoms with 2NN interactions and

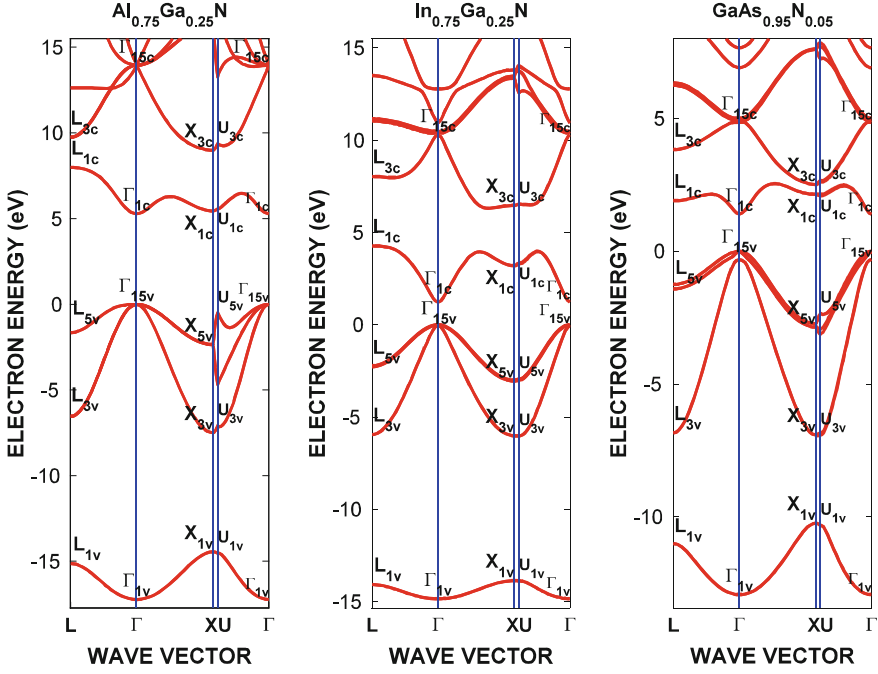


Fig. 1.14 Band structures of AlGaIn, InGaIn and GaAsN obtained using the 2NN sp^3s^* TB orbitals set with optimized energies given in Table 1.2

Predicted fundamental bandgaps of AlGaIn, InGaIn and GaAsN ternaries at the Γ , L and X points, especially at Γ , are in excellent agreement with experiment [12]. The principal band gaps of AlGaIn and InGaIn conventional nitrides increase with composition. However, the band gap of GaAsN diluted nitride at the Γ symmetry point decreases with alloy composition for $x < 0.25$, and then increases for $0.26 < x < 1$, in agreement with experiment. It is gratifying to note that using the optimized 2NN sp^3s^* tight binding model parameters for GaAs, GaN, AlN and InN compounds given in Tables 1.1 and 1.2, we are able to obtain the nonlinear composition dependence of the principal band gaps of AlGaIn, InGaIn and GaAsN ternary semiconductors without any empirical fitting or any adjustable parameter. This observation suggests that by using the optimized 2NN sp^3s^* TB parameters for bulk GaAs, GaN, AlN and InN binary compounds, given in Table 1.2, the 2NN sp^3s^* TB model allows one to determine the nonlinear composition dependence of the principal band gaps of nitride-based ternary semiconductors without any empirical fitting nor any adjustable parameter.

Cadmium and zinc based group II–VI compounds (CdX and ZnX ($X = S, Se, Te$)) and their ternary/binary heterostructures are known to have considerable potential for making optical and photovoltaic devices (e.g., light emitting diodes, laser diodes, infrared detectors, photovoltaic devices, and quantum dots [1, 2]). Therefore, similar calculations were carried out for the electronic band structure of

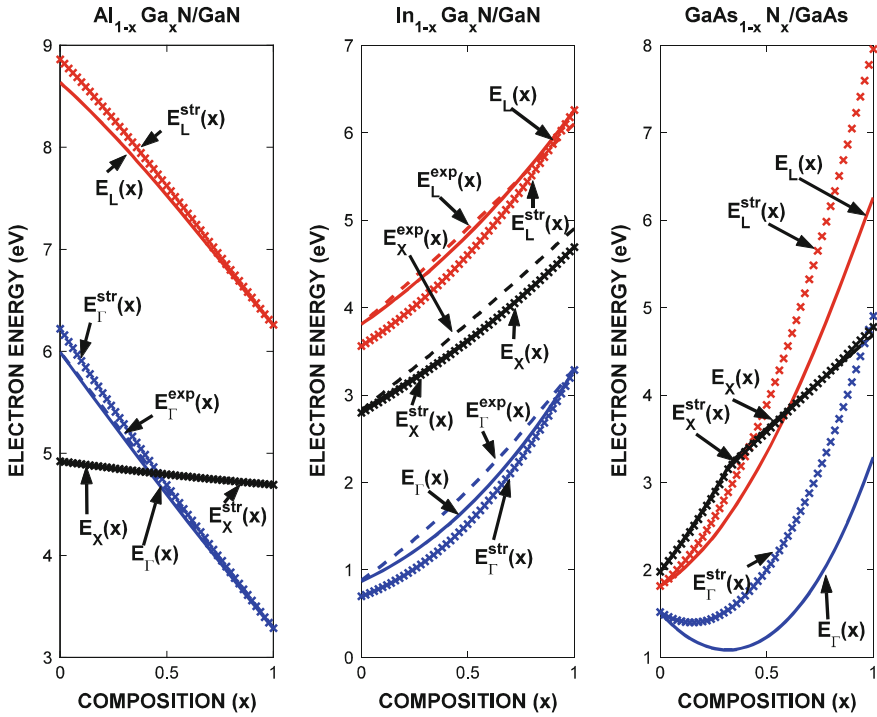


Fig. 1.15 Predicted interface strain effects on the bandgaps of ternaries in AlGaIn/GaN (left), InGaIn/GaN (middle) and GaAsN/GaAs (right) heterostructures compared with experiment [12]

cadmium and zinc based group II–VI compounds (CdX and ZnX ($X = S, Se, Te$)) and their $ZnCdS/CdS$ (left) and $ZnTeS/ZnS$ ternary/binary heterostructures by using the 2NN sp^3s^* TB parametrization, which yield the same accuracy for the conduction band structure of CdS , $CdSe$, $CdTe$, ZnS , $ZnSe$ and $ZnTe$. Figure 1.16 shows the composition and interface strain effects on band gaps of ternaries in $ZnCdS/CdS$ (left) and $ZnTeS/ZnS$ (right) heterostructures compared with experiment [12]. It should be noted that there is a considerable lattice mismatch across many of the ternary/binary heterointerfaces. Therefore, the composition dependent bond lengths, defined in (1.7a–1.7d), and tight binding matrix elements in the semiempirical NN sp^3s^* orbitals set, defined in (1.24), and hence the electronic properties of heterostructure alloy constituents will be modified with respect to their unstrained values. Furthermore, accurate description of the second conduction band and the transverse effective masses at the X- and L-symmetry points are found to be in good agreement with experiment, leading to a reliable and accurate TB model calculation of electronic properties of semiconductors. We should note that the accuracy of 2NN sp^3s^* TB parametrization depends on the accurate description of band structures of III–V and II–VI compounds by the empirical pseudopotential theory and bandgap data.

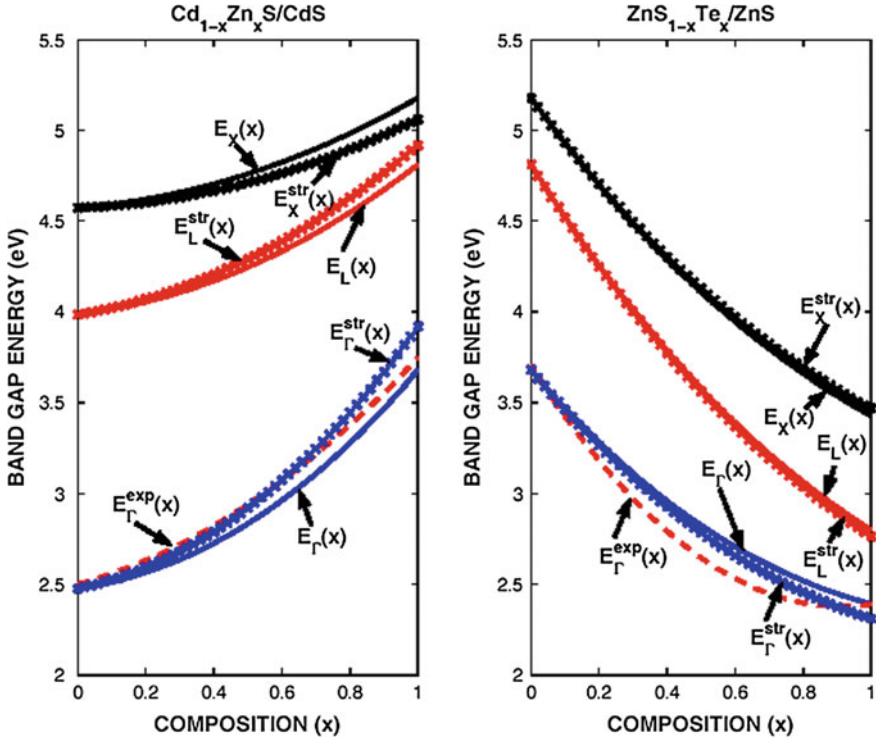


Fig. 1.16 Predicted interface strain effects on band gaps of ternaries in ZnCdS/CdS (*left*) and ZnTeS/ZnS (*right*) heterostructures compared with experiment [12]

Although the inclusion of s^* excited state in the 2NN sp^3 TB model yields accurate calculations of the conduction band structure features at the X high symmetry point, it does so by modelling the average of p-d interactions and is insufficient to determine energy levels above 6.0 eV. Jancu et al. [42, 43] suggested that from a solid state physics point of view, the actual behaviour of excited d-states is not reliably addressed in the sp^3s^* tight binding model of valence band and conduction band dispersion curves involving the Γ , X and L high symmetry points, which will be discussed next.

1.3.3 Semiempirical $Sp^3d^5s^*$ Tight Binding Modeling

In the semiempirical nearest neighbor $sp^3d^5s^*$ tight binding model proposed by Jancu et al. [42, 43], the excited d-states contribute critically to both the valence band maximum at the Γ symmetry point and to the conduction band dispersion curves at the X and L symmetry points and is explicitly included in NN $sp^3d^5s^*$ orbitals set. The band gaps at high symmetry points in the Brillouin zone and

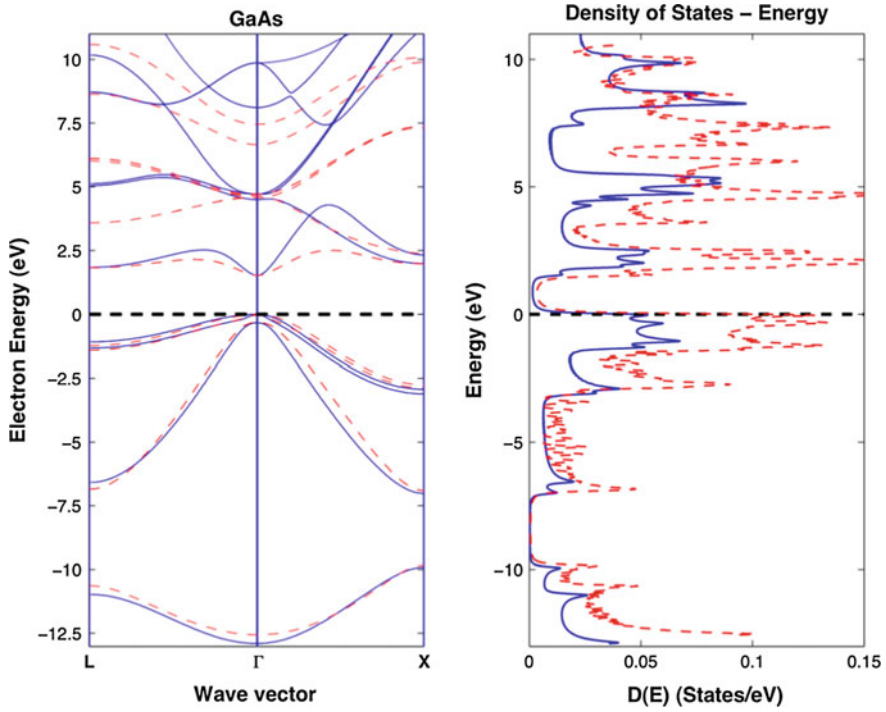


Fig. 1.17 Band structure and density of states of a GaAs compound calculated using the 2NN sp^3s^* (dashed) and NN $sp^3d^5s^*$ (solid) tight binding models

optimized tight binding parameters are given in Tables 1.1 and 1.2 for the 2NN sp^3s^* 2NN TB model and those of the NN $sp^3d^5s^*$ TB model were taken from Jancu et al. [42]. Results of 2NN sp^3s^* TB and NN $sp^3d^5s^*$ TB models are compared in Fig. 1.17 for electronic band structure and density of states of GaAs. As can be seen from Fig. 1.17, both the valence band and conduction band dispersion curves of GaAs obtained by using the NN $sp^3d^5s^*$ TB model overcomes most of the limitations of the earlier 2NN sp^3 and sp^3s^* TB models. As Fig. 1.17 (left) for GaAs show, the NN $sp^3d^5s^*$ TB model overcomes most of the limitations of the earlier of 2NN sp^3 and sp^3s^* 2NN TB models for both valence and conduction band dispersion curves. Furthermore, accurate description of the second conduction band and the transverse effective masses at the X- and L-symmetry points that are in good agreement with experiment suggests that the NN $sp^3d^5s^*$ TB parametrization makes it possible to accurately calculate the electronic and optical properties involving symmetry points at the edge of the Brillouin zone of tetrahedral semiconductors. It is noted here that the NN $sp^3d^5s^*$ TB model is reliable for the calculation of dispersion curves for energies up to 6 eV above the valence band maximum, correctly reproducing the orbital character of band edges, and their

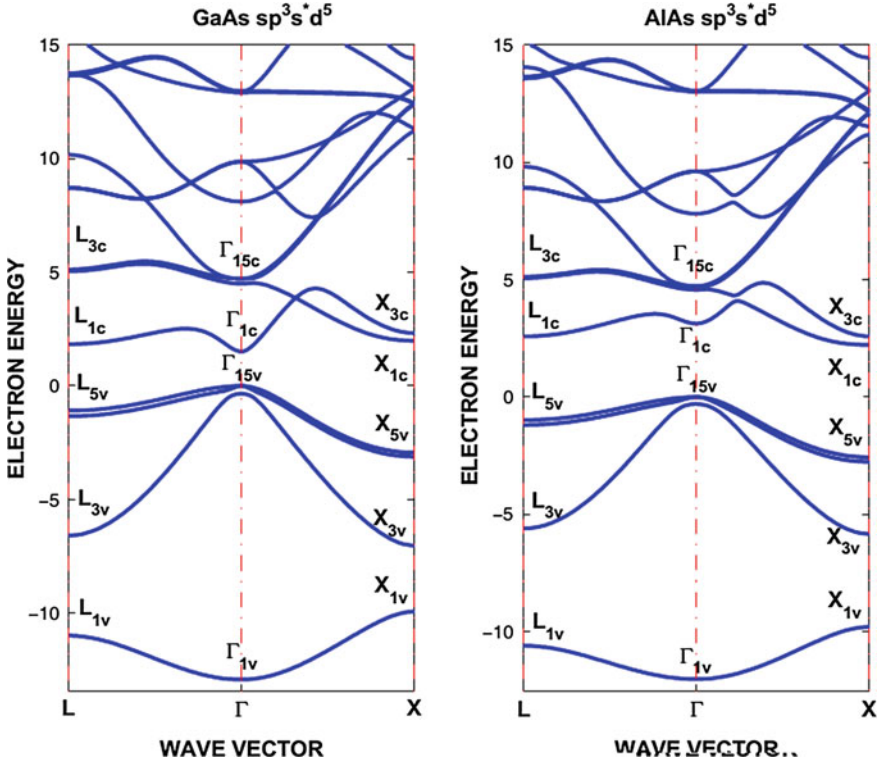


Fig. 1.18 Electronic band structure of AlAs and GaAs obtained using NN $sp^3d^5s^*$ TB model

behavior under strain, at the Γ , X and L symmetry points of the first Brillouin zone of compound semiconductors.

In a search for physically realistic and numerically accurate calculations of band structures of group III–V and II–VI compound semiconductors, Sapra et al. [44] have argued that the adding of an s^* state to an sp^3 orbitals set does not represent the true contribution of d-states to both valence band and conduction band dispersion curves and have shown that the NN sp^3d^5 tight binding model is sufficient to calculate the electronic structure of group III–V and II–VI compounds and suggests that the use of any fictitious s^* state is not needed in tight binding modeling to calculate the energy band structures. However, as can be seen from Fig. 1.18, the use of the more complete nearest neighbor $sp^3d^5s^*$ orbitals set imposes drastically heavier computational requirement on the calculation of electronic band structures of III–V and II–VI compound semiconductors, without affording realistic physical insight to the solid state interpretation of its contribution to the problem.

In the following section we use a semiempirical NN sp^3 tight binding parametrization of the problems of nearest neighbor interactions and spin-orbit

Table 1.3 Bandgaps of CdS, CdSe and ZnS obtained from fitting the NN sp^3d^5 and 2NN sp^3s^* TB models results [24] in the empirical pseudopotential bands [14] and measured band gaps [12, 13] at high symmetry points

Bandgap (eV)	CdS		ZnS		CdSe	
	sp^3s^*	sp^3d^5	sp^3s^*	sp^3d^5	sp^3s^*	sp^3d^5
E_g^T	2.476	2.555	3.680	3.702	1.887	1.937
E_g^X	4.341	4.696	5.103	5.190	3.784	3.779
E_g^L	3.983	5.193	4.810	4.641	3.097	3.027

coupling of p-states taken into account in the calculations of the band structures of III–V and II–VI compound semiconductors.

1.3.4 Semiempirical Sp^3d^5 Tight Binding Modeling

In this section we discuss the semiempirical sp^3d^5 tight binding parametrization with the first nearest neighbor interaction and spin-orbit coupling of p-states taken into account in the calculation of the electronic structure of constituents of semiconductor heterostructures. We write the semiempirical NN sp^3d^5 TB Hamiltonian matrix as [24]

$$H_{\alpha\beta} = \langle \varphi_\alpha(k) | H | \varphi_\beta(k) \rangle = E_{\alpha\beta} + \sum_{i \neq 0} I_{\alpha\beta}(0, i) e^{ik \cdot r_i} + H_{so}, \quad (1.23)$$

where $E_{\alpha\beta}$ is the on-site energy for the β orbital (s, p, d) at the atomic site α (cation and anion); and it represents the intra-atomic integrals, which couple atomic orbitals located in the same cell, and $I_{\alpha\beta}(0, i)$ represents the first nearest neighbor (NN) interaction integrals, known as the hopping term, which couples atomic orbitals located in different cells.

The nine state atomic-like ($s; x, y, z; xy, yz, zx, x^2 - y^2, 3z^2 - r^2$) basis set describes each atom of the semiconductor. $H_{\alpha\beta}$ has 19 independent matrix elements: 9 on-site and 10 off-site elements. The inclusion of spin-orbit coupling to the sp^3d^5 basis set, in which the spin-orbit interaction is given by two parameters; $\lambda_a = \langle x_a \uparrow | H_{so} | z_a \downarrow \rangle$ for anion and $\lambda_c = \langle x_c \uparrow | H_{so} | z_c \downarrow \rangle$ for cation atom, adds two extra tight binding parameters. Therefore, $H_{\alpha\beta}$ has total of 21 independent matrix elements, including the nearest neighbor interaction and spin orbit coupling of p-states. The diagonal and off-diagonal matrix elements are determined by fitting the empirical pseudopotential energy band dispersions to calculated energy band dispersions using the semiempirical NN sp^3d^5 TB parametrization for a given semiconductor. One starts with calculating the electronic band structure by estimating the values of on-site matrix elements and off-site matrix elements and then carries out a least-squares error minimization fitting procedure at a number of high

Table 1.4 Optimized s, p and s* TB parameters for CdS, CdSe and ZnS obtained from fitting the 2NN sp^3s^* TB model [24] to empirical pseudopotential band [14] at symmetry points

(eV)	CdS	ZnS	CdSe
$E_{s,a}$	-11.5300	-11.6100	-9.63
$E_{p,a}$	0.5300	1.4800	1.326
$E_{s,c}$	1.8300	1.1100	0.03
$E_{p,c}$	5.8700	6.5200	4.73
$E_{s^*,a}$	7.1300	8.0800	7.53
$E_{s^*,c}$	6.8700	8.0200	5.72
$4V_{s,s}$	-3.0700	-6.3000	-4.64
$4V_{x,x}$	1.7600	3.1100	2.64
$4V_{x,y}$	4.2300	5.0000	5.36
$4V_{s,p}$	2.1700	5.1600	4.57
$4V_{p,s}$	5.4800	5.1700	5.54
$4V_{s^*,p}$	1.9900	2.8900	3.05
$4V_{p,s^*}$	3.0600	1.7500	2.49
ε_{xx}	0.1000	0.2000	0.0
ε_{xy}	-0.0100	-0.1500	0.0
λ_a	0.0250 [6]	0.0250 [6]	0.1434
λ_c	0.0130 [14]	0.0270 [14]	0.0591

symmetry points in the valence and conduction band dispersion curves to fit band gap energies obtained from the empirical pseudopotential method [14] (Table 1.3).

We now compare the predictions of semiempirical NN sp^3d^5 and 2NN sp^3s^* TB parametrizations, both of which include spin-orbit coupling of p-states [28]. Tables 1.4 and 1.5 give the number of optimized TB parameters for the NN sp^3d^5 and 2NN sp^3s^* TB models. Figure 1.19 compare the electronic band structure of CdS, ZnS and CdSe calculated using NN sp^3d^5 and 2NN sp^3s^* TB models, both reproducing the band gaps as well as the valence band and conduction band dispersion curves at Γ , X and L symmetry points, which cannot be done with conventional NN sp^3 and 2NN sp^3 TB models.

As can be seen from Fig. 1.19 the tight binding model with an sp^3d^5 orbital basis, including the nearest neighbor interactions of cation and anion atoms and spin-orbit coupling of p-states, is adequate to accurately reproduce the band gaps at Γ , X and L symmetry points and both the valence band and conduction band dispersion curves. When compared with the 2NN sp^3s^* TB model, the NN sp^3d^5 TB model better simulates both the valence band and conduction band dispersion curves. This is due to the fact that the inclusion of s*-excited state is included by modeling the average of p-d interactions, and it does not permit the inclusion of excited d-orbitals to a sp^3 orbitals set. Therefore, the actual behavior and contribution of the excited d-states in the band structure calculations is not reliably reproduced, suggesting that the sp^3s^* TB model is of limited value for determination of optical properties of semiconductors involving high symmetry points. We should point out that both TB models with NN sp^3d^5 and 2NN sp^3s^* orbitals sets are

Table 1.5 Optimized s, p and d-TB parameters for CdS, CdSe and ZnS compounds obtained from fitting the NN sp^3d^5 TB model [24] to empirical pseudopotential bands at symmetry points [14]

(eV)	CdS	ZnS	CdSe
s_c	1.8300	1.3600	0.0300
p_c	5.8700	6.5200	4.7300
$d_c(t_2)$	-6.8300	-5.8200	-7.3100
$d_c(e)$	-7.4400	-6.2100	-7.8100
s_a	-11.5300	-14.6100	-9.6300
p_a	0.9300	1.7800	1.3260
$d_a(t_2)$	14.4300	15.5400	15.2600
$d_a(e)$	13.1500	13.6000	14.1000
$ss\sigma$	-0.8440	-1.0050	-1.2600
$sp\sigma$	2.2729	2.5387	2.3989
$pp\sigma$	1.68963	1.3343	1.9789
$pp\pi$	2.5500	4.3775	3.3400
$ps\sigma$	-0.10750	-0.3725	-0.6800
$ds\sigma$	0.0000	-2.5900	-1.2200
$dp\sigma$	-1.1400	-0.0500	-0.0100
$dp\pi$	-1.2900	-2.7800	-1.0900
$sd\sigma$	1.7500	1.3700	1.5200
$pd\sigma$	1.5580	2.1100	1.7800
$pd\pi$	-0.3500	-0.4500	-0.3200
λ_a	0.0250	0.0250	0.1434
λ_c	0.07600	0.0270	0.0591

semiempirical and their basis lies in a good description of the electronic band structures of tetrahedral semiconductors.

1.4 Density Functional Theory Modelling of Heterostructures

The conventional density functional theory (DFT) enables us to have a parameter-free description of structural and electronic properties of semiconductors and has its foundations in the work of Kohn and his collaborators [30, 45]. The variation method is used to calculate the ground state properties of a many-body system with the charge density, expressed in terms of single-particle electronic orbitals

$$n(\vec{r}) = \sum_{i=1}^N \sum_k f_{i,k} |\varphi_{i,k}(\vec{r})|^2, \quad (1.24)$$

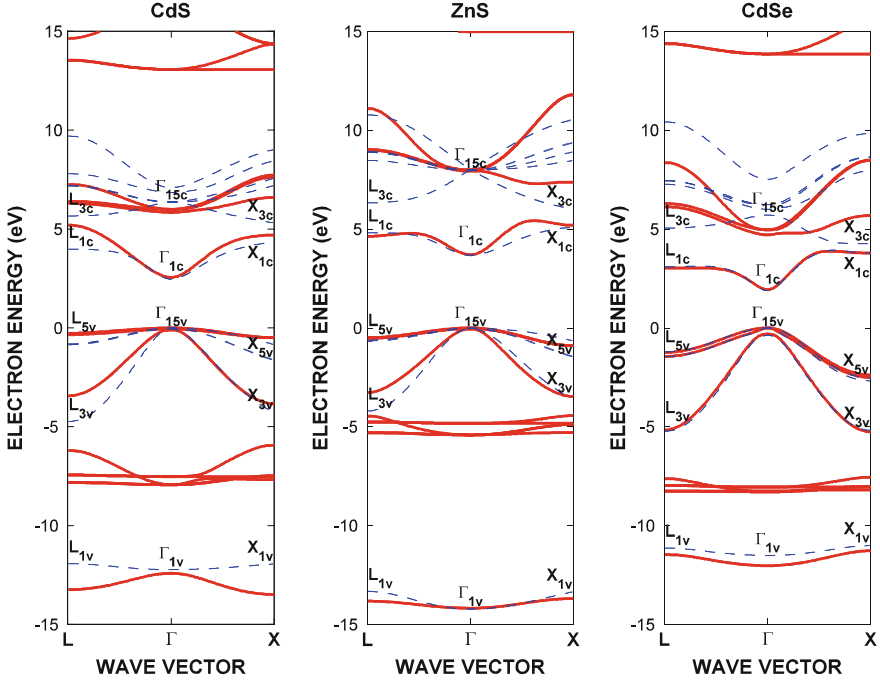


Fig. 1.19 Band structure of CdS, CdSe and ZnS compounds obtained using the 2NN sp^3s^* and NN sp^3d^5 TB models with the tight binding parameters given in Tables 1.4 and 1.5

which plays the central role in the calculation of the structural and electronic properties of semiconductors. The major consequence of the Kohn-Hohenberg-Sham studies [30, 45] is that the single particle states $\varphi_{i,k}(\mathbf{r})$ are determined by the solution of a set of equations similar to the Schrodinger equation:

$$H\varphi_{i,k} = \left(-\frac{\hbar^2}{2m} \nabla_i^2 + V_{\text{ext}} + V_{\text{eff}} \right) \varphi_{i,k} = E_i \varphi_{i,k}, \quad (1.25)$$

where $\varphi_{i,k}(\mathbf{r})$ are the one-electron wave functions, V_{ext} is the external potential of the nuclei, and $V_{\text{eff}}(\mathbf{r})$ the effective potential defined as

$$V_{\text{eff}}(\mathbf{r}) = V_H(\mathbf{r}) + V_{\text{xc}}(\mathbf{r}) = V_H(\mathbf{r}) + \frac{\delta E_{\text{xc}}(n(\mathbf{r}))}{\delta n(\mathbf{r})}, \quad (1.26)$$

where $V_H(\mathbf{r})$ is the Hartree Coulomb term and $E_{\text{xc}}(n(\mathbf{r}))$ is the exchange-correlation functional. Since the electron density $n(\mathbf{r})$ is involved in the definition of effective potential, (1.24) must be solved consistently with (1.25) and (1.26). Although the set of (1.24)–(1.26) yields, in principle, an exact solution to the quantum-mechanical problem, there are approximations involved in the

exchange-correlation potential $V_{xc}(r)$. In this respect, Local Density Approximation (LDA) and Generalized Gradient Approximations (GGA) have proven to be effective for a large number of semiconductors.

The conventional DFT calculations based on (1.24)–(1.26) provide satisfactory results for ground state properties (e.g., total energies, lattice constant, bulk modulus) but give unsatisfactory results for the electronic properties (e.g., band gaps and effective masses) [46]. For example, DFT predicted bandgaps are too small compared to experimental data [12] and predicted by empirical pseudopotential theory [14]. This difficulty can be avoided using various approaches (e.g., GW approximation or hybrid functional) that have been implemented in DFT model calculations. Even with GW and hybrid functionals, conventional DFT results in bandgap error on the order of 10–20 % as compared with experimental data [12]. As a remedy, the exchange-correlation potential contribution to the DFT band gap is shifted by using a so-called “scissor operator”, in accordance with the suggestion of Fiorentini and Balderschi [46]. The difference between the LDA and experimental band gaps (ΔE) scales with the optical dielectric constant ($\Delta E \cong 9.1/\epsilon_\infty$) [26, 27].

Recently, Tran and Blaha [47] proposed a new exchange-correlation potential, called the modified Becke-Johnson density functional (MBJLDA), which combines the Becke-Johnson exchange potential and the local density approximation (LDA) correlation potential in DFT band structure calculations. The MBJLDA functional is an exchange-correlation (XC) potential that is obtained as the functional derivative of the XC-energy functional E_{xc} with respect to the electron density $n(r)$ ($V_{xc}^{MBJ} = \delta E_{xc}[n](r)/\delta n(r)$) taken from LDA. The computational cost of DFT with the MBJLDA functional is comparable with DFT-LDA and DFT-GGA. It yields bandgaps which are in good agreement with experimental data [12, 13] and empirical pseudopotential theory [14]. We calculated the electronic properties of CdX and ZnX (X = S, Se, Te) II–VI compounds and their ternary/binary heterostructures using the WIEN2K [48] simulation package that uses the density functional theory (DFT) with MBJLDA functional in the band structure calculations.

Figure 1.20 exhibits the crystal structure of CdSe and ZnS compounds. CdX and ZnX (X = S, Se, Te) have a zinc-blende crystal structure (space group F-43 m No: 216) in which the Cd/Zn atoms are located at (0, 0, 0) and S (or Se) at (0.25, 0.25, 0.25). Converged results are obtained using 10,000 k points in the first Brillouin zone with $R_{MT}K_{max} = 8.50$, where R_{MT} represents the smallest muffin-tin radius and K_{max} is the maximum size of the reciprocal lattice vectors. LSDA, GGA and MBJLDA, as implemented in the WIEN2K simulations package, are used for exchange and correlation potentials in the calculations. SCF iterations are repeated until the total energy converges to a point less than 10^{-4} Ryd. The WIEN2K simulations with the MBJLDA functional embedded in DFT yields a remarkable improvement over LDA and GGA in the calculation of the structural properties (e.g., lattice constants and bulk modulus) and electronic structure properties (e.g., bandgaps) of the compounds (see Tables 1.6 and 1.7). However, we should point

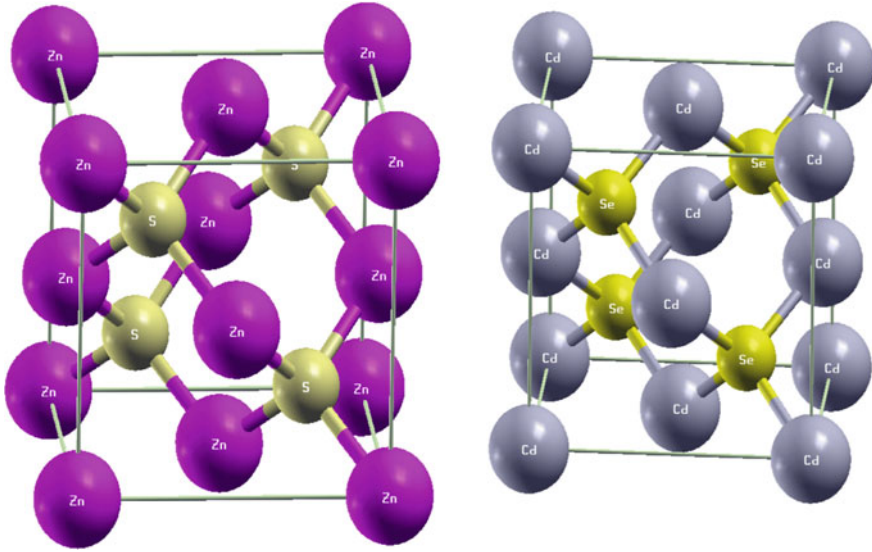


Fig. 1.20 Crystal structure of ZnS and CdSe compounds

Table 1.6 Lattice constant (a_0), bulk modulus (B_0) and its pressure derivative (B'_0) for CdX ($X = S, Se, Te$) and ZnX ($X = S, Se, Te$) calculated using WIEN2K, with DFT-MBJLDA

Property	CdS	CdSe	CdTe	ZnS	ZnSe	ZnTe
$a_0(A^0)$	5.779 ^a	5.994	6.425	5.470	5.540	5.985
	5.839 ^b	6.088	6.500	5.460	5.690	6.039
	5.810 ^c	6.084	6.480	5.320	5.669	6.030
	(5.830) ^d					6.089
B_0	68.765	53.799	47.778	71.474	82.379	54.038
	66.01	56.79	46.76	91.71	73.91	55.56
	64.30	55.53	44.50	77.3	62.50	50.90
		—	—	—	—	—
B'_0	4.375	2.364	5.200	3.550	3.646	4.381
		4.20	4.470		—	5.10
	4.31			5.00	5.00	—
					—	—

a DFT-MBJLDA, *b* 2NN sp^3s^* TBM, *c* NN sp^3d^5 TBM, (...) ^d: Measured lattice constant

out that the MBJLDA only underestimates the band gaps between -2.50 and -7% for CdTe, ZnS, ZnSe, ZnTe and overestimates up to $\sim 4\%$ for CdS and CdSe. In most cases (except ZnS), MBJLDA is in better agreement with experiment than GW and $G'W'$. Bandgap energies at L, Γ and X symmetry points, calculated by using WIEN2K simulations package, with the MBJLDA functional embedded in

Table 1.7 Comparison of calculated bandgap energies of CdX (X = S, Se, Te) and ZnX (X = S, Se, Te)

(eV)	CdS	CdSe	CdTe	ZnS	ZnSe	ZnTe
$E_{g\Gamma}$	2.660 ^a	1.89	1.56	3.66	2.67	2.22
	2.476 ^b	1.887	1.59	3.680	2.83	2.39
	2.555 ^c	1.937	0.76	3.702	1.91	1.53
	(2.55)	(1.82)	(1.60)	(3.82)	(2.82)	(2.39)
E_{gX}	4.28	3.40	2.36	4.48	3.61	2.45
	3.983	3.097	2.46	4.810	3.99	2.80
	5.193	3.027	1.83	4.641	2.14	1.62
E_{gL}	4.71	4.04	3.03	4.42	3.78	2.71
	4.341	3.784	3.57	5.103	4.54	3.43
	4.696	3.779	4.76	5.190	5.04	3.34

a DFT-MBJLDA, *b* 2NN sp^3s^* TBM, NN sp^3d^5 TBM, (...): Measured fundamental bandgap

DFT, and using NN sp^3d^5 and 2NN sp^3s^* TB models, are given in Tables 1.6 and 1.7 for CdS, CdSe, CdTe, ZnS, ZnSe and ZnTe II–VI compounds.

As shown in Table 1.6, MBJLDA functional based DFT calculations of structural properties of compound semiconductors (such as lattice constant, bulk modulus and its pressure derivative) are in good agreement with the experimental data [12]. Furthermore, as shown in Table 1.7, WIEN2K simulations with MBJLDA embedded in DFT are in good agreement with the experimental data in terms of band gaps. Figures 1.21 and 1.22 demonstrate use of MBJLDA functional in WIEN2K simulations package in calculating the electronic band structure of compound semiconductors yields much better results than those of LSDA for CdX (X = S, Se, Te) and GGA band structures for ZnX (X = S, Se, Te). Accurate modeling of band structure properties have a key role in the design of cadmium and zinc chalcogenides based device applications. According to current calculations MBJLDA performed better than other conventional DFT functionals and GW approximations to calculate band structure. For this reason, MBJLDA is a very effective method and it can be used for a wide range of semiconductors. Accurate modelling of band structure has a key role in the design of cadmium and zinc chalcogenides-based device applications.

The comparison of the band structure dispersion curves obtained by using the WIEN2K simulation package with MBJLDA functional embedded in DFT and the semiempirical tight binding theories with NN sp^3d^5 and 2NN sp^3s^* orbitals sets is shown in Fig. 1.23 for CdS and ZnS and in Fig. 1.24 for CdSe, CdTe and ZnSe compound semiconductors, respectively. As displayed in Figs. 1.23 and 1.24, the MBJLDA functional based DFT and semiempirical NN sp^3d^5 and 2NN sp^3s^* orbitals sets TB models accurately reproduce the band gaps and band dispersion curves at high symmetry points when they are compared with empirical pseudopotential calculations for bulk II–VI compounds and their heterostructures. The three different approaches almost equally well reproduce the band gaps at Γ , X and L high symmetry points and valence band and conduction band dispersion curves in

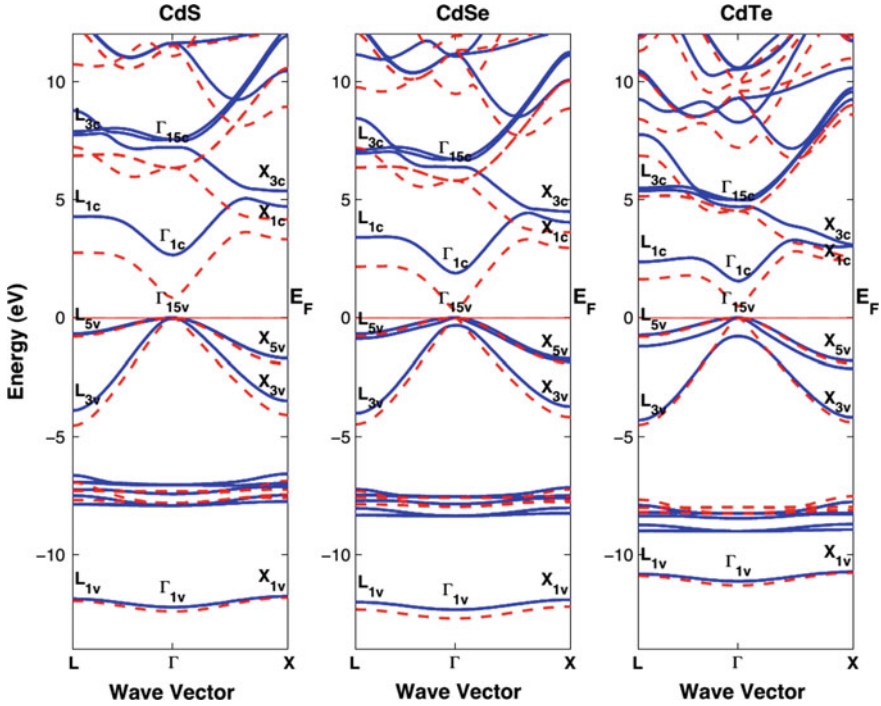


Fig. 1.21 Band structure of CdS, CdSe and ZnS obtained using WIEN2K with DFT-MBJLDA (solid lines) and DFT-LSDA (dashed lines)

the entire Brillouin zone of bulk CdX and ZnX (X = S, Se, Te) II–VI compounds and their heterostructures CdZnS/CdS, CdSe/CdTe and ZnS/CdTe.

We now discuss the use of DFT with MBJLDA functional and of the semiempirical NN sp^3d^5 and 2NN sp^3s^* TB models to calculate the electronic band structures of ZnS/CdS and CdTe/CdSe heterostructures. The predicted composition and interface strain effects on the fundamental band gaps of ZnS/CdS and CdTe in ZnS/CdS and CdTe/CdSe heterostructures are shown in Fig. 1.25 (left, right), respectively. The calculations are carried out using the WIEN2K simulations based on DFT with the MBJLDA functional (dark-solid line) and the semiempirical NN sp^3d^5 TB model (red-solid lines) and 2NN sp^3s^* TB models (dashed-blue lines) with tight binding parameters given in Tables 1.3, 1.4 and 1.5, integrated with the statistical thermodynamic model. (Since the measured band gaps are near 0 K, we ignored the logarithmic term, which accounts for the electron-phonon interactions for temperature dependence.) We can state that strain effect on the fundamental bandgaps can be quite large when the interface strain increases for large deformation potential and high alloy composition. The predicted band gaps, especially at the Γ point, are in excellent agreement with experiment [12].

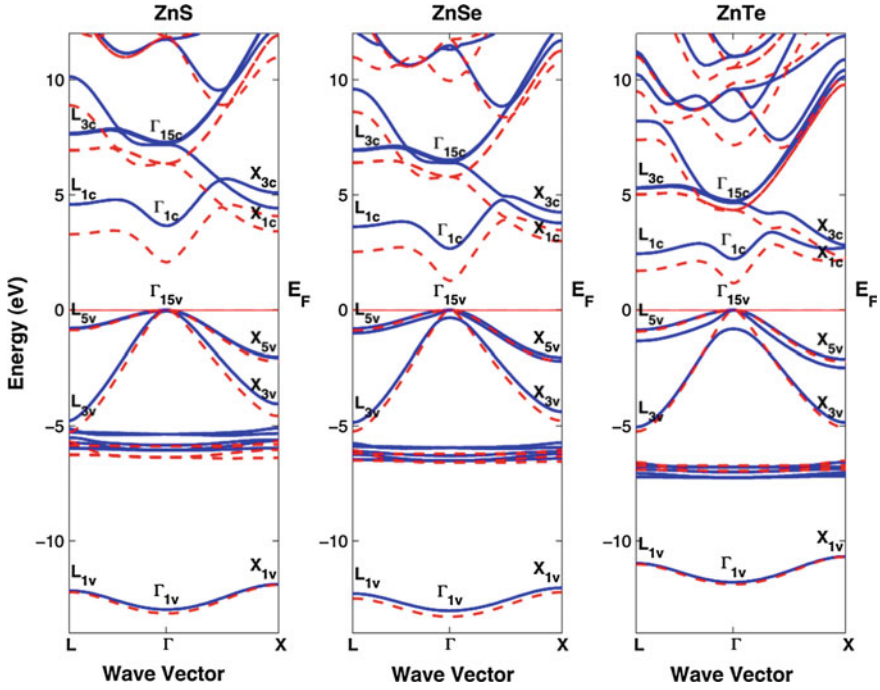


Fig. 1.22 Band structure of ZnS, ZnSe and ZnTe obtained using WIEN2K with DFT-MBJLDA functional (solid lines) and DFT-GGA (dashed lines)

In conclusion, as demonstrated in Figs. 1.23 and 1.24 for CdX and ZnX (X=S, Se, Te) binaries and in 1.25 for ZnSSe and CdSTe ternaries of ZnSSe/ZnSe and CdSTe/CdTe heterostructures, the WIEN2K simulations with DFT-MBJLDA functional and semiempirical NN sp^3d^5 and 2NN sp^3s^* TB methods described in this work accurately reproduce the band gaps and band dispersion curves at high symmetry points of Brillouin zone when they are compared with empirical pseudopotential calculations for these binary and ternary semiconductors. The three different approaches almost identically reproduce the band gaps at the Γ , X and L high symmetry points, as well as the valence band and conduction band dispersion curves in the entire Brillouin zones of the bulk CdX and ZnX (X = S, Se, Te) binaries and ZnSSe and CdSTe ternaries of ZnSSe/ZnSe and CdSTe/CdTe heterostructures, respectively.

As can be seen from Fig. 1.25, the tight binding model with an sp^3d^5 orbital basis, including the NN interactions of cation and anion atoms and spin-orbit coupling of p-states, is adequate to accurately reproduce the band gaps at the Γ , X and L symmetry points and both the valence band and conduction band dispersion curves, obviating the need for any fictitious excited s^* state in the TB model calculations. When the two semiempirical NN sp^3d^5 and 2NN sp^3s^* tight binding parameterizations are compared with each other, they both give reasonable

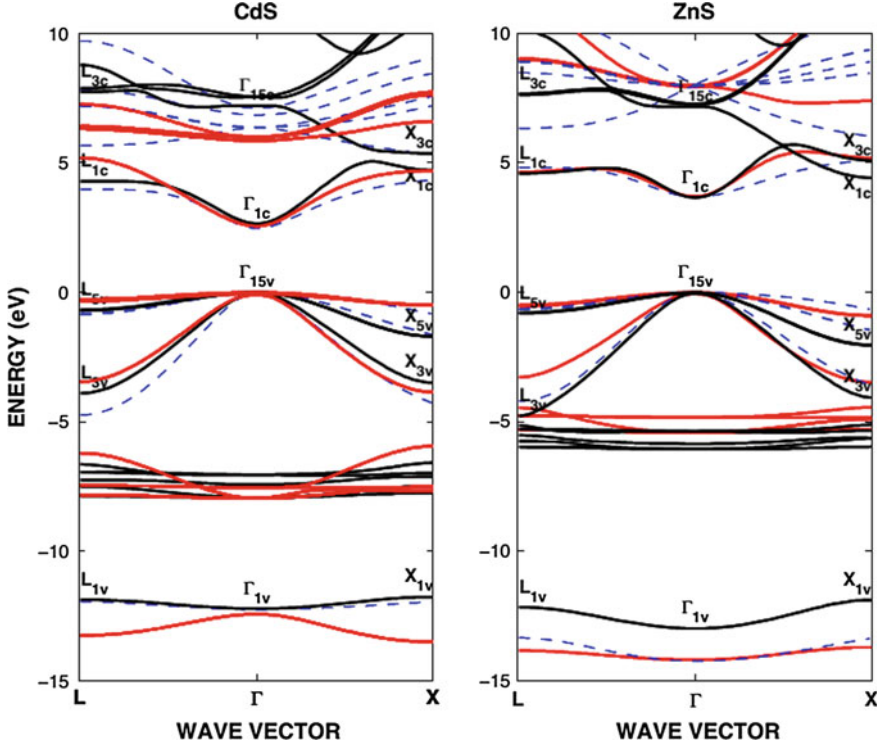


Fig. 1.23 Band structures of CdS and ZnS obtained using WIEN2K with DFT-MBJLDA (dark-solid lines), NN sp^3d^5 TB (red-solid line) and 2NN sp^3s^* TB (dashed-blue line) models

descriptions of the energy levels in the vicinity of the bottom of the conduction band and the top of the valence band of both binary and ternary II–VI compounds. However, the NN sp^3d^5 TB model does better than the 2NN sp^3s^* TB model in accurately reproducing band gaps as well as valence band and conduction band dispersion curves fitted to the empirical pseudopotential calculations at the L high symmetry point of these compounds. Furthermore, since the inclusion of the fictitious s^* excited state is done by modelling the average of p-d interactions, the 2NN sp^3s^* TB parametrization does not permit the inclusion of excited d-orbitals to an sp^3 basis set. Consequently, the actual behavior and contribution of excited d-states in the band structure calculations is reliably and accurately reflected in the NN sp^3d^5 TB model. We should point out that both the NN sp^3d^5 and 2NN sp^3s^* TB models are semiempirical and they rely on a good description of the band structures that are produced exactly by WIEN2K simulations package with MBLDA functional embedded in DFT [47, 48] and empirical pseudopotential theory [14].

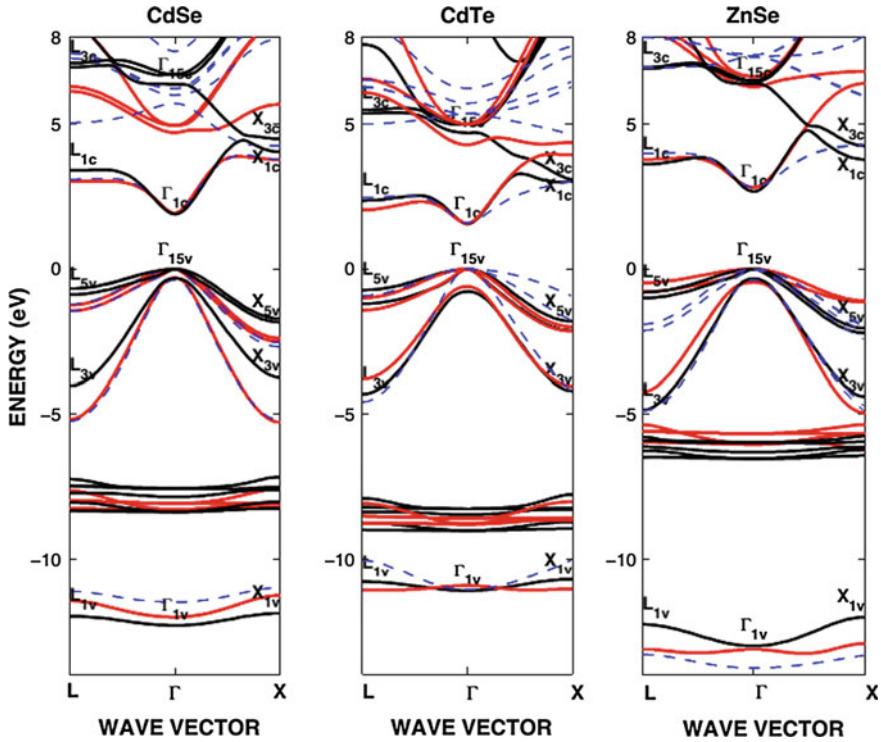


Fig. 1.24 Band structures of CdSe, CdTe and ZnSe obtained using WIEN2K with DFT-MBJLDA (dark-solid lines), NN sp^3d^5 TB (red-solid line) and 2NN sp^3s^* TB (dashed-blue line) models

1.5 Modeling of Band Offsets in Heterostructures

One of the key issues in low dimensional heterostructure semiconductor device technology is the understanding the formation and determining the magnitude of conduction and valence band structure across heterointerfaces (e.g., band offsets), shown in Fig. 1.26.

The potential barriers dominate various device properties such as injection efficiency in heterobipolar devices (e.g., HBTs) and carrier confinement in unipolar devices (e.g., MODFETs), and have received considerable attention among device scientists and engineers over the years. When the semiconductor composition changes abruptly at interface between constituents, the difference in their energy bands is accommodated by the discontinuities in the conduction and valence bands across the interface. In analyzing charge transport in bipolar and unipolar heterostructure devices, the crucial problem is to determine the formation and magnitude of potential barriers relating to conduction bands and valence bands across the heterointerface. In the context of the semiempirical tight binding model

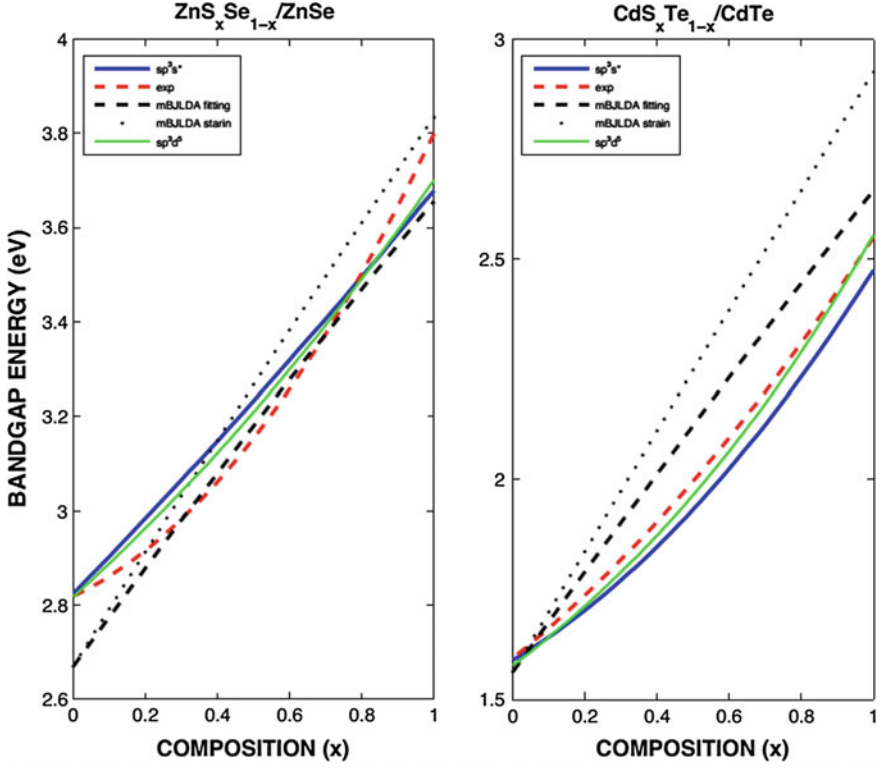


Fig. 1.25 Strain effects on the bandgaps of ZnSSe and CdSTe in ZnSSe/ZnSe and CdSTe/CdTe heterostructures as a function of composition, obtained using the WIEN2K package with DFT-mBJLDA (*dark-solid line*), and the NN sp^3d^5 TB (*red-solid line*) and 2NN sp^3s^* TB (*dashed-blue line*) models for the TB parameters given in Tables 1.3, 1.4 and 1.5

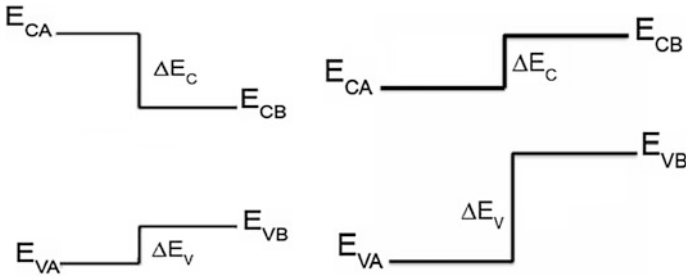


Fig. 1.26 Schematic view of band diagram of Type I (*left*) and Type II (*right*) heterostructures

discussed in Sect. 1.2, the valence band offset across an ABC/AC ternary/binary heterostructure can be obtained by taking the difference between the valence band energies of the constituents that are screened with the optical dielectric constants of the constituents [11]:

$$\Delta E_v = \left(\frac{E_v}{\varepsilon_\infty} \right)_{BC} - \left(\frac{E_v}{\varepsilon_\infty} \right)_{ABC}, \quad (1.27)$$

where $E_v = E_v(\Gamma_{15})$ is top of the valence band at the Γ high symmetry point and $\varepsilon_\infty(ABC)$ and $\varepsilon_\infty(BC)$ are the optical dielectric constants of the ABC ternary and BC binary constituents of heterostructure, respectively:

$$\varepsilon_\infty(ABC) = x\varepsilon_\infty(AC) + (1-x)\varepsilon_\infty(BC) + -x(1-x)\delta_c(\varepsilon_\infty(BC) - \varepsilon_\infty(AC)), \quad (1.28)$$

The conduction band offset for a given Γ , L and X high symmetry point of the Brillouin zone is given as the difference between the respective band gap differences, written as [11]:

$$\Delta E_{ci} = E_{ciABC} - E_{ciBC} = \Delta E_{gi} - \Delta E_v, \quad (1.29)$$

where $E_{ci} = E_{\Gamma_{6c}}, E_{L_{6c}}$ and $E_{X_{6c}}$ are the bottom of the conduction bands at the Γ , L and X high symmetry points, determined using the semiempirical tight binding theory with the sp^3 , sp^3s^* , $sp^3d^5s^*$ or sp^3d^5 orbitals sets. Here $\Delta E_{gi} = E_{gi}(ABC) - E_{gi}(BC)$ is the difference between the band gaps of ABC ternary and BC binary compounds with band gaps $E_{gi}(ABC)$ and $E_{gi}(BC)$, where $E_{gi} = E_{g\Gamma}, E_{gL}$ and E_{gX} are the principal bandgaps at the Γ , L and X symmetry points.

Once the valence band energies and band gaps at high symmetry points are known, it is easy to determine the conduction band offsets of a given heterostructure from (1.29).

The potential barrier in the conduction band and the potential step in the valence band at the interface significantly influence carrier transport and device performance [2], making reliable and accurate modeling of interface energy band structure crucial to designing the functionality of group III–V and II–VI heterostructures and their optimization in low dimensional semiconductor electronic and optical devices. Here, we focus on modeling the effects of composition and interface strain on electronic band structure, band gaps and band offsets of ternary/binary group III–V and II–VI compounds based heterostructures for which experimental data are available for comparison. In this, we compare the predictions of the nearest neighbor (NN) sp^3d^5 tight binding model, with spin-orbit coupling of p-states, to those of the widely used second nearest neighbor (2NN) sp^3s^* model, with spin-orbit coupling of p-states. Tables 1.3, 1.4 and 1.5 give the number of tight binding parameters for the 2NN sp^3s^* TB and NN sp^3d^5 tight TB models used in the calculations. Figure 1.27 shows the composition and strain variations of the band gap difference and valence band offsets at lattice matched AlGaAs/GaAs and pseudomorphic InGaAs/GaAs heterostructures.

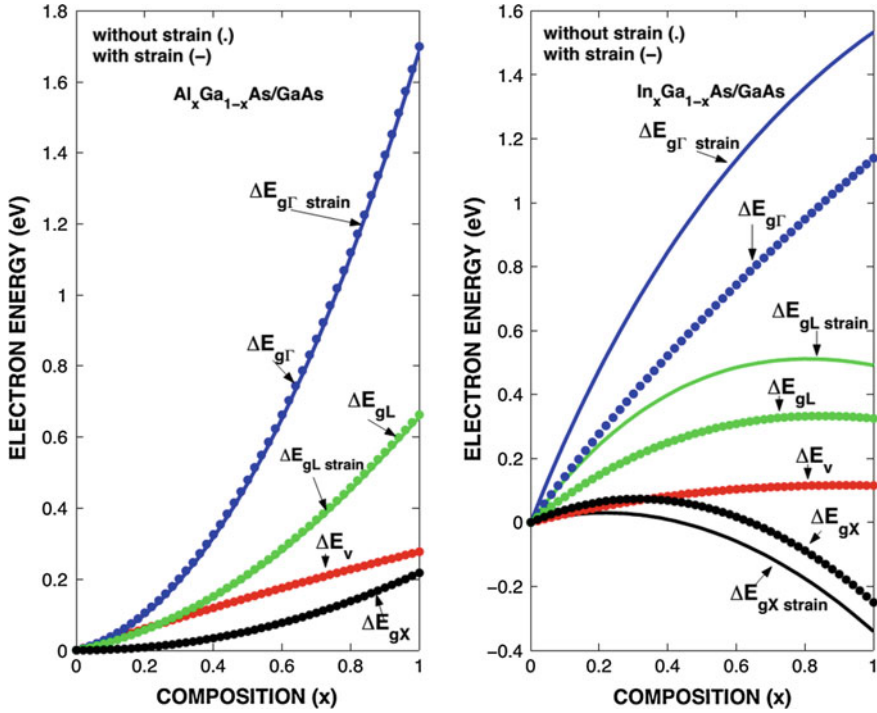


Fig. 1.27 Composition effects on bandgap differences and valence band offsets of lattice matched AlGaAs/GaAs (*left*) and strained InGaAs/GaAs (*right*) heterostructures

The band offsets in III–V nitrides based ternary/binary heterostructures are shown in Fig. 1.28 for AlGaAs/GaAs, InGaAs/GaAs and GaAsN/GaAs as functions of interface strain for the entire composition range ($0 \leq x \leq 1$). As shown in Fig. 1.28, the interface strain effect on the valence band offsets in AlGaAs/GaAs (left), InGaAs/GaAs (middle) and GaAsN/GaAs (right) dilute nitrides based heterostructures is rather small because of the smaller valence band deformation potentials. However, the interface strain effects on conduction band offsets at the Γ , L and X high symmetry points can be considerably larger because the conduction band deformation potentials are large.

The predicted valence band offsets are in good agreement with experimental findings [12, 13]. As shown in Figs. 1.27 and 1.28, strain effects on the band gap differences and on the conduction band offsets can be quite large since the lattice mismatch increases with an increase in the ternary mole fraction. As we pointed out earlier, there is a considerable lattice mismatch across many of the ternary/binary heterointerfaces. Interface strain due to lattice mismatch causes a shift in the lattice

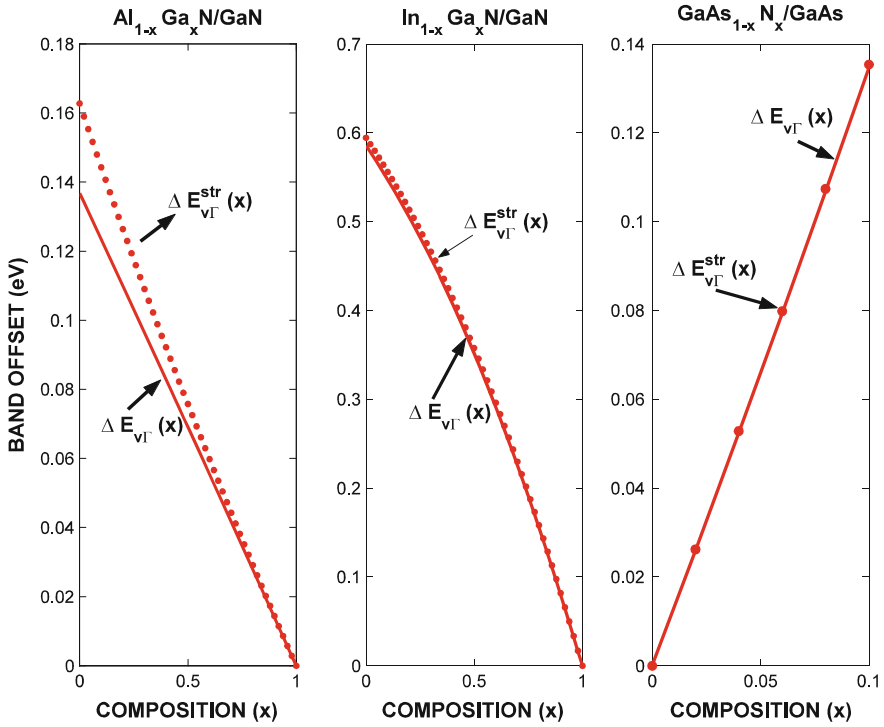


Fig. 1.28 Predicted interface strain effects on the valence band offsets of AlGaIn/GaN (*left*), InGaIn/GaN (*middle*) and GaAsN/GaAs (*right*) heterostructures

constant of the epilayer: $a = (1 + \varepsilon)a_0$, where ε is the symmetric strain tensor. Therefore, the bond lengths and on-site and off-site TB matrix elements will be modified with strain. Consequently, the electronic properties of heterostructures such as conduction and valence band offsets will be modified with respect to their unstrained values. As an example, Fig. 1.29 shows the composition and strain effects on conduction and valence band offsets of CdSTe/CdTe and ZnSSe/ZnSe heterostructures calculated by using the WIEN2K simulation package based on the DFT with MBJLDA functional and semiempirical NN sp^3d^5 and 2NN sp^3s^* TB models for optimized tight binding parameters given in Tables 1.3, 1.4 and 1.5. As can be seen in this comparison, the interface strain effects on the conduction band offsets of ZnSSe/ZnSe and CdSTe/CdTe heterostructures can be quite large when the interface strain increases for large the deformation potential and high alloy composition.

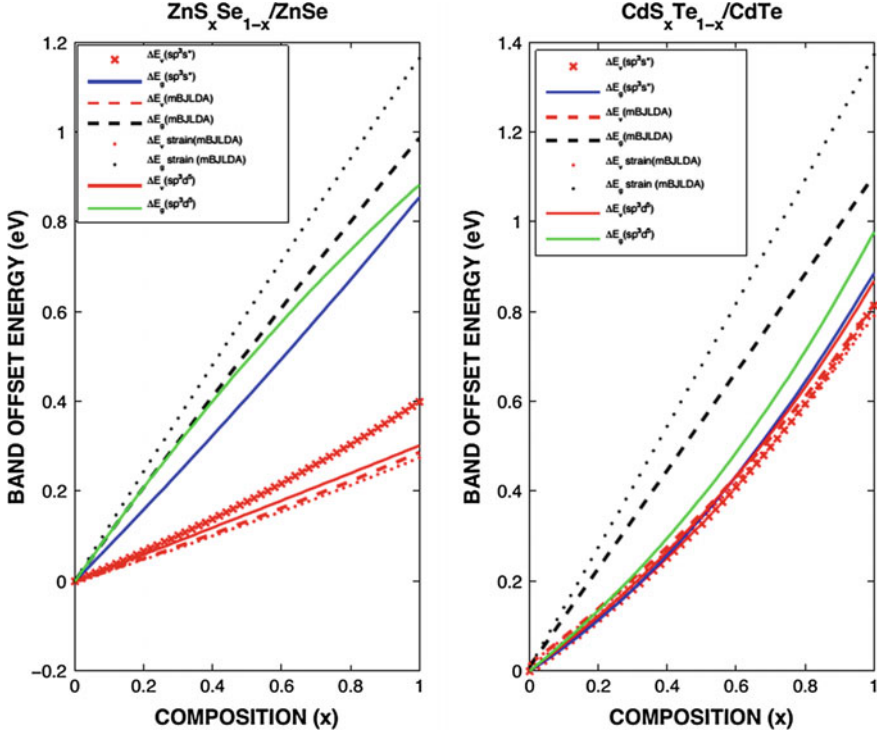


Fig. 1.29 Predicted composition and interface strain effects on the band offsets of ZnSSe/ZnSe and CdSTe/CdTe heterostructures, obtained using the WIEN2K with DFT-MBJLDA (*dark-solid line*) compared with the NN sp^3d^5 (*red-solid line*) and 2NN sp^3s^* (*dashed-blue line*) TB models for parameters given in Tables 1.1 and 1.2 of [30]

1.6 Conclusion

Ever since the invention of the first transistor in 1949, device scientists and engineers have witnessed amazing achievements in the semiconductor science and technology. Advances in the growth of semiconductor thin films with layer thickness approaching atomic dimensions and fabrication of low dimensional electronic and optical devices has provided new opportunities in fundamental science and technology of semiconductors and semiconductor devices. Such advancement could not have been possible without a qualitatively reliable understanding of the basic physics of semiconducting materials and quantitatively precise potential predictions and performance of devices, leading to new concepts in the semiconductor growth that allowed previously many unknown devices with more complex functionality and much higher densities for electronic and optical applications.

This review has presented comparative predictions of electronic band structures of group III–V and II–VI compounds and their heterostructures obtained using the semiempirical tight binding theory based on the 2NN sp^3 , 2NN sp^3s^* , NN sp^3d^{5*} and NN sp^3d^5 orbitals sets in comparison with the WIEN2K simulations package based on the density functional theory (DFT) with the modified modified Becke-Johnson exchange-correlation potential with the local density approximation (LDA), called MBJLDA functional. We conclude that DFT with MBJLDA functional and the semiempirical NN sp^3d^5 TB model are in good agreement with experimental data for band gaps and band offsets and have considerable potential in the design and optimization of group III–V and II–VI compounds and their ternaries used for the fabrication of electronic and optical devices. When compared to the 2NN sp^3s^* TB model, NN sp^3d^5 TB parametrization better simulates both the valence band and conduction band dispersion curves. The actual contribution of the excited d-states in the band structure calculations is reliably reproduced in the case of NN sp^3d^5 orbitals set tight binding parameterization. We conclude that DFT with MBJLDA and semiempirical NN sp^3d^5 TB model show good agreement with experiment for band gaps and band offsets and has a considerable potential in the design and optimization of group III–V and II–VI compounds and their ternaries used for the fabrication of low dimensional electronic and optical devices.

References

1. H. Ünlü, N.J.M. Horing (eds.), *Low Dimensional Semiconductor Structure: Characterization and Applications* (Springer, Heidelberg, 2013)
2. H. Morkoç, H. Ünlü, G. Ji, Principles and technology of MODFETs. Wiley **1**, 2 (1991)
3. H. Ünlü, H. Morkoç, Solid State Technol. **31**, 83 (1988)
4. H. Morkoç, H. Ünlü, in *Semiconductor and Semimetals*, vol. 24 ed. by R. Dingle (1987), p. 135
5. H. Ünlü, H. Morkoç, and S. Iyer, in *Gallium Arsenide Technology*, vol. 2 ed. by D.K. Ferry (1990), p. 231
6. H. Morkoç, H. Ünlü, H. Zabel, N. Otsuka, Solid State Technol. **31**, 71 (1988)
7. H. Kroemer, Proc. IRE **45**, 1535 (1957)
8. H. Kroemer, Rev. Mod. Phys. **73**, 783 (2000)
9. N. Lucas, H. Zabel, H. Morkoç, H. Ünlü, Appl. Phys. Lett. **52**, 2117 (1988)
10. H. Ünlü, Solid State Electron. **35**, 1343 (1992)
11. H. Ünlü, Phys. Status Solidi B **216**, 107 (1999); B **223**, 195 (2001); B **229** 581 (2002); B **235**, 248 (2003)
12. O. Madelung (ed), *Numerical Data and Functional Relationships in Science and Technology*, 17a, Springer, Berlin (1982); ibid 17d, Springer, Berlin (1984)
13. Vurgaftman, I., Meyer, J.R. ve Ram-Mohan, L.R. (2001), J. Appl. Phys. **89**, 5815 (2001)
14. M.L. Cohen, J.R. Chelikowsky, *Electronic Structure and Optical Properties of Semiconductors* (2nd edn), Springer, Berlin (1989)
15. C. Pryor, J. Kim, L.M. Wang, A.J. Williamson, A. Zunger, J. Appl. Phys. **83**, 2548 (1998)
16. P. Keating, Phys. Rev. **145**, 637 (1966)
17. K. Shim, H. Rabitz, Phys. Rev. B **57**, 12874 (1998)
18. M. Rabah, B. Sahraoui, B. Bouhafs, B. Abbar, H. Abid, Phys. Stat. Solidi B **238**, 156 (2003)
19. H.H. Gürel, Ö. Akıncı, H. Ünlü, Comput. Mater. Sci. **33**, 269 (2005)

20. Ö. Akıncı, H.H. Gürel, H. Ünlü, *Thin Solid Films* **511–512**, 684 (2006)
21. H.H. Gürel, Ö. Akıncı, H. Ünlü, *Superlattices and Microstructures*, 40 (4–6), 588 (2006)
22. H.H. Gürel, Ö. Akıncı, H. Ünlü, *Physica Status Solidi (c)* **4**(2), 316 (2007)
23. Ö. Akıncı, H.H. Gürel, H. Ünlü, *Physica Status Solidi (c)* **5**(2), 478 (2008)
24. H.H. Gürel, Ö. Akıncı, H. Ünlü, *Thin Solid Films* **516**, 7098 (2008)
25. Ö. Akıncı, H.H. Gürel, H. Ünlü, *J. Nanoscience Nanotechnology* **8**, 540 (2008)
26. Ö. Akıncı, Ö., H.H. Gürel, H.H. and H. Ünlü, *Thin Solid Films*, 517, 2431 (2009)
27. H.H. Gürel, O. Akıncı, H. Ünlü, *Superlattices Microstruct.* **51**, 725 (2012)
28. H.H. Gürel, H. Ünlü, *Mater. Sci. Semicond. Process.* **16**, 1619 (2013)
29. J.L. Martins, A. Zunger, *Physical Review B* **30**, 6217 (1984)
30. P. Hohenberg and W. Kohn, *Phys. Rev.* 136 B, 864 (1964)
31. J.C. Slater and G. F. Koster (1954), *Physical Review*, 94, 1498 (1954)
32. D.J. Chadi, M.L. Cohen, *Physica Status Solidi* **68**, 405 (1975)
33. D.J. Chadi, *Physical Review B* **16**, 790 (1977)
34. W.A. Harrison, *Electronic Structure and the Properties of Solids*, Freeman (1980)
35. D.N. Talwar, C.S. Ting, *Physical Review B* **25**, 2660 (1982)
36. J.P. Loehr, Talwar, *Phys. Rev. B* **55**, 4353 (1997)
37. T.B. Boykin, L.J. Gamble, G. Klimeck, R.C. Bowen, *Phys. Rev. B* **59**, 7301 (1999)
38. Y. Fu, K.A. Chao, *Phys. Rev. B* **43**, 4119 (1991)
39. A.D. Carlo, *Semiconductor Sci. and Tech.* **18**, R1 (2001)
40. E.P. O'Reilly, A. Lindsay, A., S. Tomic, M.K. Saadi, *Semicond. Sci. Technol.* **17**, 870 (2002)
41. P. Vogl, H.P. Hjalmarson, J.D. Dow, *J. Phys. Chem. Solids* **44**, 365 (1983)
42. M. Jancu, F. R.Scholz, F. Beltram, F. Bassani, *Phys. Rev. B* **57**, 6493 (1998)
43. J.M. Jancu, F. Bassani, F. Della Sala, R. Schols, *Appl. Phys. Lett.* **81**, 4838 (2002)
44. S. Sapra, N. Shanthi, D.D. Sarma, *Phys. Rev. B* **66**, 205202 (2002)
45. W. Kohn, J.L. Sham, *Phys. Rev.* **140A**, 1133 (1965)
46. V. Fiorentini, A. Baldereschi, *Phys. Rev. B* **51**, 17196 (1995)
47. F. Tran, P. Blaha, *Phys. Rev. Lett.* **102**, 226401 (2009)
48. P. Blaha, K. Schwarz, G.K.H. Madsen, D. Kvasnicka, J. Luitz, *WIEN2k, An Augmented Plane Wave Plus Local Orbitals Program For Calculating Crystal Properties*, Vienna University of Technology, Vienna, Austria, 2001

Chapter 2

Aspects of the Modeling of Low Dimensional Quantum Systems

Norman J.M. Horing

Abstract The mathematical modeling of low dimensional quantum systems is discussed in this chapter. In particular, the use of generalized functions in such modeling is illustrated in some detail, including applications of the Dirac delta function and its derivative (“delta-prime”) in determining quantum mechanical Schrödinger Green’s functions describing the dynamics of various low dimensional systems. The illustrations include quantum dots, wires and wells (and a superlattice) in various dimensions. Also, the one-dimensional “delta-prime” potential is shown to provide an impenetrable barrier.

2.1 Introduction

The advent and rapid development of the fabrication of low dimensional semiconductor materials, replete with the promise of nanostructures upon which a whole new generation of quantum electronic, optical and computational devices can be based, has stimulated an enormous effort to explore the physical properties of such materials, and how they can be manipulated to greatest advantage. Practically all the fields of science and engineering are involved in this massive effort throughout the world. Mathematical modeling has an important role in this matter, enabling highly tractable analyses that provide insight into the quantum dynamical behavior of nanostructures and their possible optimization. In this chapter on low dimensional quantum systems, in which the direction/range of particle motion is sharply limited by deep (strong) wells (barriers) in its potential profile, the role of generalized functions, in particular the Dirac delta function, will be discussed in model calculations of the properties of such structures in three, two, one and zero dimensional systems. The generalized functions are useful in representing potentials that confine

N.J.M. Horing (✉)

Department of Physics and Engineering Physics, Stevens Institute of Technology,
Hoboken, New Jersey 07030, USA
e-mail: nhoring@stevens.edu

electrons into spatially small subsystems, such as quantum dots, wires, wells and lattices thereof, lodged in a host medium that may itself be confined in lower dimensions; for example, quantum dots can occur in bulk three-dimensional structures, but can also occur on a thin planar two-dimensional layer and on a one-dimensional quantum wire. Such electron confinement by a potential usually has some wave function extension outside the spatial boundaries of the potential, associated with tunneling phenomenology. However, if the confining potential is so strong (high) that such external wavefunction extension is negligible, then an object, such as a thin quantum well, may be viewed as zero dimensional in the confinement direction(s), and may be reasonably represented by a generalized function in regard to that spatial direction(s). Both of these important uses of generalized function models, in particular Dirac delta functions, in representing (i) confining potentials, and, also, (ii) spatial extension of a low dimensional narrow quantum well, will be discussed here.

The illustrative model calculations to be discussed here include the use of Dirac delta function confining potentials in the determination of Schrödinger Green's functions for:

- (i) a zero dimensional quantum dot in a 3D bulk medium permeated by an isotropic simple harmonic oscillator potential,
- (ii) double quantum dot systems lodged in a thin 2D layer, also in a thin 1D wire,
- (iii) a one dimensional quantum wire on a thin 2D layer in a normal magnetic field.

Also, we present the example of (iv) the space-time inversion of the dynamic, nonlocal dielectric function of a type 1 superlattice. Finally, the physical features of the "delta-prime" derivative of a Dirac-delta function potential will be treated as well.

The calculational procedure generally followed here involves rewriting the differential Schrödinger Green's function equation for the system at hand as an integral equation in terms of the decoupled Green's function without special low dimensional features (such as quantum dots, wires, etc.). The resulting Dyson-type integral equation involves the decoupled Green's function as its driving term and also as part of its kernel. The kernel also involves the special low dimensional features described in terms of generalized functions (such as the Dirac delta function and its derivative), which considerably simplify the equation, so that it may be solved algebraically or by a spatial Fourier transform in directions of translational invariance. The associated energy spectrum is discussed in terms of the frequency poles of the resulting Green's function. In regard to the inversion of the dynamic, nonlocal dielectric function of a superlattice, it is facilitated by an "ansatz" expressing its spatial structure in terms of Dirac delta functions. The resulting superlattice structure factor describes a continuous band of superlattice plasmons.

2.2 3D Harmonic Oscillator with a Superposed Quantum Dot (Fig. 2.1)

To start, we consider a three dimensional isotropic Schrödinger oscillator in a bulk medium, which, by itself, has been viewed as a quantum dot capable of having many electrons in the plethora of its many states. The associated retarded Green's function G_{3D}^{HO} satisfies ($\hbar \rightarrow 1$)

$$\left[i \frac{\partial}{\partial t} + \frac{1}{2m} \nabla^2 - \frac{m\omega_0^2}{2} r^2 \right] G_{3D}^{HO}(\mathbf{r}, \mathbf{r}'; t, t') = \delta^{(3)}(\mathbf{r} - \mathbf{r}') \delta(t - t'),$$

and is given in position-time representation as [1]

$$\begin{aligned} G_{3D}^{HO}(\mathbf{r}, \mathbf{r}'; T) \\ = \eta_+(T) \left(\frac{m\omega_0}{2\pi i \sin \omega_0 T} \right)^{3/2} \exp \left\{ \frac{im\omega_0}{2 \sin \omega_0 T} [(\mathbf{r} \cdot \mathbf{r} + \mathbf{r}' \cdot \mathbf{r}') \cos \omega_0 T - 2\mathbf{r} \cdot \mathbf{r}'] \right\}, \end{aligned} \quad (2.1)$$

where $\mathbf{r} = (x_1, x_2, x_3)$, $T = t - t'$, and ω_0 is the natural frequency of the oscillator of mass m ($\eta_+(T) = \eta_+(t - t')$ is the Heaviside unit step function). The role of a superposed quantum dot (at $\mathbf{r}_0$) in the electron energy spectrum and propagation is represented by the addition of a potential term of the form

$$U(\mathbf{r}) = \alpha \delta^{(3)}(\mathbf{r} - \mathbf{r}_0), \quad (2.2)$$

where $\delta^{(3)}(\mathbf{r} - \mathbf{r}_0)$ is a three dimensional Dirac delta function, and $\alpha = -U_0 V < 0$ describes the product of the quantum well depth U_0 times its volume V . The retarded Green's function $G(\mathbf{r}, \mathbf{r}'; \omega)$ for electron propagation in this bulk medium satisfies the differential equation (frequency/energy representation).

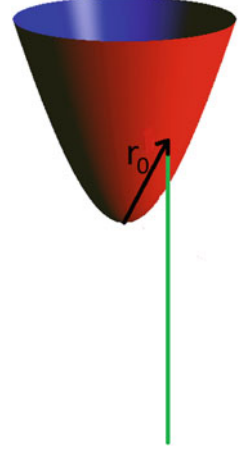
$$\left[\omega + \frac{1}{2m} \nabla^2 - \frac{m\omega_0^2}{2} r^2 - U(\mathbf{r}) \right] G(\mathbf{r}, \mathbf{r}'; \omega) = \delta^{(3)}(\mathbf{r} - \mathbf{r}'). \quad (2.3)$$

Rewriting this as an integral equation in three dimensions, we have

$$G(\mathbf{r}, \mathbf{r}'; \omega) = G_{3D}^{HO}(\mathbf{r}, \mathbf{r}'; \omega) + \int d^3 \mathbf{r}'' G_{3D}^{HO}(\mathbf{r}, \mathbf{r}''; \omega) U(\mathbf{r}'') G(\mathbf{r}'', \mathbf{r}'; \omega), \quad (2.4)$$

where G_{3D}^{HO} is given by (2.1). Introducing (2.2) on the right of (2.4), the $\mathbf{r}''$ -integration leads to

Fig. 2.1 Schematic of a 3D harmonic oscillator potential (red) with a Dirac delta function potential well (green) at $\mathbf{r}_0$



$$G(\mathbf{r}, \mathbf{r}'; \omega) = G_{3D}^{HO}(\mathbf{r}, \mathbf{r}'; \omega) + \alpha G_{3D}^{HO}(\mathbf{r}, \mathbf{r}_0; \omega) G(\mathbf{r}_0, \mathbf{r}'; \omega). \quad (2.5)$$

To solve, set $\mathbf{r} = \mathbf{r}_0$ in (2.5), with the result

$$G(\mathbf{r}_0, \mathbf{r}'; \omega) = \frac{1}{1 - \alpha G_{3D}^{HO}(\mathbf{r}_0, \mathbf{r}_0; \omega)} G_{3D}^{HO}(\mathbf{r}_0, \mathbf{r}'; \omega), \quad (2.6)$$

whence the full Green's function for the “delta-dot” superposed on a 3D oscillator is given by

$$G(\mathbf{r}, \mathbf{r}'; \omega) = G_{3D}^{HO}(\mathbf{r}, \mathbf{r}'; \omega) + \alpha G_{3D}^{HO}(\mathbf{r}, \mathbf{r}_0; \omega) \frac{1}{1 - \alpha G_{3D}^{HO}(\mathbf{r}_0, \mathbf{r}_0; \omega)} G_{3D}^{HO}(\mathbf{r}_0, \mathbf{r}'; \omega). \quad (2.7)$$

The new set of coupled electron energy states occurring as frequency poles of $G(\omega)$ is introduced by the vanishing of the denominator of the last term on the right of (2.7),

$$1 = \alpha G_{3D}^{HO}(\mathbf{r}_0, \mathbf{r}_0; \omega). \quad (2.8)$$

Fourier transforming (2.1) from time to frequency/energy representation, we have

$$1 = \alpha \int_0^\infty dT e^{i\omega T} \left(\frac{m\omega_0}{2\pi i \sin \omega_0 T} \right)^{3/2} \exp \left\{ \frac{im\omega_0 r_0^2}{\sin \omega_0 T} [\cos \omega_0 T - 1] \right\}. \quad (2.9)$$

The right side of (2.9) is readily expanded in the eigenfunctions $\phi_n(\mathbf{r})$ of the 3D harmonic oscillator hamiltonian alone as ($\omega \rightarrow \omega + i0^+$; E_n are the associated decoupled oscillator eigenenergies)

$$1 = \alpha \int_0^\infty dT e^{i\omega T} \sum_n (-i) |\phi_n(\mathbf{r})|^2 e^{-iE_n T} = \alpha \sum_n \frac{|\phi_n(\mathbf{r})|^2}{\omega - E_n}. \quad (2.10)$$

(An alternative evaluation could be developed from the generating function of Laguerre functions [2].) As the 3D oscillator energy spectrum E_n is discrete (and uniformly spaced), the structure of the coupled “delta-dot” eigenenergy dispersion relation of (2.10) makes it clear that the new coupled set of energy roots will occur shifted from each value of E_n , with $n = (n_1, n_2, n_3)$ referring to the three sets of integers corresponding to the three spatial directions of the bulk oscillator medium.

2.3 Double Quantum Dot Systems in Low Dimensions

Considering a GaAs semiconductor inversion layer or other 2D system (e.g. graphene) in which electron motion is confined to the x - y plane at $z = 0$, we examine the role of a double quantum dot induced by a 2D double-well potential ($\alpha_\pm < 0$) constituted of Dirac delta functions:

$$U(\mathbf{r}) = \alpha_+ \delta^{(2)}(\mathbf{r} - \mathbf{r}_+) + \alpha_- \delta^{(2)}(\mathbf{r} - \mathbf{r}_-) = \sum_{\pm} \alpha_{\pm} \delta^{(2)}(\mathbf{r} - \mathbf{r}_{\pm}), \quad (2.11)$$

where $\mathbf{r} = x\hat{i} + y\hat{j}$ and the two potential wells defining the double dot are located at $\mathbf{r}_{\pm} = \pm \hat{i}d/2$ with strengths $\alpha_{\pm} = -U_{0\pm}A_{\pm} < 0$ representing the product of the well depths $U_{0\pm}$ at $\mathbf{r}_{\pm}$ times their common area $A_{\pm} \equiv a^2$. The retarded Green’s function for electron propagation on the sheet, $G(\mathbf{r}_1, \mathbf{r}_2; \omega)$, including the dot regions, obeys the differential Schrödinger equation (frequency representation)

$$\left[\omega + \frac{1}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) - U(\mathbf{r}_1) \right] G(\mathbf{r}_1, \mathbf{r}_2; \omega) = \delta^{(2)}(\mathbf{r}_1 - \mathbf{r}_2). \quad (2.12)$$

Alternatively, this may be rewritten as a two dimensional integral equation as

$$G(\mathbf{r}_1, \mathbf{r}_2; \omega) = G_{2D}^0(\mathbf{r}_1, \mathbf{r}_2; \omega) + \int d^2\mathbf{r}_3 G_{2D}^0(\mathbf{r}_1, \mathbf{r}_3; \omega) U(\mathbf{r}_3) G(\mathbf{r}_3, \mathbf{r}_2; \omega), \quad (2.13)$$

where G_{2D}^0 is the infinite sheet two dimensional Schrödinger Green’s function for electron propagation on the sheet in the absence of the double quantum dot potential wells. Employing (2.11) and integrating with respect to $\mathbf{r}_3$, we have

$$G(\mathbf{r}_1, \mathbf{r}_2; \omega) = G_{2D}^0(\mathbf{r}_1, \mathbf{r}_2; \omega) + \sum_{\pm} \alpha_{\pm} G_{2D}^0(\mathbf{r}_1, \mathbf{r}_{\pm}; \omega) G(\mathbf{r}_{\pm}, \mathbf{r}_2; \omega). \quad (2.14)$$

Setting $\mathbf{r}_1 = \mathbf{r}_{\pm}$ here, $G(\mathbf{r}_{\pm}, \mathbf{r}_2; \omega)$ is readily obtained algebraically, with the result

$$G(\mathbf{r}_{\pm}, \mathbf{r}_2; \omega) = \frac{D_{\mp, \mp} G_{2D}^0(\mathbf{r}_{\pm}, \mathbf{r}_2; \omega) + D_{\pm, \mp} G_{2D}^0(\mathbf{r}_{\mp}, \mathbf{r}_2; \omega)}{D_{\mp, \mp} D_{\pm, \pm} - D_{\pm, \mp} D_{\mp, \pm}}, \quad (2.15)$$

where

$$D_{\pm, \pm} = \delta_{\pm, \pm} - \alpha_{\pm} G_{2D}^0(\mathbf{r}_{\pm}, \mathbf{r}_{\pm}; \omega) \quad (2.16)$$

and

$$D_{\pm, \mp} = \alpha_{\pm} G_{2D}^0(\mathbf{r}_{\pm}, \mathbf{r}_{\mp}; \omega). \quad (2.17)$$

The final double-dot Green's function is given by (2.14)–(2.17) as

$$G(\mathbf{r}_1, \mathbf{r}_2; \omega) = G_{2D}^0(\mathbf{r}_1, \mathbf{r}_2; \omega) + \sum_{\pm} \alpha_{\pm} G_{2D}^0(\mathbf{r}_1, \mathbf{r}_{\pm}; \omega) \times \frac{D_{\mp, \mp} G_{2D}^0(\mathbf{r}_{\pm}, \mathbf{r}_2; \omega) + D_{\pm, \mp} G_{2D}^0(\mathbf{r}_{\mp}, \mathbf{r}_2; \omega)}{D_{\mp, \mp} D_{\pm, \pm} - D_{\pm, \mp} D_{\mp, \pm}}. \quad (2.18)$$

The coupled subband energy eigenstates associated with the double quantum dot arise from the frequency poles of its Green's function, as given by the dispersion relation:

$$\det(D_{\mp, \mp} D_{\pm, \pm} - D_{\pm, \mp} D_{\mp, \pm}) = 0. \quad (2.19)$$

The retarded infinite sheet Schrödinger Green's function for a spinless 2D electron (absent dots) is well known as ($\hbar \rightarrow 1$) [3]

$$G_{2D}^B(\mathbf{r}_1, \mathbf{r}_2; T) = -\eta_+(T) \frac{m}{2\pi} \frac{\exp[\frac{im(X^2 + Y^2)}{2T}]}{T}, \quad (2.20)$$

in position-time representation, with $T = t_1 - t_2$, $X = x_1 - x_2$, $Y = y_1 - y_2$ and $\eta_+(T)$ is the Heaviside unit step function. Fourier transforming from time to frequency representation, and setting $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}_{\pm}$ as needed in (2.16), ($X = 0$, $Y = 0$), we obtain ($\omega \rightarrow \omega + i0^+$)

$$G_{2D}^0(0, 0; \omega) = -\frac{m}{2\pi} \int_0^\infty dT \frac{e^{i\omega T}}{T}. \quad (2.21)$$

This T -integral diverges as $T \rightarrow 0$ for $X = Y = 0$. The divergence is an artifact of assuming a $\delta^{(2)}(\mathbf{r})$ potential which confines the dot to a single point. It may be removed by noting that the original integral equation involves integration over a small, but finite radius “ a ”, so we should install a “smear” radius “ a ” putting $X^2 + Y^2 \rightarrow a^2$ (instead of “0”) in (2.20), leading to the replacement

$$G_{2D}^0(0, 0; \omega) \rightarrow G_{2D}^0(a, \omega) \equiv -\frac{m}{2\pi} \int_0^\infty \frac{dT}{T} e^{i\omega T} \exp(ima^2/2T), \quad (2.22)$$

which is convergent. It is given by a Hankel function of the first kind, order zero ($H_0^{(1)}$), yielding

$$G_{2D}^0(0, 0; \omega) \rightarrow -\frac{im}{2} H_0^{(1)}(\sqrt{2ma^2\omega}) = -\frac{im}{2} \left[J_0(\sqrt{2ma^2\omega}) + iY_0(\sqrt{2ma^2\omega}) \right]. \quad (2.23)$$

For a small (nano) quantum dot, $2ma^2\omega \ll 1$; this may be approximated as [4]

$$G_{2D}^0(0, 0; \omega) \approx \frac{m}{2} Y_0(\sqrt{2ma^2\omega}) \approx \frac{m}{\pi} \ln(\sqrt{ma^2\omega/2}), \quad (2.24)$$

since Y_0 then becomes dominant.

The energy eigenvalue dispersion relation, (2.19), for the 2D Schrödinger coupled quantum double dot system is given by (note that $|\mathbf{r}_\pm - \mathbf{r}_\mp| = d$; $|\mathbf{r}_\pm - \mathbf{r}_\pm| \rightarrow a$; suppress ω):

$$[1 - \alpha_+ G_{2D}^0(a)] [1 - \alpha_- G_{2D}^0(a)] - \alpha_+ \alpha_- [G_{2D}^0(d)]^2 = 0. \quad (2.25)$$

If the dots are widely separated, $d \gg a$, then $G_{2D}^0(d) \rightarrow 0$ due to the extremely rapid variation of the time-dependent exponential on the right of (2.20). The result is two completely independent decoupled modes given by

$$[1 - \alpha_+ G_{2D}^0(a)] = [1 - \alpha_- G_{2D}^0(a)] = 0. \quad (2.26)$$

However, if the separation of the dots shrinks dramatically, $d \rightarrow a$, then (2.25) becomes

$$1 - (\alpha_+ + \alpha_-) G_{2D}^0(a) = 0, \quad (2.27)$$

characterizing a single mode dispersion relation for one quantum dot of combined strength $\alpha_+ + \alpha_-$. Of course, the fully coupled mode solutions of (2.25) are considerably more complicated because of the transcendental character of that dispersion relation.

Another interesting case is that of a double quantum dot superposed on electrons confined to move along a straight line in just one dimension, x . The associated Schrödinger Green's function for free 1D electrons in the absence of the double-dot is [3]

$$G_{1D}^0(x_1, x_2; \omega) = \frac{-im}{\sqrt{2m\omega}} \exp\left(i\sqrt{2m\omega}|x_1 - x_2|\right). \quad (2.28)$$

The “double-dot” Green's function in one-dimension is given by (2.18) with the replacement $G_{2D}^0 \rightarrow G_{1D}^0$ (note that G_{1D}^0 of (2.28) is finite when $x_1 = x_2$). Similar remarks apply to the dispersion relations of (2.25)–(2.27). For the “single-mode” type limits of (2.26), (2.27), the corresponding 1D bound state is given by

$$1 + \frac{ixm}{\sqrt{2m\omega}} = 0, \quad \text{or} \quad \omega = -\frac{\alpha^2 m}{2}, \quad (2.29)$$

as found by Glasser, et al. (Phys. Rev. B 42, 7630 (1990)).

2.4 Quantum Wire on a 2D Sheet in a Uniform Magnetic Field

To examine the effects of a normal magnetic field on the dynamics of a quantum wire along the y -axis at $x = 0$ (on the two dimensional x - y plane at $z = 0$), the confinement of electrons to the wire may be represented by a one-dimensional Dirac-delta function potential as $(\mathbf{r} = (x, y))$ on the 2D sheet)

$$U(\mathbf{r}) = \alpha\delta(x), \quad (2.30)$$

where $\alpha = -U_0 d < 0$, with U_0 as the potential well depth and d as its width. The governing Schrödinger Green's function equation is similar to (2.12), but with the kinetic Hamiltonian part adjusted to include the magnetic field; and $U(\mathbf{r})$ is given by (2.30) for the quantum wire. As usual, the partial differential Schrödinger equation can be reformulated as an integral equation given in frequency/energy representation by

$$G(\mathbf{r}_1, \mathbf{r}_2; \omega) = G_{2D}^B(\mathbf{r}_1, \mathbf{r}_2; \omega) + \int d^{(2)}\mathbf{r}_3 G_{2D}^B(\mathbf{r}_1, \mathbf{r}_3; \omega) U(\mathbf{r}_3) G(\mathbf{r}_3, \mathbf{r}_2; \omega), \quad (2.31)$$

where $G_{2D}^B(\mathbf{r}_1, \mathbf{r}_2; \omega)$ is the two-dimensional magnetic field Green's function on the 2D planar sheet, with no wire present. It involves gauge dependence and lacks spatial translational invariance through the Peierls phase factor $C(\mathbf{r}, \mathbf{r}')$: ($\mathbf{A}(\mathbf{r})$ is the vector potential of the uniform magnetic field)

$$C(\mathbf{r}, \mathbf{r}') = \exp\left(\frac{ie}{\hbar c} \int_{\mathbf{r}'}^{\mathbf{r}} d\mathbf{r}'' \cdot \mathbf{A}(\mathbf{r}'')\right), \quad (2.32)$$

such that G_{2D}^B has the form

$$G_{2D}^B(\mathbf{r}, \mathbf{r}'; \omega) = C(\mathbf{r}, \mathbf{r}') G'_{2D}(\mathbf{r}, \mathbf{r}'; \omega), \quad (2.33)$$

where $G'_{2D}(\mathbf{r}, \mathbf{r}'; \omega) = G'_{2D}(\mathbf{r} - \mathbf{r}'; \omega)$ is effectively translationally invariant on the 2D sheet. Employing (2.30) in (2.31) yields (suppress ω)

$$G(\mathbf{r}_1, \mathbf{r}_2) = G_{2D}^B(\mathbf{r}_1, \mathbf{r}_2) + \alpha \int d^{(2)}\mathbf{r}_3 G_{2D}^B(\mathbf{r}_1, \mathbf{r}_3) \delta(x_3) G(\mathbf{r}_3, \mathbf{r}_2), \quad (2.34)$$

and executing the x_3 -integration,

$$G(x_1, y_1; x_2, y_2) = G_{2D}^B(x_1, y_1; x_2, y_2) + \alpha \int dy_3 G_{2D}^B(x_1, y_1; x_3 = 0, y_3) G(x_3 = 0, y_3; x_2, y_2). \quad (2.35)$$

Focusing attention on electron propagation *along the wire* with $x_1 = x_2 = 0$, we note that the Peierls phase integral, $\int_{\mathbf{r}'}^{\mathbf{r}} d\mathbf{r}'' \cdot \mathbf{A}(\mathbf{r}'')$, which occurs in G as well as G_{2D}^B (and is generally dependent on the path from $\mathbf{r}'$ to $\mathbf{r}$ in the x - y plane) in consequence of gauge covariance, is now confined to a straight line path along the y -axis. Accordingly, its role in (2.35) is eliminated and with

$$G(\mathbf{r}_1, \mathbf{r}_2; \omega) = C(\mathbf{r}_1, \mathbf{r}_2) G'(\mathbf{r}_1, \mathbf{r}_2; \omega), \quad (2.36)$$

we have the Green's function "on-axis" determined by

$$G'(x_1 = 0, y_1; x_2 = 0, y_2) = G'_{2D}(x_1 = 0, y_1; x_2 = 0, y_2) + \alpha \int dy_3 G'_{2D}(x_1 = 0, y_1; x_3 = 0, y_3) G'(x_3 = 0, y_3; x_2 = 0, y_2). \quad (2.37)$$

Since G'_{2D} is translationally invariant, this integral equation can be solved by a single momentum transform in the y -direction ($y_1 - y_2 \rightarrow p_y$), yielding

$$G'(x_1 = 0, x_2 = 0; p_y) = \frac{1}{1 - \alpha G'_{2D}(x_1 = 0, x_2 = 0; p_y)} G'_{2D}(x_1 = 0, x_2 = 0; p_y). \quad (2.38)$$

Finally, the Green's function for electron propagation *along the quantum wire* is determined by substituting (2.38) as the last G' -factor on the right of (2.37) transformed to p_y -representation, with the result

$$G'(x_1 = 0, x_2 = 0; p_y; \omega) = G'_{2D}(x_1 = 0, x_2 = 0; p_y; \omega) + \alpha G'_{2D}(x_1 = 0, x_2 = 0; p_y; \omega) \times [1 - \alpha G'_{2D}(x_1 = 0, x_2 = 0; p_y; \omega)]^{-1} G'_{2D}(x_1 = 0, x_2 = 0; p_y; \omega) \quad (2.39)$$

as a function of frequency/energy (and p_y). (Off-axis propagation is more complicated.) It is immediately apparent that the quantum wire energy spectrum is determined by the zeros of the denominator of (2.39), (restore ω)

$$1 - \alpha G'_{2D}(x_1 = 0, x_2 = 0; p_y; \omega) = 0. \quad (2.40)$$

Note that the energy spectrum of (2.40) supplants the original spectrum associated with the frequency poles of $G'_{2D}(x_1 = 0, x_2 = 0; p_y; \omega)$, since $G(x_1 = 0, x_2 = 0; p_y; \omega)$ vanishes identically when $G'_{2D}(x_1 = 0, x_2 = 0; p_y; \omega) \rightarrow \infty$ at its pole positions, as can be seen from (2.39).

The 2D magnetic field Landau-quantized retarded Schrödinger Green's function is given in position-time representation as [3, 5]

$$G'_{2D}(\mathbf{r}_1, \mathbf{r}_2; T) = -\eta_+(T) \frac{m\omega_c}{4\pi} \exp\left[\frac{im\omega_c(X^2 + Y^2)}{4 \tan(\omega_c T/2)}\right] / \sin(\omega_c T/2), \quad (2.41)$$

where $X = x_1 - x_2$; $Y = y_1 - y_2$; $T = t_1 - t_2$; ω_c is the cyclotron frequency, m is the effective mass and $\eta_+(T)$ is the Heaviside unit step function. Fourier transforming $Y \rightarrow p_y$ and $T \rightarrow \omega$, we have

$$G'_{2D}(X; p_y; \omega) = \frac{m\omega_c}{\pi} \sum_{n=0}^{\infty} \frac{e^{-m\omega_c X^2/4}}{\omega - [n + 1/2]\omega_c} \times \int_0^{\infty} dY \cos(p_y Y) \exp(-m\omega_c Y^2/4) L_n(m\omega_c [X^2 + Y^2]/2), \quad (2.42)$$

where the T -integral was carried out using the integrand as the generator of a series of Laguerre polynomials $L_n(x)$ [2]. The Y -integration is facilitated by expanding

$L_n(m\omega_c[X^2 + Y^2]/2)$ in products of Hermite polynomials [6],
 $\left(\binom{n}{k} \text{ is the binomial coefficient}\right)$

$$G'_{2D}(X; p_y; \omega) = \frac{m\omega_c}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{e^{-m\omega_c X^2/4}}{\omega - [n + 1/2]\omega_c} \sum_{k=0}^n \binom{n}{k} H_{2n-2k} \left(\sqrt{\frac{m\omega_c}{2}} X \right) \\ \times \int_0^{\infty} dY \cos(p_y Y) \exp(-m\omega_c Y^2/4) H_{2k} \left(\sqrt{\frac{m\omega_c}{2}} Y \right). \quad (2.43)$$

The Y -integral can now be evaluated in terms of a variant of Hermite polynomials defined as $He_n(x) = 2^{n/2} H_n(x/\sqrt{2})$ [7]:

$$G'_{2D}(X, p_y; \omega) = -\sqrt{\frac{m\omega_c}{\pi}} e^{-m\omega_c X^2/4} e^{-p_y^2/m\omega_c} \sum_{n=0}^{\infty} \frac{(-1)^n}{n! (\omega - [n + 1/2]\omega_c)} \\ \times \sum_{k=0}^n \left(\frac{1}{2}\right)^k \binom{n}{k} H_{2n-2k} \left(\sqrt{\frac{m\omega_c}{2}} X \right) He_{2k} \left(\frac{2p_y}{\sqrt{m\omega_c}} \right). \quad (2.44)$$

In regard to the eigenenergy spectrum, the usual 2D Landau quantized energy levels in the absence of a quantum wire ($\alpha = 0$) appear as the frequency poles of G'_{2D} in the first term on the right of (2.35). The coupled eigenenergies associated with the quantum wire in a magnetic field perpendicular to the 2D sheet arise from (2.40). For electron propagation along the wire, $x_1 = x_2 = 0$, the analysis above yields the coupled ($\alpha \neq 0$) dispersion relation as

$$1 = \frac{\alpha}{\hbar^{3/2}} \sqrt{\frac{m\omega_c}{\pi}} e^{-\hbar p_y^2/m\omega_c} \sum_{n=0}^{\infty} \left[H_n \left(p_y \sqrt{\frac{\hbar}{m\omega_c}} \right) \right]^2 \frac{2^n/n!}{\omega - [n + 1/2]\omega_c}. \quad (2.45)$$

For a very narrow nanowire, $\alpha^2 m/\hbar^3 \omega_c = d^2 m U_0^2/\hbar^3 \omega_c \ll 1$, it is clear that there is an energy root ω_n near each multiple of the cyclotron frequency, $n\omega_c$, close enough so that the nearest pole alone determines it as

$$\omega_n = [n + 1/2]\omega_c + \frac{\alpha}{\hbar^{3/2}} \sqrt{\frac{m\omega_c}{\pi}} \frac{2^n}{n!} e^{-\hbar p_y^2/m\omega_c} \left[H_n \left(p_y \sqrt{\frac{\hbar}{m\omega_c}} \right) \right]^2. \quad (2.46)$$

2.5 Inverse Dielectric Function of a Periodic Superlattice

The screening function $K(1, 2)$ [$1 = \mathbf{r}_1, t_1 = (x_1, y_1, z_1, t_1)$] of a periodic superlattice (Fig. 2.2) provides an interesting and useful example of the use of Dirac delta

functions in representing nano-thin systems, in particular 2D dynamic, nonlocal plasmonic sheets. The frequency poles of $K(1, 2)$ are the superlattice plasmons and its zero frequency limit describes static shielding. Its role is subsumed in the linear relations

$$K(1, 2) = \frac{\delta V(1)}{\delta U(2)} \quad \text{or} \quad V(1) = \int d^4 2 K(1, 2) U(2), \quad (2.47)$$

which indicate that the action of $K(1, 2)$ on an impressed potential $U(2)$ at space-time point 2 yields the resulting effective potential $V(1)$ at 1. Its relation to the dielectric function $\varepsilon(1, 2)$,

$$\varepsilon(1, 2) = \delta(1 - 2) + 4\pi\alpha_0(1, 2), \quad (2.48)$$

where $4\pi\alpha_0(1, 2)$ is the polarizability, defines $K(1, 2)$ as the space-time matrix inverse of $\varepsilon(1, 2)$:

$$\int d^4 3 K(1, 3) \varepsilon(3, 2) = \delta^4(1, 2) \rightarrow 4\text{D space-time unit matrix}. \quad (2.49)$$

Alternatively expressed, $K(1, 2)$ obeys the integral equation [3]

$$K(1, 2) = \delta(1 - 2) + \int d^4 3 \int d^4 4 v(1 - 3) R(3, 4) K(4, 2), \quad (2.50)$$

where $R(1, 2)$ is the density perturbation response function of the system, $v(1 - 2)$ is the Coulomb potential and

$$R(1, 2) = \frac{\delta \rho(1)}{\delta V(2)} = -iG_1(1, 2)G_1(2, 1^+), \quad (2.51)$$

with G_1 as the single particle noninteracting Green's function in the random phase approximation. It is readily seen that this corresponds to the polarizability as

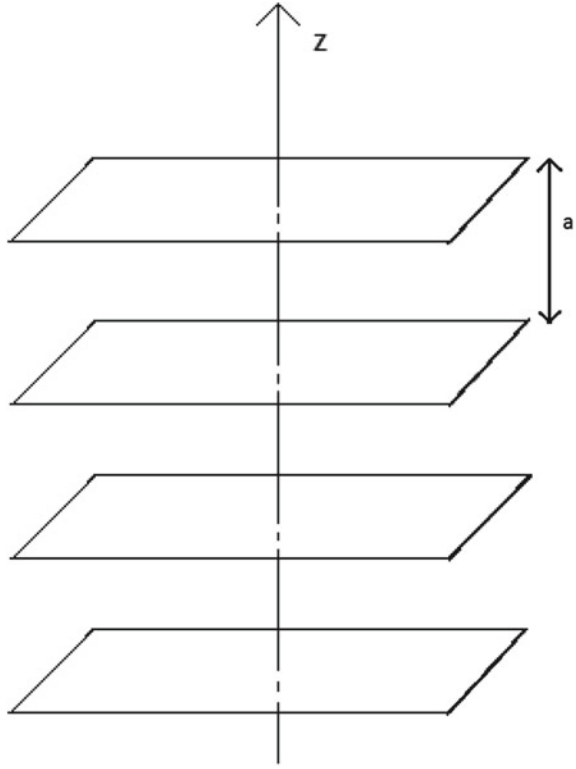
$$4\pi\alpha_0(1, 2) = - \int d^4 3 v(1 - 3) R(3, 2). \quad (2.52)$$

Furthermore,

$$\varepsilon(1, 2) = \delta(1 - 2) - \int d^4 3 v(1 - 3) R(3, 2). \quad (2.53)$$

As an illustration of the usefulness of Dirac delta functions in representing the spatial extension of a low dimensional system, we consider an infinite superlattice of identical 2D planar sheets of electrons parallel to the x - y plane having period a along the z -axis. Each sheet is characterized by a 2D density perturbation

Fig. 2.2 Schematic of a type 1 superlattice



response function of (2.51): Assuming that only the lowest electron energy subband $\phi_0(z)$ for motion in the z -direction across the sheet n is energetically accessible, an associated eigenfunction expansion of $R(1,2)$ has the form $R(1,2) \rightarrow R(x_1 - x_2; y_1 - y_2; t_1 - t_2) \rightarrow |\phi_0(z_1 - na)|^2 |\phi_0(z_2 - na)|^2$. As the width of the sheet shrinks to zero, so does $\phi_0(z_1 - na)$, except at $z = 0$ where it grows to accommodate the normalization to unity,

$$\text{i.e., } |\phi_0(z_1 - na)|^2 \rightarrow \delta(z_1 - na).$$

Furthermore, since there is translational invariance in the x - y plane of the sheet, $R(1,2;\omega)$ can be Fourier transformed in terms of a single 2D wave vector $\mathbf{Q}$ in the x - y plane, such that

$$R(\mathbf{Q}, z_1, z_2; \omega) = \sum_{n=-\infty}^{\infty} \delta(z_1 - na) \delta(z_2 - na) R^{2D}(\mathbf{Q}, \omega), \quad (2.54)$$

since the densities of the various sheets are additive. The corresponding 2D wave vector transform of the Coulomb potential coupling is (charge strength λ)

$$v(1-3) \rightarrow (\lambda/Q)\exp(-Q|z_1 - z_3|), \quad (2.55)$$

so (2.53) in $\mathbf{Q}$ -representation is given by

$$\varepsilon(\mathbf{Q}, z_1, z_2; \omega) = \delta(z_1 - z_2) - 4\pi\alpha_0^{2D}(\mathbf{Q}, \omega) \sum_{n=-\infty}^{\infty} e^{-Q|z_1 - na|} \delta(z_2 - na), \quad (2.56)$$

where the z_3 -integral has been executed and we have defined the corresponding polarizability of a single 2D sheet as

$$4\pi\alpha_0^{2D}(\mathbf{Q}, \omega) = \lambda R^{2D}(\mathbf{Q}, \omega)/Q. \quad (2.57)$$

To address the problem of inverting $\varepsilon(\mathbf{Q}, z_1, z_2; \omega)$ as a matrix in z_1, z_2 , we make the ansatz for $K(\mathbf{Q}, z_2, z_3; \omega)$:

$$K(\mathbf{Q}, z_2, z_3; \omega) = \delta(z_2 - z_3) + \sum_{n'=-\infty}^{\infty} C_{n'}(\mathbf{Q}, z_2, z_3; \omega) \delta(z_3 - n'a), \quad (2.58)$$

with the coefficients $C_{n'}$ to be determined. The inversion condition, written in the form

$$\int dz_2 \varepsilon(\mathbf{Q}, z_1, z_2; \omega) K(\mathbf{Q}, z_2, z_3; \omega) = \delta(z_1 - z_3) \quad (2.59)$$

with (2.56), (2.58) leads to (suppress Q and ω)

$$\begin{aligned} & \sum_{n'=-\infty}^{\infty} \left[C_{n'}(z_1, n'a) - 4\pi\alpha_0^{2D} e^{-Q|z_1 - n'a|} \right] \delta(z_3 - n'a) \\ &= 4\pi\alpha_0^{2D} \sum_{n=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} e^{-Q|z_1 - na|} C_{n'}(na, n'a) \delta(z_3 - n'a), \end{aligned} \quad (2.60)$$

where the z_2 -integration has been carried out. Equating the coefficients of like delta functions, we have

$$C_{n'}(z_1, n'a) - 4\pi\alpha_0^{2D} \sum_{n=-\infty}^{\infty} e^{-Q|z_1 - na|} C_{n'}(na, n'a) = 4\pi\alpha_0^{2D} e^{-Q|z_1 - n'a|}. \quad (2.61)$$

To solve for $C_{n'}(z_1, na)$, it is necessary to determine $C_{n'}(na, n'a)$. Because of periodicity, it is to be expected that

$$C_{n'}(ma, n'a) = G(m - n'), \quad (2.62)$$

and setting $z_1 = ma$ in (2.61) yields

$$G(m - n') - 4\pi\alpha_0^{2D} \sum_{n=-\infty}^{\infty} e^{-Qa|m-n|} G(n - n') = 4\pi\alpha_0^{2D} e^{-Qa|m-n'|}. \quad (2.63)$$

Introducing a Fourier (series) analysis with

$$G(m - n') = \frac{a}{2\pi} \int_{-\pi/a}^{\pi/a} dq e^{iqa(m-n')} g(q), \quad (2.64)$$

leads to the result [8]

$$g(q) = \frac{4\pi\alpha_0^{2D} \sinh(Qa)}{\cosh(Qa) - \cos(qa)} \left(1 - 4\pi\alpha_0^{2D} \frac{\sinh(Qa)}{\cosh(Qa) - \cos(qa)} \right)^{-1}, \quad (2.65)$$

from which $G(n - n')$ (2.64) and $C_{n'}(na, n'a)$ (2.62) may be determined: In turn, this provides the coefficients $C_{n'}(z_1, n'a)$ through (2.61). This suffices to yield the full solution for the inverse dielectric screening function $K(\mathbf{Q}, z_1, z_2; \omega)$ (since $C_{n'}(\mathbf{Q}, z_2, z_3, \omega)$ in (2.58) involves z_3 only at $z_3 = n'a$). Introducing the definition of the structure factor $S(Q, q)$,

$$S(Q, q) = \frac{\sinh(Qa)}{\cosh(Qa) - \cos(qa)}, \quad (2.66)$$

the final result for the superlattice screening function is given by

$$\begin{aligned} K(\mathbf{Q}, z_1, z_2; \omega) &= \delta(z_1 - z_2) + \sum_{n'=-\infty}^{\infty} \delta(z_2 - n'a) C_{n'}(\mathbf{Q}, z_1, n'a; \omega) \\ &= \delta(z_1 - z_2) + 4\pi\alpha_0^{2D}(\mathbf{Q}, \omega) \sum_{n'=-\infty}^{\infty} \delta(z_2 - n'a) \\ &\quad \times \left(e^{-Q|z_1 - n'a|} + \sum_{n=-\infty}^{\infty} e^{-Q|z_1 - na|} \frac{1}{2\pi} \int_{-\pi}^{\pi} dx e^{ix(n-n')} \frac{4\pi\alpha_0^{2D}(\mathbf{Q}, \omega) S(Q, x/a)}{1 - 4\pi\alpha_0^{2D}(\mathbf{Q}, \omega) S(Q, x/a)} \right). \end{aligned} \quad (2.67)$$

The spectrum of superlattice plasmons determined by the dispersion relation arising from the vanishing of the denominator of the integrand on the right of (2.67),

$$1 - 4\pi\alpha_0^{2D}(\mathbf{Q}, \omega) S(Q, q) = 0, \quad (2.68)$$

involves a continuous band of mode frequencies [9–24] associated with the integral limits $-\pi/a \leq q \leq \pi/a$. The use of Dirac delta functions here has facilitated the explicit determination of the inverse dielectric screening function as well as the dynamic, nonlocal plasmon dispersion relation.

2.6 An Impenetrable Barrier: Quantum Particle Dynamics in a Highly Singular 1D-Potential

The use of the one dimensional Dirac delta function ($\delta(x)$) itself as a potential has proven to be quite straightforward; however, the inclusion of its derivative ($\delta'(x)$) has brought forth substantial controversy [25–29]. As far back as 1986, Seba [28] found that electron transmission through that highly singular potential could not take place. However, other researchers have appended boundary conditions to $\delta'(x)$ and claimed that electron transmission can occur [30–39].

We recently studied this problem by constructing the full, exact Green's function for the Dirac-delta-function-derivative model, defining it solely in terms of the usual derivative property [$\delta'(x) \equiv d(\delta(x))/dx$] under integration by parts, with no appeal to additional boundary conditions [40]. Aside from such an integration by parts, we employ only the conventional, well established properties of the Dirac delta function, $\delta(x)$, in terms of integration as a generalized function, $\int_{-\infty}^{\infty} dx \delta(x-a)f(x) = f(a)$, and differentiation, $d(\eta_+(x-a))/dx = \delta(x-a)$, where $\eta_+(x)$ is the Heaviside unit step function (which is understood to have the value $\eta_+(0) = 1/2$, to which a Fourier series representation converges at the position of the step). Our avoidance of boundary conditions is due to the fact that they are not at all necessary under the usual conventions and that they may distort the meaning of $\delta'(x)$ beyond recognition. The Green's function we obtain analytically in closed form confirms Seba's finding that there can be no electron transmission through the $\delta'(x)$ -singular potential if the potential profile is otherwise well-behaved, and that an incident Schrödinger wave-packet must be totally reflected.

Here, we broaden these one-dimensional considerations to take into account a spatially variable mass, $m(x)$, and a reasonably well behaved spatially variable potential, $V(x)$, and a $\delta(x)$ -potential, all in addition to the $\delta'(x)$ -potential previously examined. It is in this very general situation that we construct the full, exact Schrödinger Green's function in closed form, and show that *no* electron wave packet transmission through $\delta'(x)$ can occur. Our analysis of this problem is in agreement with aspects of the formulation by Park [38], but the latter is quite limited in application to one specific case.

Allowing for a variable mass, $m(x)$, the one-dimensional Schrödinger Green's function equation for a system with time-translational invariance has the (hermitian) Sturm-Liouville form in frequency representation $t - t' \rightarrow \omega$ (we suppress the

explicit appearance of ω , so $G(x, x'; t - t') \rightarrow G(x, x'; \omega) \equiv G(x, x')$, as given by [41]

$$\left[\frac{\partial}{\partial x} \left(\frac{1}{m(x)} \frac{\partial}{\partial x} \right) + V(x) + U(x) \right] G(x, x') = \delta(x - x'). \quad (2.69)$$

Here, $V(x)$ is understood to be a relatively well behaved potential (which can accommodate finite discontinuities using the well known Green's function joining technique employed in the theory of surface/interface states [42–46]; it also includes an ω -term from the Fourier time-transform), and the highly singular delta-function-type generalized function potentials are relegated to $U(x)$ as

$$U(x) = -\alpha\delta(x) + \beta\delta'(x) \quad (2.70)$$

(α, β are constants and $\delta'(x) \equiv d(\delta(x))/dx$).

To start, we define an auxiliary Green's function, $G_0(x, x')$, as the inverse of the Sturm-Liouville operator excluding $U(x)$:

$$\left[\frac{\partial}{\partial x} \left(\frac{1}{m(x)} \frac{\partial}{\partial x} \right) + V(x) \right] G_0(x, x') = \delta(x - x'). \quad (2.71)$$

Taking $y_1(x)$ and $y_2(x)$ as two linearly independent solutions of the homogeneous counterpart of (2.71), with y_1 chosen to satisfy the boundary condition at the lower limit and y_2 doing so at the upper limit, the solution of (2.71) is known to be [41]

$$G_0(x, x') = \frac{m(x')}{\Delta(y_1, y_2)} \begin{cases} y_1(x)y_2(x') & \text{for } x < x' \\ y_2(x)y_1(x') & \text{for } x > x' \end{cases}, \quad (2.72)$$

where $\Delta(y_1, y_2)$ is the Wronskian of the two solutions, $y_1(x'), y_2(x')$, evaluated at x' :

$$\Delta(y_1, y_2) = \det \begin{vmatrix} y_1 & y_1' \\ y_2 & y_2' \end{vmatrix}, \quad (2.73)$$

($y' \equiv dy(x')/dx'$).

Considering y_1, y_2 to be known, and hence G_0 is known, (2.69) may be rewritten as

$$G(x, x') = G_0(x, x') + \int_{-\infty}^{\infty} dx'' G_0(x, x'') U(x'') G(x'', x'), \quad (2.74)$$

or

$$G(x, x') = G_0(x, x') - \alpha G_0(x, 0)G(0, x') + \beta \int_{-\infty}^{\infty} dx'' \delta'(x'') G_0(x, x'') G(x'', x'). \quad (2.75)$$

Integrating the last term on the right by parts, this becomes

$$G(x, x') = G_0(x, x') - \alpha G_0(x, 0)G(0, x') - \beta \int_{-\infty}^{\infty} dx'' \delta(x'') \frac{\partial}{\partial x''} [G_0(x, x'') G(x'', x')]. \quad (2.76)$$

Introducing “left” and “right” derivative notation (L and R , respectively),

$$\begin{aligned} \frac{\partial}{\partial x''} G(x'', x') &\equiv [\partial_{(L)} G(x'', x')]; \\ \frac{\partial}{\partial x''} G(x', x'') &\equiv [\partial_{(R)} G(x', x'')]; \\ \frac{\partial}{\partial x'} \frac{\partial}{\partial x''} G(x', x'') &\equiv [\partial_{(L,R)}^2 G(x', x'')], \end{aligned} \quad (2.77)$$

Equation (2.76) may be written as

$$G(x, x') = G_0(x, x') - \alpha G_0(x, 0)G(0, x') - \beta [\partial_{(R)} G_0(x, 0)] G(0, x') - \beta G_0(x, 0) [\partial_{(L)} G(0, x')]. \quad (2.78)$$

To solve, we need to determine $G(0, x')$ and $[\partial_{(L)} G(0, x')]$: Setting $x \rightarrow 0$, we obtain (2.78) as

$$G(0, x') = G_0(0, x') - \alpha G_0(0, 0)G(0, x') - \beta [\partial_{(R)} G_0(0, 0)] G(0, x') - \beta G_0(0, 0) [\partial_{(L)} G(0, x')]. \quad (2.79)$$

Forming the left derivative of (2.78), we have

$$\begin{aligned} [\partial_{(L)} G(x, x')] &= [\partial_{(L)} G_0(x, x')] - \alpha [\partial_{(L)} G_0(x, 0)] G(0, x') \\ &\quad - \beta [\partial_{(L,R)}^2 G_0(x, 0)] G(0, x') - \beta [\partial_{(L)} G_0(x, 0)] [\partial_{(L)} G(0, x')], \end{aligned} \quad (2.80)$$

and putting $x \rightarrow 0$ in (2.80) yields

$$\begin{aligned} [\partial_{(L)}G(0, x')] &= [\partial_{(L)}G_0(0, x')] - \alpha[\partial_{(L)}G_0(0, 0)]G(0, x') \\ &\quad - \beta[\partial_{(L,R)}^2G_0(0, 0)]G(0, x') - \beta[\partial_{(L)}G_0(0, 0)][\partial_{(L)}G(0, x')], \end{aligned} \quad (2.81)$$

which expresses $[\partial_{(L)}G(0, x')]$ in terms of $G(0, x')$ as

$$\begin{aligned} (1 + \beta[\partial_{(L)}G_0(0, 0)])[\partial_{(L)}G(0, x')] &= [\partial_{(L)}G_0(0, x')] - \alpha[\partial_{(L)}G_0(0, 0)]G(0, x') \\ &\quad - \beta[\partial_{(L,R)}^2G_0(0, 0)]G(0, x'), \end{aligned} \quad (2.82)$$

or

$$\begin{aligned} [\partial_{(L)}G(0, x')] &= (1 + \beta[\partial_{(L)}G_0(0, 0)])^{-1} \\ &\quad \times \left([\partial_{(L)}G_0(0, x')] - \alpha[\partial_{(L)}G_0(0, 0)]G(0, x') - \beta[\partial_{(L,R)}^2G_0(0, 0)]G(0, x') \right). \end{aligned} \quad (2.83)$$

Employing this result in (2.79), we have

$$\begin{aligned} G(0, x') &= G_0(0, x') - \alpha G_0(0, 0)G(0, x') - \beta[\partial_{(R)}G_0(0, 0)]G(0, x') \\ &\quad - \beta G_0(0, 0)(1 + \beta[\partial_{(L)}G_0(0, 0)])^{-1} \{ [\partial_{(L)}G_0(0, x')] - \alpha[\partial_{(L)}G_0(0, 0)]G(0, x') \\ &\quad - \beta[\partial_{(L,R)}^2G_0(0, 0)]G(0, x') \}, \end{aligned} \quad (2.84)$$

which yields $G(0, x')$ as

$$\begin{aligned} G(0, x') &= \left\{ 1 + \alpha G_0(0, 0) + \beta[\partial_{(R)}G_0(0, 0)] - \beta G_0(0, 0)(1 + \beta[\partial_{(L)}G_0(0, 0)])^{-1} \right. \\ &\quad \times \left(\alpha[\partial_{(L)}G_0(0, 0)] + \beta[\partial_{(L,R)}^2G_0(0, 0)] \right) \Big\}^{-1} \\ &\quad \times \left[G_0(0, x') - \beta G_0(0, 0)(1 + \beta[\partial_{(L)}G_0(0, 0)])^{-1} [\partial_{(L)}G_0(0, x')] \right]. \end{aligned} \quad (2.85)$$

Equation (2.85) may now be substituted into the right side of (2.83) to obtain $[\partial_{(L)}G(0, x')]$ in terms of G_0 and its derivatives alone. Finally, the substitution of these results for $G(0, x')$ and $[\partial_{(L)}G(0, x')]$ as indicated above into the right side of (2.78) yields the full, *exact* Green's function for the highly singular 1-D potential $U(x)$ of (2.70) joined onto any relatively well behaved 1-D potential $V(x)$, such as a harmonic oscillator and/or a constant electric field potential independent of time.

For the special case of $\beta = 0$, we obtain the corresponding result for the $-\delta(x)$ -potential with *no* $-\delta'(x)$ -potential.

$$G(x, x') = G_0(x, x') - \frac{\alpha G_0(x, 0)G_0(0, x')}{1 + \alpha G_0(0, 0)}. \quad (2.86)$$

To examine the role of the highly singular potentials in the 1-D Green's function, we rewrite (2.72) in the form

$$G_0(x, x') = C(x') \{ \eta_+(x' - x)y_1(x)y_2(x') + \eta_+(x - x')y_2(x)y_1(x') \} \quad (2.87)$$

where $\eta_+(x)$ is the Heaviside unit step function and $C(x') \equiv m(x')/A(y_1, y_2)$. Since $\eta_+(0) = 1/2$,

$$G_0(0, 0) = C(0) \left\{ \frac{1}{2}y_1(0)y_2(0) + \frac{1}{2}y_2(0)y_1(0) \right\} = C(0)y_1(0)y_2(0). \quad (2.88)$$

Differentiating (2.87) to form $[\partial_{(L)}G_0(x, x')]$, there is a cancellation of terms involving Dirac delta functions arising from $\partial_{(L)}\eta_+(x - x')$, etc., and we obtain

$$[\partial_{(L)}G_0(x, x')] = C(x') \{ \eta_+(x' - x)y_1'(x)y_2(x') + \eta_+(x - x')y_2'(x)y_1(x') \}. \quad (2.89)$$

Differentiating again to form $[\partial_{(L,R)}^2G(x, x')]$, we have $(C'(x') \equiv d(C(x'))/dx')$, etc.)

$$\begin{aligned} [\partial_{(L,R)}^2G_0(x, x')] &= C'(x') \{ \eta_+(x' - x)y_1'(x)y_2(x') + \eta_+(x - x')y_2'(x)y_1(x') \} \\ &\quad + C(x') \left\{ \begin{aligned} &\delta(x' - x)y_1'(x)y_2(x') + \eta_+(x' - x)y_1'(x)y_2'(x') \\ &-\delta(x - x')y_2'(x)y_1(x') + \eta_+(x - x')y_2'(x)y_1'(x') \end{aligned} \right\}, \end{aligned} \quad (2.90)$$

which may be written more compactly using the definition of the Wronskian, $A_{x'}(y_1, y_2)$, as (subscript “ x' ” denotes evaluation at x')

$$\begin{aligned} [\partial_{(L,R)}^2G_0(x, x')] &= -C(x')\delta(x' - x)A_{x'}(y_1, y_2) \\ &\quad + C(x')[\eta_+(x' - x)y_1'(x)y_2'(x') + \eta_+(x - x')y_2'(x)y_1'(x')] \\ &\quad + C'(x')\{ \eta_+(x' - x)y_1'(x)y_2(x') + \eta_+(x - x')y_2'(x)y_1(x') \}. \end{aligned} \quad (2.91)$$

Clearly, the first term on the right of (2.91) shows that

$$\left| \left[\partial_{(L,R)}^2 G_0(0,0) \right] \right| \rightarrow \infty. \quad (2.92)$$

In view of the huge value of $\left| \left[\partial_{(L,R)}^2 G_0(0,0) \right] \right|$, we may write (2.85) for $\beta \neq 0$ as

$$G(0, x') = - \frac{1 + \beta [\partial_{(L)} G_0(0,0)]}{\beta^2 G_0(0,0) [\partial_{(L,R)}^2 G_0(0,0)]} \left\{ G_0(0, x') - \frac{\beta G_0(0,0) [\partial_{(L)} G_0(0, x')]}{1 + \beta [\partial_{(L)} G_0(0,0)]} \right\}, \quad (2.93)$$

and considering particle transmission through the highly singular potential at the origin (with $x < 0$ and $x' > 0$), this means that (2.92), (2.93)

$$G(0, x') = 0. \quad (\beta \neq 0) \quad (2.94)$$

On the same basis, (2.83) for $[\partial_{(L)} G(0, x')]$ may be written for $\beta \neq 0$ as

$$\begin{aligned} [\partial_{(L)} G(0, x')] &= D^{-1} [\partial_{(L)} G_0(0, x')] \\ &+ \frac{1}{\beta G_0(0,0)} \{ G_0(0, x') - \beta D^{-1} G_0(0,0) [\partial_{(L)} G_0(0, x')] \}, \end{aligned} \quad (2.95)$$

where we have defined D as

$$D = 1 + \beta [\partial_L G_0(0,0)]. \quad (2.96)$$

Here, we see that the role of $\Delta U = -\alpha \delta(x)$ as part of the generalized potential U is negligible in $G(x, x')$ for well behaved potentials $V(x)$ so long as $\beta \neq 0$. Employing (2.93) and (2.95), we find $G(x, x')$ of (2.78) for $\beta \neq 0$ as

$$\begin{aligned} G(x, x') &= G_0(x, x') - \beta G_0(x, 0) (D^{-1} [\partial_{(L)} G_0(0, x')]) \\ &+ \frac{1}{\beta G_0(0,0)} \{ G_0(0, x') - \beta D^{-1} G_0(0,0) [\partial_{(L)} G_0(0, x')] \}, \end{aligned}$$

or

$$G(x, x') = G_0(x, x') - \frac{G_0(x, 0) G_0(0, x')}{G_0(0,0)}. \quad (2.97)$$

Considering that the Fourier transform of $G(x, x'; \omega)$ to direct time representation represents the quantum mechanical amplitude for a Schrödinger particle (wave packet) to be transmitted from position x' at time t' to x at a later time t , such that

$$\Psi_{out}(x, t) = \int \frac{d\omega}{2\pi} e^{-i\omega(t-t')} \int_{-\infty}^{\infty} dx' G(x, x'; \omega) \Psi_{in}(x', t'), \quad (2.98)$$

it is clear that an electron wave packet, $\Psi_{in}(x' < 0, t')$ in the region of incidence, $x' < 0$, cannot be transmitted to $\Psi_{out}(x > 0, t)$ in the outgoing wave region, $x > 0$, on the other side of the highly singular potential $\beta\delta'(x)$ because (2.97) and (2.72) jointly yield a vanishing result

$$G(x > 0, x' < 0; \omega) = C(x') \left[y_2(x)y_1(x') - \frac{y_2(x)y_1(0)y_2(0)y_1(x')}{y_1(0)y_2(0)} \right] = 0. \quad (2.99)$$

This very general result means that there is *no* possibility of particle transmission from $x' < 0$ through the $\beta\delta'(x)$ -potential, even in the presence of well behaved potentials $V(x)$ as well as the presence of $\Delta U = -\alpha\delta(x)$, provided $\beta \neq 0$.

References

1. V. Fessatidis, N.J.M. Horing, K. Sabeeh, *Philos. Mag.* **B79**, 77 (1999)
2. A. Erdelyi, *Higher Transcendental Functions*. vol. II, Eq. (17), p. 189. (McGraw-Hill, New York, 1953)
3. N.J.M. Horing, *Quantum Statistical Field Theory: Schwinger's Variational Method*. (Oxford University Press, to be published)
4. A. Erdelyi, *Higher Transcendental Functions*, vol. II, Eq. (33), p. 8. (McGraw-Hill, New York, 1953)
5. N. Horing, M. Yildiz, *Ann. Phys. (NY)* **97**, 216–241 (1976)
6. A. Erdelyi, *Higher Transcendental Functions*, vol. II, Eq. (32), p. 195. (McGraw-Hill, New York, 1953)
7. A. Erdelyi, *Tables of Integral Transforms*, vol. I. (McGraw-Hill, New York, 1954)
8. I.S. Gradshteyn, I.M. Ryzhik, *Tables of Integrals, Series and Products* (Academic, New York, 1965)
9. P.B. Visscher, L.M. Falicov, *Phys. Rev. B* **3**, 2541 (1971)
10. D. Grecu, *Phys. Rev. B* **8**, 1958 (1972)
11. A.L. Fetter, *Ann. Phys. (N. Y.)* **88**, 1 (1974)
12. M. Apostol, *Z. Phys. B* **22**, 13 (1975)
13. M. Kobayashi et al., *J. Phys. Soc. Jpn.* **39**, 18 (1975)
14. J. Mizuno et al., *J. Phys. Soc. Jpn.* **39**, 983 (1975)
15. A. Caille et al., *Solid State Commun.* **24**, 805 (1977)
16. L. Bloss and E. M., Brody, *Solid State Commun.* **43**, 523 (1982)
17. W.L. Bloss, *Solid State Commun.* **44**, 363 (1982)
18. D. Olego et al., *Phys. Rev. B* **25**, 7867 (1982)
19. P. Ruden, G.H. Dohler, *Phys. Rev. B* **27**, 3538 (1983)
20. P. Ruden, G.H. Dohler, *Phys. Rev. B* **27**, 3547 (1983)
21. S. Das Sarma, A. Madhukar, *Phys. Rev. B* **23**, 805 (1981); S. Das Sarma, J.J. Quinn, *ibid.* **25**, 7603 (1982); **27**, 6516 (1983)
22. S. Das Sarma, *Phys. Rev. B* **28**, 2240 (1983); *Appl. Surf. Sci.* **11/12**, 535 (1982)
23. G.G. de la Cruz et al., *J. Phys. Chem. Solids* **44**, 807 (1983)

24. A.C. Tselis, J.J. Quinn, Phys. Rev. B **29**, 2021 (1984); **29**, 3318 (1984)
25. P.L. Christiansen, H.C. Arnbak, A.V. Zolotaryuk, V.N. Ermakov, Y.B. Gaididei, J. Phys. A: Math. Gen. **36**, 7589 (2003)
26. F.A.B. Coutinho, Y. Nogami, J. Fernando Perez; J Phys. A. Math. Gen. **30**, 3937 (1997)
27. Y. Golovaty; [arXiv:1201.2610v1](https://arxiv.org/abs/1201.2610v1) [math.SP] 12 Jan 2012
28. P. Šeba, Rep. Math. Phys. **24**, 111 (1986)
29. F.M. Toyama, Y. Nogami, J. Phys. A: Math. Theor. **40**, F685 (2007)
30. M. Gadella, M.L. Glasser, L.M. Nieto, Int. J. Theor. Phys. **50**, 2144 (2011)
31. M. Gadella, J. Negro, L.M. Nieto, Phys. Lett. A **373**, 1310 (2009)
32. M. Gadella, M.L. Glasser, L.M. Nieto, Int. J. Theor. Phys. **50**, 2191 (2011)
33. J.J. Alvarez, M. Gadella, M.L. Glasser, L.P. Lara, L.M. Nieto, J. Phys: Conf. Ser. **284**, 012009 (2011)
34. A.V. Zolotaryuk, Y. Zolotaryuk, [arXiv:1202.1117v1](https://arxiv.org/abs/1202.1117v1) [math-ph] 6 Feb 2012
35. J.F. Brasche, L. Nizhnik, [arXiv:1112.2545v1](https://arxiv.org/abs/1112.2545v1) [math.FA] 12 Dec 2011
36. P. Kurasov, Math. Anal. Appl. **201**, 297 (1996)
37. S. Kocinak, V. Milanovic, Mod. Phys. Lett. B **26**, 1250092 (2012)
38. D.K. Park, J. Phys. A Math.: Gen. **29**, 6407 (1996)
39. Haydar Uncu, Devrim Tarhan, Ersan Demiralp, Özgür E. Müstecaplıoğlu, Phys. Rev. A **76**, 013618 (2007)
40. N.J.M. Horing, unpublished
41. P.M. Morse, H. Feshbach, *Methods of Theoretical Physics*, vol. 1. (McGraw-Hill, New York, 832, 1953)
42. F. Garcia-Moliner, F. Flores, *Introduction to the Theory of Solid Surfaces*. (Cambridge University Press, Cambridge, 1979)
43. F. Garcia-Moliner, V.R. Velasco, *Theory of Single and Multiple Interfaces*. (World Scientific, Singapore, 1992)
44. S.G. Davison, J.D. Levine, *Solid State Physics*, vol. 25, eds. by H. Ehrenreich, F. Seitz, D. Turnbull (Academic Press, New York, 1970)
45. S.G. Davison, M. Steslicka, *Basic Theory of Surface States* (Clarendon Press, Oxford, 1992)
46. J. Inglesfield, J. Phys. C; Solid State Phys. **4**, L 14, (1971)

Chapter 3

Wave Propagation and Diffraction Through a Subwavelength Nano-Hole in a 2D Plasmonic Screen

Norman J.M. Horing and Désiré Miessein

Abstract This article addresses the subject of subwavelength transmission and diffraction by a two dimensional plasmonic screen/layer with an embedded nano-hole aperture. The dynamical development of a prototypical scalar wave (incident at an arbitrary angle) is traced in terms of the associated Helmholtz Green's function. The latter is derived through the exact analytic solution of a succession of integral equations which obviate the need to specify boundary conditions separately. The resulting wave transmission/diffraction is evaluated numerically and exhibited graphically: It includes the finite plasmonic response of the 2D layer jointly with that of the subwavelength nano-hole aperture, for an arbitrary angle of incidence.

3.1 Introduction

Wave transmission and diffraction in the presence of a subwavelength nano-hole in a screen (radius $R \ll \lambda$) has been the subject of extensive investigation in recent years due to a variety of interesting optical phenomena [1–12]. The properties of the screen are often represented by metallic boundary conditions, which, however, may not accurately describe its plasmonic polarization response. In this article we discuss the transmission/diffraction of a prototypical scalar field incident on a thin two dimensional plasmonic layer, having well defined finite polarizability, containing a subwavelength aperture. The two dimensional layer is taken to be at the plane $z = 0$, embedded in a 3D bulk host medium having background dielectric constant $\epsilon_b^{(3D)}$ with the aperture at its center, $x = y = 0$. Employing a succession of integral

N.J.M. Horing (✉) · D. Miessein

Department of Physics and Engineering Physics, Stevens Institute of Technology, Hoboken,
NJ 07030, USA

e-mail: nhoring@stevens.edu

D. Miessein

e-mail: dmiessei@stevens.edu

equations for the scalar Green's function of the associated Helmholtz problem describing the evolution of the wave in the transmission/diffraction process, the need for boundary conditions is obviated, and explicit analytic solutions are obtained that accurately account for both the presence of the subwavelength nano-hole and the finite plasmonic response of the layer.

In Sect. 3.2, the scalar Helmholtz Green's function of the two dimensional plasmonic layer containing a nano-hole embedded in a 3D bulk host medium is obtained through the analytic solutions of a succession of integral equations. The transmission/diffraction of a plane scalar wave at an arbitrary angle of incidence on the perforated 2D layer is discussed in Sect. 3.3. The numerical analysis of Helmholtz wave propagation in the vicinity of the plasmonic layer containing a nano-hole, based on the analytic results obtained in Sect. 3.3, is presented in detail in the figures of Sect. 3.4. A summary of conclusions is contained in Sect. 3.5.

3.2 Scalar Green's Function for a 2D Plasmonic Layer with a Nano-Hole Aperture Embedded in a 3D Bulk Host Medium

Considering a two dimensional (2D) plasmonic layer on the plane $z = 0$, embedded in a three dimensional bulk host medium with background dielectric constant $\epsilon_b^{(3D)}$, the scalar Helmholtz Green's function of the system G_{fs} satisfies the equation (ω —frequency representation) [13–15]

$$\left(-\nabla^2 - \frac{\omega^2}{c^2} \epsilon_b^{(3D)}\right) G_{fs}(\mathbf{r}, \mathbf{r}'; \omega) - \frac{\omega^2}{c^2} \int d^3 \mathbf{r}' \alpha_{fs}^{(2D)}(\mathbf{r}, \mathbf{r}''; \omega) G_{fs}(\mathbf{r}'', \mathbf{r}'; \omega) = \delta^{(3)}(\mathbf{r} - \mathbf{r}'), \quad (3.1)$$

where the polarizability $\alpha_{fs}^{(2D)}$ of the 2D plasmonic layer has the form

$$\alpha_{fs}^{(2D)}(\mathbf{r}, \mathbf{r}'; \omega) = \alpha_{fs}^{(2D)}(\mathbf{r}_{\parallel}, \mathbf{r}'_{\parallel}; \omega) d \delta(z) \delta(z'). \quad (3.2)$$

Here, d is the thickness of the plasmonic sheet, $\mathbf{r} = (\mathbf{r}_{\parallel}; z)$ and $\alpha_{fs}^{(2D)}(\mathbf{r}_{\parallel}, \mathbf{r}'_{\parallel}; \omega)$ is the 2D plasmonic polarizability on the 2D sheet and $\delta(z)$ is the Dirac delta function needed to lock the polarizability onto the plane of the 2D layer at $z = 0$. To solve for $G_{fs}(\mathbf{r}, \mathbf{r}'; \omega)$, we employ the bulk Helmholtz Green's function, $G_{3D}(\mathbf{r}, \mathbf{r}'; \omega)$, given by

$$\left(-\nabla^2 - \frac{\omega^2}{c^2} \epsilon_b^{(3D)}\right) G_{3D}(\mathbf{r}, \mathbf{r}'; \omega) = \delta^{3D}(\mathbf{r} - \mathbf{r}'), \quad (3.3)$$

in lateral wavevector representation ($\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel} \mapsto \mathbf{k}_{\parallel}$),

$$G_{3D}(\mathbf{k}_{\parallel}; z, z'; \omega) = -\frac{e^{ik_{\perp}|z-z'|}}{2ik_{\perp}} \quad (3.4)$$

where $k_{\perp} = \pm\sqrt{q_{\omega}^2 - k_{\parallel}^2}$ and $q_{\omega} = (\omega/c)\sqrt{\epsilon_b^{(3D)}}$.

The Helmholtz Green's function including the plasmonic sheet, $G_{fs}(\mathbf{r}, \mathbf{r}'; \omega)$, satisfies the integral equation

$$\begin{aligned} G_{fs}(\mathbf{r}, \mathbf{r}'; \omega) &= G_{3D}(\mathbf{r}, \mathbf{r}'; \omega) \\ &+ \frac{\omega^2}{c^2} \int d^3\mathbf{r}'' \int d^3\mathbf{r}''' G_{3D}(\mathbf{r}, \mathbf{r}''; \omega) \alpha_{fs}^{2D}(\mathbf{r}'', \mathbf{r}'''; \omega) G_{fs}(\mathbf{r}''', \mathbf{r}'; \omega). \end{aligned} \quad (3.5)$$

Fourier transforming this in the lateral plane of translational invariance ($\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel} \rightarrow \mathbf{k}_{\parallel}$) and employing (3.2), $G_{3D}(\mathbf{k}_{\parallel}; 0, z'; \omega)$ is readily obtained algebraically and leads to G_{fs} as

$$\begin{aligned} G_{fs}(\mathbf{k}_{\parallel}; z, z'; \omega) &= G_{3D}(\mathbf{k}_{\parallel}; z, z'; \omega) \\ &+ \frac{\omega^2}{c^2} d \frac{G_{3D}(\mathbf{k}_{\parallel}; z, 0; \omega) \alpha_{fs}^{(2D)}(\mathbf{k}_{\parallel}; \omega) G_{3D}(\mathbf{k}_{\parallel}; 0, z'; \omega)}{1 - \frac{\omega^2}{c^2} d \alpha_{fs}^{(2D)}(\mathbf{k}_{\parallel}; \omega) G_{3D}(\mathbf{k}_{\parallel}; 0, 0; \omega)}. \end{aligned} \quad (3.6)$$

Using (3.4), this yields

$$G_{fs}(\mathbf{k}_{\parallel}; z, z'; \omega) = -\frac{e^{ik_{\perp}|z-z'|}}{2ik_{\perp}} + \frac{\gamma e^{ik_{\perp}(|z|+|z'|)}}{2ik_{\perp}(2ik_{\perp} + \gamma)}, \quad (3.7)$$

where $\gamma = \frac{\omega^2}{c^2} d \alpha_{fs}^{(2D)}(\mathbf{k}_{\parallel}; \omega)$. For metallic densities, $\alpha_{fs}^{(2D)}(\mathbf{k}_{\parallel}; \omega)$ tends to negative infinity and γ also approaches negative infinity, so (3.7) reduces to

$$G_{fs}(\mathbf{k}_{\parallel}; z, z'; \omega) = -\left(\frac{e^{ik_{\perp}|z-z'|} - e^{ik_{\perp}(|z|+|z'|)}}{2ik_{\perp}} \right), \quad (3.8)$$

which indicates perfect reflection and no transmission of an impinging scalar wave.

It is useful to represent $G_{fs}(\mathbf{r}_{\parallel}, 0; z, 0; \omega)$ in lateral position representation: with $z' = 0$, (3.7) yields

$$G_{fs}(\mathbf{k}_{\parallel}; z, 0; \omega) = -\frac{e^{ik_{\perp}|z|}}{(2ik_{\perp} + \gamma)}. \quad (3.9)$$

Bearing in mind that

$$\gamma = \frac{\omega^2 d}{c^2} \alpha_{fs}^{(2D)}(\mathbf{k}_{\parallel}; \omega) \quad (3.10)$$

and

$$\alpha_{fs}^{(2D)}(\mathbf{k}_{\parallel}; \omega) = \left(\frac{2\pi i e^2 n_{2D}}{m^* \omega^2} \right) \sqrt{q_{\omega}^2 - k_{\parallel}^2}, \quad (3.11)$$

is the polarizability of the 2D plasmonic layer (with n_{2D} as the 2D equilibrium density and m^* is effective mass), we have

$$\begin{aligned} G_{fs}(\mathbf{k}_{\parallel}; z, 0; \omega) &= -\frac{1}{1+\mu} \frac{e^{ik_{\perp}|z|}}{2ik_{\perp}} \\ &= \frac{1}{1+\mu} G_{3D}(\mathbf{k}_{\parallel}; z, 0; \omega), \end{aligned} \quad (3.12)$$

or, in position representation,

$$G_{fs}(\mathbf{r}_{\parallel}, 0; z, 0; \omega) = \frac{1}{4\pi[1+\mu]} \frac{e^{iq_{\omega}\sqrt{r_{\parallel}^2 + |z|^2}}}{\sqrt{r_{\parallel}^2 + |z|^2}}, \quad (3.13)$$

where $\mu = \pi e^2 d n_{2D} / m^* c^2$.

Addressing the problem of a two dimensional plasmonic layer S_1 perforated by a nano-scale aperture S_2 of radius R at $\mathbf{r} = 0$ in the $(x - y)$ plane, the full sheet polarizability is modified by subtracting the part associated with the hole (Fig. 3.1)

$$\alpha^{(2D)}(\mathbf{r}, \mathbf{r}'; \omega) = \alpha_{fs}^{(2D)}(\mathbf{r}, \mathbf{r}'; \omega) - \alpha_h^{(2D)}(\mathbf{r}, \mathbf{r}'; \omega) \quad (3.14)$$

where $\alpha_h^{(2D)}(\mathbf{r}, \mathbf{r}'; \omega)$ is the part of the layer polarizability excised from the nano-hole. The scalar Green's function for this system, including the aperture, satisfies the integral equation given by:

$$\begin{aligned} G(\mathbf{r}, \mathbf{r}'; \omega) &= G_{fs}(\mathbf{r}, \mathbf{r}'; \omega) \\ &\quad - \frac{\omega^2}{c^2} \int d^3 \mathbf{r}'' \int d^3 \mathbf{r}''' G_{fs}(\mathbf{r}, \mathbf{r}''; \omega) \alpha_h^{2D}(\mathbf{r}'', \mathbf{r}'''; \omega) G(\mathbf{r}''', \mathbf{r}'; \omega). \end{aligned} \quad (3.15)$$

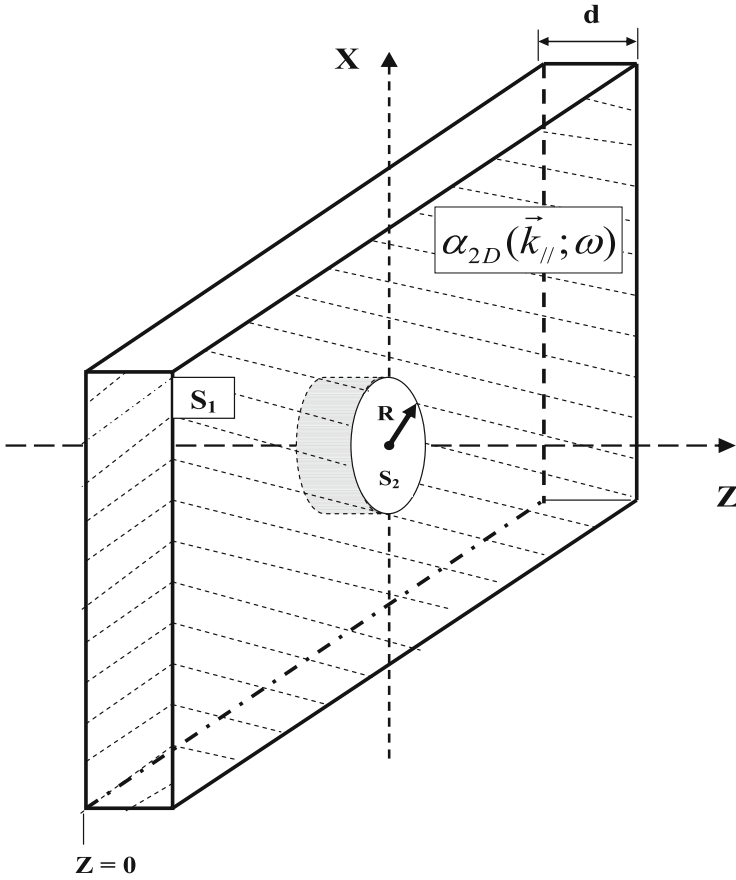


Fig. 3.1 Two dimensional plasmonic layer (thickness d , embedded at $z = 0$ in a three dimensional bulk host medium) with a nano-hole of radius R at the origin of the $(x - y)$ plane

Here, the polarizability part excised from the nano-hole is defined as

$$\begin{aligned}
 \alpha_h^{(2D)}(\mathbf{r}, \mathbf{r}'; \omega) &= \alpha_{fs}^{(2D)}(x, y, x', y'; \omega) d \delta(z) \delta(z') \\
 &\times \eta_+ \left(\frac{a}{2} - |x| \right) \eta_+ \left(\frac{b}{2} - |y| \right) \\
 &\times \eta_+ \left(\frac{a}{2} - |x'| \right) \eta_+ \left(\frac{b}{2} - |y'| \right),
 \end{aligned} \tag{3.16}$$

where $\eta_+ \left(\frac{a}{2} - |x| \right)$ is the Heaviside unit step function representing a cut-off function imposed to confine the integration range on the 2D sheet to the nano-hole dimension ($R \sim a$) and $\delta(z)$ is the Dirac delta function needed to lock the

polarizability onto the plane of the 2D plasmonic layer. A simple approximation of (3.16) for very small radius leads to $\eta_+ \left(\frac{a}{2} - |x|\right) \mapsto a \delta(x)$ and

$$\alpha_h^{(2D)}(\mathbf{r}, \mathbf{r}'; \omega) \approx A^2 \alpha_{fs}^{(2D)}(\mathbf{r}_{\parallel}, \mathbf{r}'_{\parallel}; \omega) \delta^2(\mathbf{r}_{\parallel}) \delta^2(\mathbf{r}'_{\parallel}) d \delta(z) \delta(z'), \quad (3.17)$$

(A is nano-hole area) and with this, (3.15) yields

$$G(\mathbf{r}_{\parallel}, \mathbf{r}'_{\parallel}; z, z'; \omega) = G_{fs}(\mathbf{r}_{\parallel}, \mathbf{r}'_{\parallel}; z, z'; \omega) - \beta G_{fs}(\mathbf{r}_{\parallel}, 0; z, 0; \omega) G(0, \mathbf{r}'_{\parallel}; 0, z'; \omega), \quad (3.18)$$

where $\beta = \gamma A^2$ with $\gamma = \frac{\omega^2 d}{c^2} \alpha_{fs}^{(2D)}([\mathbf{r}_{\parallel} = 0], [\mathbf{r}'_{\parallel} = 0]; \omega)$. Equation (3.18) may be solved algebraically by setting $\mathbf{r}_{\parallel} = 0$ and $z = 0$, with the closed form analytic result

$$G(\mathbf{r}_{\parallel}, \mathbf{r}'_{\parallel}; z, z'; \omega) = G_{fs}(\mathbf{r}_{\parallel}, \mathbf{r}'_{\parallel}; z, z'; \omega) - \frac{\beta G_{fs}(\mathbf{r}_{\parallel}, 0; z, 0; \omega) G_{fs}(0, \mathbf{r}'_{\parallel}; 0, z'; \omega)}{1 + \beta G_{fs}([\mathbf{r}_{\parallel} = 0] - [\mathbf{r}'_{\parallel} = 0]; z = 0, z' = 0; \omega)}. \quad (3.19)$$

Anticipating that the transmitted scalar wave is controlled by $G(\mathbf{r}_{\parallel}, 0; z, 0; \omega)$, setting $\mathbf{r}'_{\parallel} = 0$ and $z' = 0$ yields

$$G(\mathbf{r}_{\parallel}, 0; z, 0; \omega) = \frac{G_{fs}(\mathbf{r}_{\parallel}, 0; z, 0; \omega)}{1 + \beta G_{fs}([\mathbf{r}_{\parallel} = 0] - [\mathbf{r}'_{\parallel} = 0]; z = 0, z' = 0; \omega)}. \quad (3.20)$$

which involves a divergent integral in the denominator when all its positional arguments vanish. Specifically, rewriting (3.11) in lateral position representation with $\mathbf{r}_{\parallel} = 0$, $\mathbf{r}'_{\parallel} = 0$, and $z = 0$, $z' = 0$, the resulting $k_{\parallel}$ -integral is divergent at the upper limit,

$$\alpha_{fs}^{(2D)}(0, 0; \omega) = \frac{1}{2\pi} \left(\frac{2\pi i e^2 n_{2D}}{m^* \omega^2} \right) \int_0^{\infty} dk_{\parallel} k_{\parallel} \sqrt{q_{\omega}^2 - k_{\parallel}^2}, \quad (3.21)$$

and similar remarks apply to $G_{fs}(0, 0; 0, 0; \omega)$:

$$G_{fs}(0, 0; 0, 0; \omega) = \frac{i}{4\pi[1 + \mu]} \left\{ \int_0^{q_{\omega}} \frac{dk_{\parallel} k_{\parallel}}{\sqrt{q_{\omega}^2 - k_{\parallel}^2}} - i \int_{q_{\omega}}^{\infty} \frac{dk_{\parallel} k_{\parallel}}{\sqrt{k_{\parallel}^2 - q_{\omega}^2}} \right\}. \quad (3.22)$$

These divergencies are artifacts of confining the aperture to zero radial size in (3.17), and recognizing that the integral equation, (3.15), is actually extended over a finite range R , it is necessary to adjust the result to accommodate a “smear” average in which $|\mathbf{r}_{\parallel}|, |\mathbf{r}'_{\parallel}| \mapsto 0$ is replaced by $0 \mapsto R$; alternatively, in Fourier lateral

$k_{\parallel}$ -integrations, the range is restricted to a large but finite range of $k_{\parallel} < \frac{1}{R}$, with frequencies in the range $q_{\omega}R < 1$ for nano-holes of subwavelength dimensions, $R < \lambda \sim \frac{1}{q_{\omega}}$. The resulting integrations are elementary, yielding

$$\alpha_{fs}^{(2D)}([\mathbf{r}_{\parallel} = 0], [\mathbf{r}'_{\parallel} = 0]; \omega) \approx \frac{1}{6\pi R^3} \left(\frac{2\pi i e^2 n_{2D}}{m^* \omega^2} \right) \left[(q_{\omega}R)^3 + i \left[1 - (q_{\omega}R)^2 \right]^{3/2} \right] \quad (3.23)$$

and

$$\gamma = \frac{\mu}{3\pi R^3} \left[i (q_{\omega}R)^3 - \left[1 - (q_{\omega}R)^2 \right]^{3/2} \right]. \quad (3.24)$$

Similarly,

$$G_{fs}([\mathbf{r}_{\parallel} = 0] - [\mathbf{r}'_{\parallel} = 0]; z = 0, z' = 0; \omega) = \frac{1}{4\pi(1+\mu)R} \left[i q_{\omega}R + \sqrt{1 - (q_{\omega}R)^2} \right]. \quad (3.25)$$

3.3 Scalar Field Response of a Perforated 2D Plasmonic Layer to an Incident Wave $U_0(\mathbf{r}, \omega)$: Transmission

The model scalar field system employed here, which is governed by Helmholtz dynamics, is driven by a source density $\rho(\mathbf{r}; t)$ which is the sum of the induced density, $\rho_{ind}(\mathbf{r}; t)$, and an external density, $\rho_{ext}(\mathbf{r}; t)$, in accordance with the inhomogeneous equation (Fig. 3.2)

$$\left(-\nabla^2 - \frac{\omega^2}{c^2} \varepsilon_b^{(3D)} \right) U(\mathbf{r}; \omega) - \frac{\omega^2}{c^2} \int d^3 \mathbf{r}' \alpha^{(2D)}(\mathbf{r}, \mathbf{r}'; \omega) U(\mathbf{r}'; \omega) = 4\pi \rho_{ext}(\mathbf{r}; \omega). \quad (3.26)$$

Here, $\alpha^{(2D)}(\mathbf{r}, \mathbf{r}'; \omega)$ is the polarizability induced in the perforated 2D layer, $\alpha^{(2D)}(\mathbf{r}, \mathbf{r}'; \omega) = \alpha_{fs}^{(2D)}(\mathbf{r}, \mathbf{r}'; \omega) - \alpha_h^{(2D)}(\mathbf{r}, \mathbf{r}'; \omega)$ (3.14), (3.17). Comparison of (3.26) with (3.1) immediately shows that

$$U(\mathbf{r}, \omega) = 4\pi \int d^3 \mathbf{r}' G(\mathbf{r}, \mathbf{r}'; \omega) \rho_{ext}(\mathbf{r}'; \omega), \quad (3.27)$$

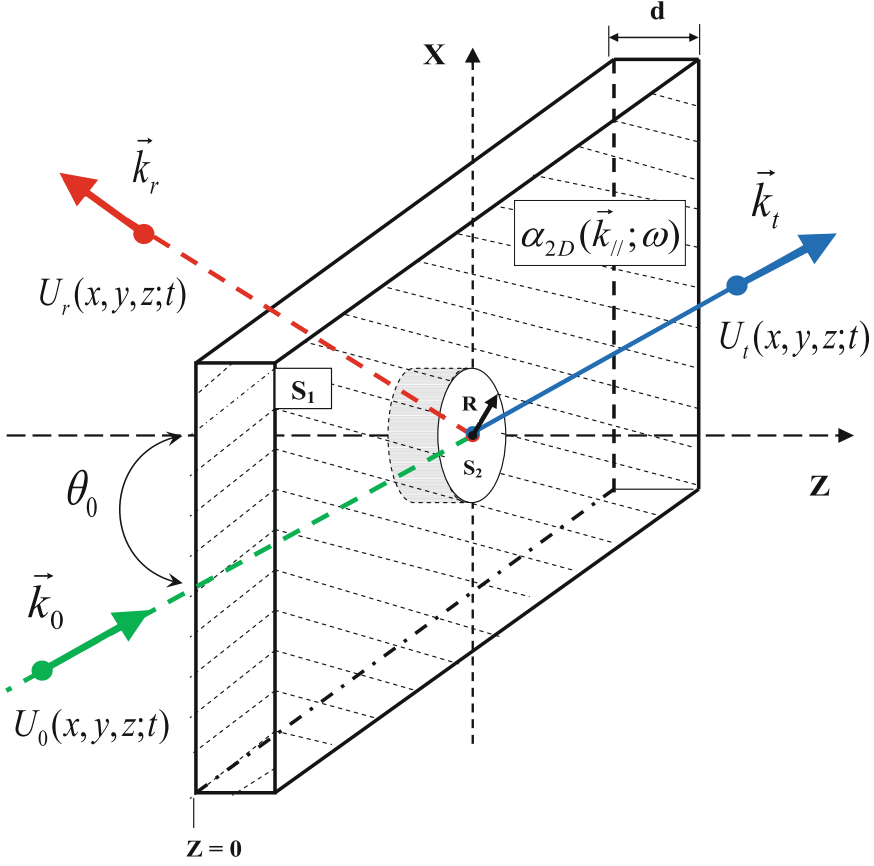


Fig. 3.2 Two dimensional plasmonic layer (thickness d , embedded at $z = 0$ in a three dimensional bulk host medium) with a nano-hole of radius R at the origin of the $(x - y)$ plane, shown with incident, reflected and transmitted wave vectors ($\vec{k}_0, \vec{k}_r, \vec{k}_t$) for waves ($U_0(x, y, z; t)$, $U_r(x, y, z; t)$, $U_t(x, y, z; t)$)

which is typical for linear systems. To connect this with the incident field $U_0(\mathbf{r}; \omega)$ generated by a distant source $\rho_{ext}(\mathbf{r}; \omega)$, we follow the techniques of [13–15] noting that

$$U_0(\mathbf{r}, \omega) = 4\pi \int d^3 \mathbf{r}' G_{3D}(\mathbf{r}, \mathbf{r}'; \omega) \rho_{ext}(\mathbf{r}'; \omega), \quad (3.28)$$

or, in positional matrix notation

$$U_0 = 4\pi G_{3D} \rho_{ext}, \quad (3.29)$$

so that

$$U = GG_{3D}^{-1} U_0. \quad (3.30)$$

By the same token [13–15], (3.15) may be written as the positional integral (matrix) equation

$$G = G_{fs} - \frac{\omega^2}{c^2} G_{fs} \alpha_h^{(2D)} G, \quad (3.31)$$

or, multiplying by G^{-1} from the right and G_{fs}^{-1} from the left,

$$G^{-1} = G_{fs}^{-1} + \frac{\omega^2}{c^2} \alpha_h^{(2D)}. \quad (3.32)$$

Alternatively, analogous treatment of (3.5) yields

$$G_{fs}^{-1} = G_{3D}^{-1} - \frac{\omega^2}{c^2} \alpha_{fs}^{(2D)}, \quad (3.33)$$

so that

$$\begin{aligned} G^{-1} &= G_{3D}^{-1} - \frac{\omega^2}{c^2} [\alpha_{fs}^{(2D)} - \alpha_h^{(2D)}] \\ &= G_{fs}^{-1} + \frac{\omega^2}{c^2} \alpha_h^{(2D)}. \end{aligned} \quad (3.34)$$

Consequently, U may be written as

$$U = GG_{3D}^{-1} U_0 = U_0 + \frac{\omega^2}{c^2} G \alpha_{fs}^{(2D)} U_0 - \frac{\omega^2}{c^2} G \alpha_h^{(2D)} U_0. \quad (3.35)$$

In explicit position/frequency representation, this result reads as

$$\begin{aligned} U(\mathbf{r}_{\parallel}; z; \omega) &= U_0(\mathbf{r}_{\parallel}; z; \omega) + \frac{\omega^2}{c^2} \int d^2 \mathbf{r}'_{\parallel} \int d^2 \mathbf{r}''_{\parallel} \int dz' \int dz'' G(\mathbf{r}_{\parallel}, \mathbf{r}'_{\parallel}; z, z'; \omega) \\ &\quad \times \alpha_{fs}^{2D}(\mathbf{r}'_{\parallel}, \mathbf{r}''_{\parallel}; z', z''; \omega) U_0(\mathbf{r}_{\parallel}; z''; \omega) \\ &\quad - \frac{\omega^2}{c^2} \int d^2 \mathbf{r}'_{\parallel} \int d^2 \mathbf{r}''_{\parallel} \int dz' \int dz'' G(\mathbf{r}_{\parallel}, \mathbf{r}'_{\parallel}; z, z'; \omega) \\ &\quad \times \alpha_h^{2D}(\mathbf{r}'_{\parallel}, \mathbf{r}''_{\parallel}; z', z''; \omega) U_0(\mathbf{r}_{\parallel}; z''; \omega). \end{aligned} \quad (3.36)$$

Bearing in mind that α_{fs}^{2D} involves $\delta(z')\delta(z'')$ —functions (3.2) and α_h^{2D} involves $\delta^2(\mathbf{r}'_{\parallel})\delta^2(\mathbf{r}''_{\parallel})\delta(z')\delta(z'')$ —functions (3.17) facilitating integrations on the right of (3.36), we obtain

$$U(\mathbf{r}_{\parallel}; z; \omega) = U_0(\mathbf{r}_{\parallel}; z; \omega) + I_1(\mathbf{r}_{\parallel}, z; \omega) - I_2(\mathbf{r}_{\parallel}, z; \omega), \quad (3.37)$$

where (3.19) has been employed in evaluating the integrals $I_1(\mathbf{r}_{\parallel}, z; \omega)$ and $I_2(\mathbf{r}_{\parallel}, z; \omega)$ as

$$\begin{aligned} I_1(\mathbf{r}_{\parallel}, z; \omega) = & \frac{\omega^2 d}{c^2} \int d^2 \mathbf{r}''_{\parallel} \left\{ \frac{1}{(2\pi)^2} \int d^2 \mathbf{k}_{\parallel} e^{i\mathbf{k}_{\parallel} \cdot \mathbf{r}_{\parallel}} G_{fs}(\mathbf{k}_{\parallel}; z, 0; \omega) e^{-i\mathbf{k}_{\parallel} \cdot \mathbf{r}''_{\parallel}} \alpha_{fs}^{2D}(\mathbf{k}_{\parallel}; \omega) \right\} \\ & \times U_0(\mathbf{r}''_{\parallel}; 0; \omega) \\ & - \frac{\omega^2 d}{c^2} \left[\frac{\beta G_{fs}(\mathbf{r}_{\parallel}, 0; z, 0; \omega)}{1 + \beta G_{fs}(0, 0; 0, 0; \omega)} \right] \\ & \times \int d^2 \mathbf{r}''_{\parallel} \left\{ \frac{1}{(2\pi)^2} \int d^2 \mathbf{k}_{\parallel} G_{fs}(\mathbf{k}_{\parallel}; 0, 0; \omega) \alpha_{fs}^{2D}(\mathbf{k}_{\parallel}; \omega) e^{-i\mathbf{k}_{\parallel} \cdot \mathbf{r}''_{\parallel}} \right\} \\ & \times U_0(\mathbf{r}''_{\parallel}; 0; \omega) \end{aligned} \quad (3.38)$$

and

$$\begin{aligned} I_2(\mathbf{r}_{\parallel}, z; \omega) &= \beta G(\mathbf{r}_{\parallel}, 0; z, 0; \omega) U_0(0; 0; \omega) \\ &= \beta \left[\frac{G_{fs}(\mathbf{r}_{\parallel}, 0; z, 0; \omega)}{1 + \beta G_{fs}(0, 0; 0, 0; \omega)} \right] U_0(0; 0; \omega), \end{aligned} \quad (3.39)$$

with G_{fs} expressed in lateral position representation. It should be noted that in the absence of a nano-hole, $\beta = \gamma A^2 = 0$ and $I_2(\mathbf{r}_{\parallel}, z; \omega) = 0$. In this case, the field $U(\mathbf{r}_{\parallel}, z; \omega) \mapsto U_{fs}(\mathbf{r}_{\parallel}, z; \omega)$ is that of the full sheet alone, so that

$$\begin{aligned} U_{fs}(\mathbf{r}_{\parallel}; z; \omega) &= U_0(\mathbf{r}_{\parallel}; z; \omega) + I_1(\mathbf{r}_{\parallel}, z; \omega) \\ &= U_0(\mathbf{r}_{\parallel}; z; \omega) \\ &+ \frac{d\omega^2}{c^2} \int d^2 \mathbf{r}'_{\parallel} \int d^2 \mathbf{r}''_{\parallel} \frac{1}{(2\pi)^2} \left(\int d^2 \mathbf{k}_{\parallel} e^{i\mathbf{k}_{\parallel} \cdot (\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel})} G_{fs}(\mathbf{k}_{\parallel}; z, 0; \omega) \right) \\ &\times \left(\frac{1}{(2\pi)^2} \int d^2 \mathbf{k}'_{\parallel} e^{i\mathbf{k}'_{\parallel} \cdot (\mathbf{r}'_{\parallel} - \mathbf{r}''_{\parallel})} \alpha_{fs}^{2D}(\mathbf{k}'_{\parallel}; \omega) \right) U_0(\mathbf{r}''_{\parallel}; 0; \omega). \end{aligned} \quad (3.40)$$

Using (3.9), this yields

$$U_{fs}(\mathbf{r}_{\parallel}; z; \omega) = U_0(\mathbf{r}_{\parallel}; z; \omega) - \frac{\mu}{[1 + \mu]} \frac{1}{(2\pi)^2} \int d^2 \mathbf{r}_{\parallel}'' \int d^2 \mathbf{k}_{\parallel} e^{ik_{\perp}|z| + i\mathbf{k}_{\parallel} \cdot (\mathbf{r}_{\parallel} - \mathbf{r}_{\parallel}'')} U_0(\mathbf{r}_{\parallel}''; 0; \omega). \quad (3.41)$$

Considering an infinite incident plane wave-train of the form

$$\begin{aligned} U_0(\mathbf{r}; t) &= U_0 e^{i[\mathbf{k}_0 \cdot \mathbf{r} - \omega_0 t]} \\ &= U_0 e^{i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp} z - \omega_0 t]}, \end{aligned} \quad (3.42)$$

where U_0 is the amplitude, its frequency representation is given by

$$U_0(\mathbf{r}_{\parallel}; z; \omega) = 2\pi U_0 e^{i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp} z]} \delta(\omega - \omega_0), \quad (3.43)$$

where $k_{0\perp} = \sqrt{q_{\omega_0}^2 - k_{0\parallel}^2}$, $q_{\omega_0} = \frac{\omega_0}{c} \sqrt{\varepsilon_b^{(3D)}}$ is the incident wavenumber, and $k_{0\parallel}$ and $k_{0\perp}$ are the parallel and perpendicular components of the incident wave vector, respectively.

The scalar field in the vicinity of the 2D plasmonic screen in the *absence* of the nano-hole is determined by (3.41) jointly with (3.43); in time representation (executing the $\mathbf{r}_{\parallel}''$ - and $\mathbf{k}_{\parallel}$ -integrations) it is given by ($k_{0\perp} > 0$)

$$U_{fs}(\mathbf{r}_{\parallel}; z; t) = U_0(\mathbf{r}_{\parallel}; z; t) - b U_0 e^{i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp}|z| - \omega_0 t]} \quad (3.44)$$

where $b \equiv \frac{\mu}{1 + \mu}$. The incident and reflected waves occurring in the region $z < 0$ are described jointly as

$$U_{fs}^-(\mathbf{r}_{\parallel}; z; t) = U_0(\mathbf{r}_{\parallel}; z; t) [1 - b e^{-2ik_{0\perp} z}] \quad (3.45)$$

and its square modulus relative to that of the incident wave is given by

$$\left| \frac{U_{fs}^-(\mathbf{r}_{\parallel}; z; t)}{U_0} \right|^2 = 1 + b^2 - 2b \cos[2k_{0\perp} z]. \quad (3.46)$$

The total transmitted wave occurring in the region $z > 0$ is given by

$$U_{fs}^+(\mathbf{r}_{\parallel}; z; t) = U_0(\mathbf{r}_{\parallel}; z; t) - b U_0 e^{i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp} z - \omega_0 t]}, \quad (3.47)$$

and its square modulus (relative to the incident wave) is given by

$$\left| \frac{U_{fs}^+(\mathbf{r}_{\parallel}; z; t)}{U_0} \right|^2 = (1 - b)^2. \quad (3.48)$$

It is readily seen that this result correctly reproduces the complete absence of a transmitted wave in the limit in which the plasmonic layer becomes a perfectly conducting screen. In this case, $\alpha_{fs}^{(2D)}$ tends to negative infinity and b approaches 1. Alternatively, for $z > 0$, (3.47) takes the form

$$U_{fs}^+(\mathbf{r}_{\parallel}; z; t) = U_0(\mathbf{r}_{\parallel}; z; t) - U_0 e^{i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp} z - \omega_0 t]}, \quad (3.49)$$

or

$$U_{fs}^+(\mathbf{r}_{\parallel}; z; t) = 0, \quad (3.50)$$

which also shows that there is no transmission. On the other hand, for $z < 0$, (3.45) becomes

$$U_{fs}^-(\mathbf{r}_{\parallel}; z; t) = U_0(\mathbf{r}_{\parallel}; z; t) - U_0 e^{i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} - k_{0\perp} z - \omega_0 t]} \quad (3.51)$$

which indicates total reflection.

Addressing scalar field propagation in the vicinity of a perforated plasmonic layer, (3.37), (3.38), (3.39) and (3.43) yield a delta function, $\delta^{(2)}(\mathbf{k}_{0\parallel} - \mathbf{k}_{\parallel})$ after $\mathbf{r}_{\parallel}''$ -integration, which facilitates the $\mathbf{k}_{\parallel}$ -integration, with result for $I_1(\mathbf{r}_{\parallel}; z; \omega)$ as

$$\begin{aligned} I_1(\mathbf{r}_{\parallel}; z; \omega) &= \frac{\omega^2 d}{c^2} G_{fs}(\mathbf{k}_{0\parallel}; z, 0; \omega) \alpha_{fs}^{2D}(\mathbf{k}_{0\parallel}; \omega) 2\pi \delta(\omega - \omega_0) U_0 e^{i\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel}} \\ &\quad - \frac{\omega^2 d}{c^2} \left[\frac{\beta G_{fs}(\mathbf{r}_{\parallel}, 0; z, 0; \omega)}{1 + \beta G_{fs}(0, 0; 0, 0; \omega)} \right] \\ &\quad \times G_{fs}(\mathbf{k}_{0\parallel}; 0, 0; \omega) \alpha_{fs}^{2D}(\mathbf{k}_{0\parallel}; \omega) 2\pi \delta(\omega - \omega_0) U_0, \end{aligned} \quad (3.52)$$

and using (3.20), we have

$$\begin{aligned} I_1(\mathbf{r}_{\parallel}; z; \omega) &= \frac{\omega^2 d}{c^2} G_{fs}(\mathbf{k}_{0\parallel}; z, 0; \omega) \alpha_{fs}^{2D}(\mathbf{k}_{0\parallel}; \omega) 2\pi \delta(\omega - \omega_0) U_0 e^{i\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel}} \\ &\quad - \frac{\omega^2 d}{c^2} \beta G(\mathbf{r}_{\parallel}, 0; z, 0; \omega) G_{fs}(\mathbf{k}_{0\parallel}; 0, 0; \omega) \\ &\quad \times \alpha_{fs}^{2D}(\mathbf{k}_{0\parallel}; \omega) 2\pi \delta(\omega - \omega_0) U_0. \end{aligned} \quad (3.53)$$

Bearing in mind (3.37) for $U(\mathbf{r}_{\parallel}; z; \omega)$ and using (3.39) for $I_2(\mathbf{r}_{\parallel}; z; \omega)$ with (3.53) for $I_1(\mathbf{r}_{\parallel}; z; \omega)$, and transforming from frequency to time representation, we obtain

$$U(\mathbf{r}_{\parallel}; z; t) = U_0(\mathbf{r}_{\parallel}; z; t) - b U_0 e^{i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp} |z| - \omega_0 t]} - \beta_0 [1 - b] U_0 G(\mathbf{r}_{\parallel}, 0; z, 0; \omega_0) e^{-i\omega_0 t}, \quad (3.54)$$

which can also be written in terms of U_{fs} as

$$U(\mathbf{r}_{\parallel}; z; t) = U_{fs}(\mathbf{r}_{\parallel}; z; t) - \beta_0 [1 - b] U_0 G(\mathbf{r}_{\parallel}, 0; z, 0; \omega_0) e^{-i\omega_0 t}, \quad (3.55)$$

where

$$\beta_0 = \gamma_0 A^2; \quad \gamma_0 = \frac{\mu}{3\pi R^3} \left[i(q_{\omega_0} R)^3 - [1 - (q_{\omega_0} R)^2]^{3/2} \right] \quad (3.56)$$

and $U_{fs}(\mathbf{r}_{\parallel}; z; t)$ is given by (3.44).

Equation (3.55) describes the total scalar field generated by the perforated 2D plasmonic layer in the presence of the nano-hole. This result can be analyzed in terms of the square modulus of $U(\mathbf{r}_{\parallel}; z; t)$ for $z < 0$ and for $z > 0$, both of which are given by

$$\begin{aligned} \left| \frac{U(\mathbf{r}_{\parallel}; z; t)}{U_0} \right|^2 &= \left| \frac{U_{fs}(\mathbf{r}_{\parallel}; z; t)}{U_0} \right|^2 + \beta_0 \beta_0^* [1 - b]^2 G(\mathbf{r}_{\parallel}, 0; z, 0; \omega_0) G^*(\mathbf{r}_{\parallel}, 0; z, 0; \omega_0) \\ &\quad - [1 - b] \left[\beta_0^* G^*(\mathbf{r}_{\parallel}, 0; z, 0; \omega_0) \left(e^{i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp} z]} - b e^{i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp} |z|]} \right) \right] \\ &\quad - [1 - b] \left[\beta_0 G(\mathbf{r}_{\parallel}, 0; z, 0; \omega_0) \left(e^{-i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp} z]} - b e^{-i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp} |z|]} \right) \right]. \end{aligned} \quad (3.57)$$

This equation, which represents the total field distribution generated by the 2D plasmonic layer in the presence of the nano-hole, is examined numerically below using the Green's function of (3.20) jointly with the results above for G_{fs} (3.13, 3.25).

3.4 Numerical Analysis of Helmholtz Scalar Wave Propagation in the Vicinity of a Perforated Plasmonic Layer with a Nano-Hole

Considering a subwavelength nano-hole, $(q_{\omega_0} R) < 1$ and $\sqrt{1 - (q_{\omega_0} R)^2}$ is real, the associated frequency range for the incident scalar wave is given by

$$\omega_0 < \frac{c}{R\sqrt{\epsilon_b^{(3D)}}} \equiv \tilde{\omega}_0, \quad (3.58)$$

with $\tilde{\omega}_0$ as the *cut-off angular frequency*. Based on the above results, the wave amplitude field distributions for $z < 0$ and $z > 0$ are evaluated using (3.57), exhibiting dependence on the radius R of the nano-hole.

For the transmitted scalar wave in the region $z > 0$, $|z| = z$, the field distribution of the wave is given by

$$\begin{aligned} \left| \frac{U^+(\mathbf{r}_{\parallel}; z; t)}{U_0} \right|^2 &= \left| \frac{U_{fs}^+(\mathbf{r}_{\parallel}; z; t)}{U_0} \right|^2 + \beta_0 \beta_0^* [1 - b]^2 G(\mathbf{r}_{\parallel}, 0; z, 0; \omega_0) G^*(\mathbf{r}_{\parallel}, 0; z, 0; \omega_0) \\ &\quad - [1 - b] \left[\beta_0^* G^*(\mathbf{r}_{\parallel}, 0; z, 0; \omega_0) \left(e^{i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp} z]} - b e^{i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp} z]} \right) \right] \\ &\quad - [1 - b] \left[\beta_0 G(\mathbf{r}_{\parallel}, 0; z, 0; \omega_0) \left(e^{-i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp} z]} - b e^{-i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp} z]} \right) \right]. \end{aligned} \quad (3.59)$$

Considering the scalar wave to be incident upon the perforated 2D plasmonic layer (with a nano-hole) at an angle θ_0 with respect to the z -axis in the $(x - z)$ plane

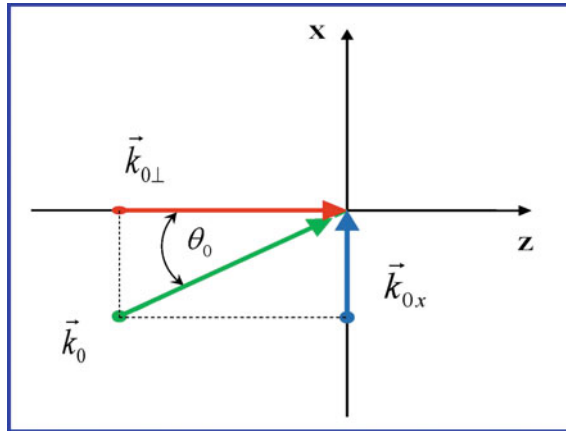


Fig. 3.3 The incident wave vector $\mathbf{k}_0$ in $(x - z)$ plane impinging on the 2D plasmonic layer at angle θ_0 with respect to the normal z -axis, with frequency ω_0

of incidence (Fig. 3.3), the components of the incident wave vector $\mathbf{k}_0$ are defined as follows ($k_0 \equiv q_{\omega_0}$):

$$k_{0\parallel} = k_{0x} = q_{\omega_0} \sin(\theta_0); \quad k_{0y} \equiv 0; \quad k_{0\perp} = q_{\omega_0} \cos(\theta_0). \quad (3.60)$$

Substitution of this into (3.59) yields

$$\begin{aligned} \left| \frac{U^+(x, y, z; t)}{U_0} \right|^2 &= [1 - b]^2 \\ &\times \left[1 + \frac{R^2 b^2 [1 - 3(q_{\omega_0} R)^2 + 3(q_{\omega_0} R)^4]}{[(12 - b)^2 + 3(16 - b)(q_{\omega_0} R)^2 - 3(16 - b)b(q_{\omega_0} R)^4](r_{\parallel}^2 + z^2)} \right] \\ &+ [1 - b]^2 \\ &\times \left[\frac{2bR \sqrt{1 - (q_{\omega_0} R)^2} [(12 - b) - 3(4 - b)(q_{\omega_0} R)^2 - 3b(q_{\omega_0} R)^4] \cos[k_{0x}x + k_{0z}z - q_{\omega_0} \sqrt{r_{\parallel}^2 + z^2}]}{[(12 - b)^2 + 3(16 - b)(q_{\omega_0} R)^2 - 3(16 - b)b(q_{\omega_0} R)^4](r_{\parallel}^2 + z^2)^{1/2}} \right] \\ &+ [1 - b]^2 \\ &\times \left[\frac{2bR(q_{\omega_0} R)[b - 3(4 - b)(q_{\omega_0} R)^2 + 3b(q_{\omega_0} R)^4] \sin[k_{0x}x + k_{0z}z - q_{\omega_0} \sqrt{r_{\parallel}^2 + z^2}]}{[(12 - b)^2 + 3(16 - b)(q_{\omega_0} R)^2 - 3(16 - b)b(q_{\omega_0} R)^4](r_{\parallel}^2 + z^2)^{1/2}} \right], \end{aligned} \quad (3.61)$$

where $r_{\parallel} = \sqrt{x^2 + y^2}$ for $q_{\omega_0} R < 1$.

$\theta_0 = 0^\circ$ (Normal incidence)

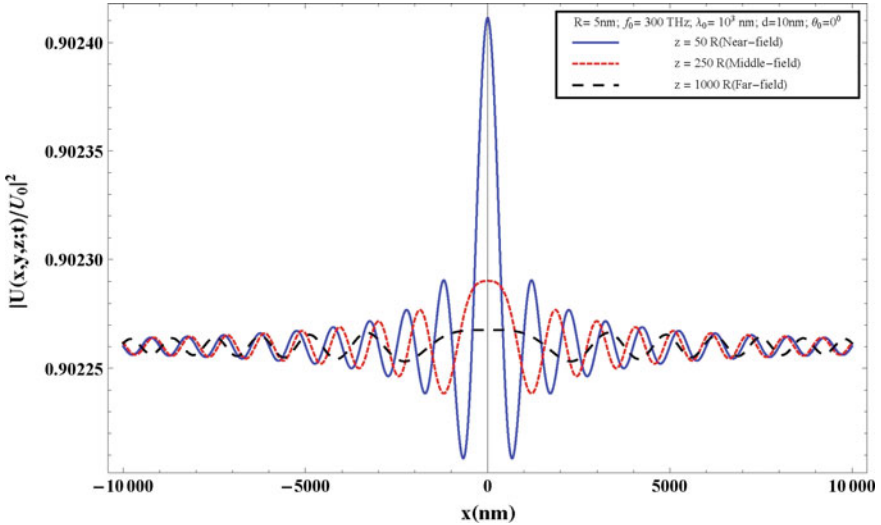


Fig. 3.4 $\theta_0 = 0^\circ$ —Transmitted field square moduli as functions of x produced by a perforated 2D plasmonic layer of GaAs of thickness $d = 10$ nm, effective mass $m^* = 0.067 m_0$ (where m_0 is the free-electron mass, $n_{2D} = 4 \times 10^{15}/\text{cm}^2$ with dielectric constant of the host medium $\epsilon_b^{(3D)} = 1$ in the presence of a nano-hole of radius $R = 5$ nm) for angle of incidence $\theta_0 = 0^\circ$, with incident wavelength $\lambda_0 = 1000$ nm and frequency $f_0 = 300$ THz for $z = 50 R$ [near-field (Blue)], $z = 250 R$ [middle-field (Red)] and $z = 1000 R$ [far-field (Black)] regimes

For the reflected and incident waves in the region $z < 0$, $|z| = -z$, the field distribution is given by

$$\begin{aligned} \left| \frac{U^-(\mathbf{r}_{\parallel}; z; t)}{U_0} \right|^2 &= \left| \frac{U_{fs}^-(\mathbf{r}_{\parallel}; z; t)}{U_0} \right|^2 + \beta_0 \beta_0^* [1 - b]^2 G(\mathbf{r}_{\parallel}, 0; z, 0; \omega_0) G^*(\mathbf{r}_{\parallel}, 0; z, 0; \omega_0) \\ &\quad - [1 - b] \left[\beta_0^* G^*(\mathbf{r}_{\parallel}, 0; z, 0; \omega_0) \left(e^{i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp} z]} - b e^{i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} - k_{0\perp} z]} \right) \right] \\ &\quad - [1 - b] \left[\beta_0 G(\mathbf{r}_{\parallel}, 0; z, 0; \omega_0) \left(e^{-i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} + k_{0\perp} z]} - b e^{-i[\mathbf{k}_{0\parallel} \cdot \mathbf{r}_{\parallel} - k_{0\perp} z]} \right) \right]. \end{aligned} \quad (3.62)$$

The numerical analysis is focused on the transmitted field, with $z > 0$ for $R = 5$ nm. The transmitted field distributions are exhibited as functions of x for $z > 0$ with the parameters chosen as $f_0 = \frac{\omega_0}{2\pi} = 300$ THz and $\lambda_0 = 1000$ nm. These field distributions show the nature of the transmitted scalar radiation fields in the subwavelength regime ($R \ll \lambda$) in terms of near-field, middle-field and far-field zones of diffraction for various values of z at $y = 0$.

Figures 3.4, 3.5, 3.6 and 3.7, exhibit the square moduli of the transmitted fields $|U^+(x, y, z; t)/U_0|^2$ as functions of x at $y = 0$ for $\theta_0 = 0^\circ$, $\theta_0 = 30^\circ$, $\theta_0 = 60^\circ$, and

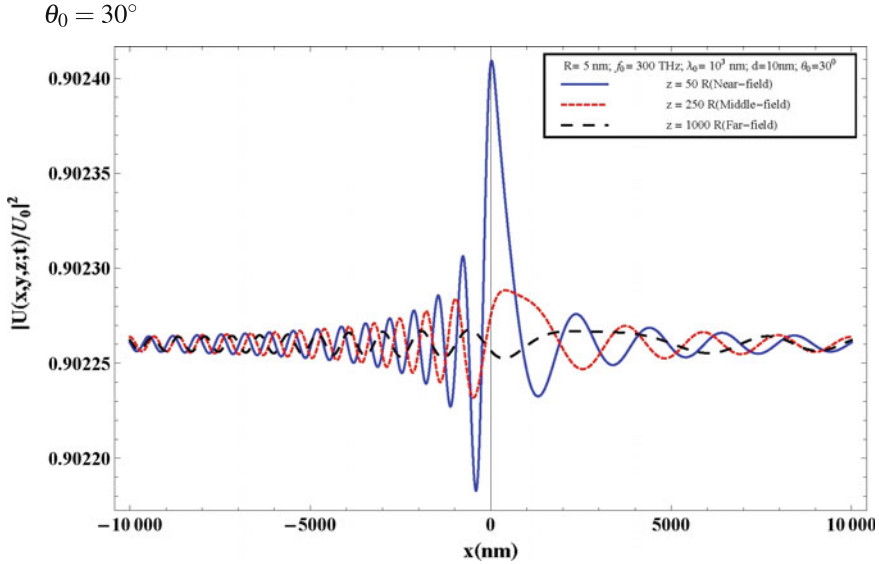


Fig. 3.5 $\theta_0 = 30^\circ$ —Transmitted field square moduli as functions of x produced by a perforated 2D plasmonic layer of GaAs of thickness $d = 10$ nm, effective mass $m^* = 0.067m_0$ (where m_0 is the free-electron mass, $n_{2D} = 4 \times 10^{15}/\text{cm}^2$ with dielectric constant of the host medium $\epsilon_b^{(3D)} = 1$ in the presence of a nano-hole of radius $R = 5$ nm) for angle of incidence $\theta_0 = 30^\circ$, with incident wavelength $\lambda_0 = 1000$ nm and frequency $f_0 = 300$ THz for $z = 50 R$ [near-field (Blue)], $z = 250 R$ [middle-field (Red)] and $z = 1000 R$ [far-field (Black)] regimes

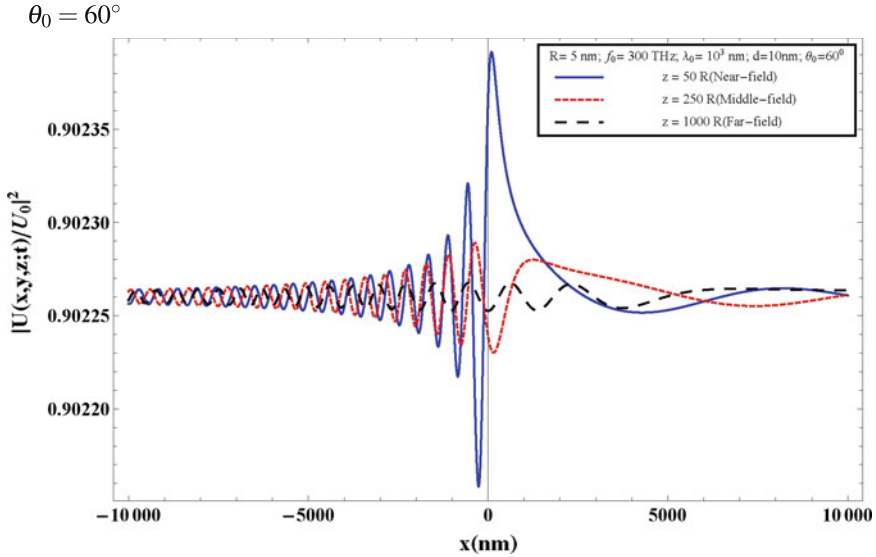


Fig. 3.6 $\theta_0 = 60^\circ$ —Transmitted field square moduli as functions of x produced by a perforated 2D plasmonic layer of GaAs of thickness $d = 10$ nm, effective mass $m^* = 0.067 m_0$ (where m_0 is the free-electron mass, $n_{2D} = 4 \times 10^{15}/\text{cm}^2$ with dielectric constant of the host medium $\epsilon_b^{(3D)} = 1$ in the presence of a nano-hole of radius $R = 5$ nm) for angle of incidence $\theta_0 = 60^\circ$, with incident wavelength $\lambda_0 = 1000$ nm and frequency $f_0 = 300$ THz for $z = 50 R$ [near-field (Blue)], $z = 250 R$ [middle-field (Red)] and $z = 1000 R$ [far-field (Black)] regimes

$\theta_0 = 80^\circ$, respectively, with several values of $z = 50 R$, $z = 250 R$ and $z = 1000 R$ in each figure (taking $R = 5$ nm and incident wave frequency $f_0 = \frac{c_0}{2\pi} = 300$ THz).

In addition to the transmitted field intensities illustrated above as functions of x for fixed $y = 0$, we also exhibit the field amplitude distributions $|U^+(x, y, z; t)/U_0|^2$ of the transmitted field intensities for the entire $x - y$ plane in terms of 3D and density plots, which characterize the spread of the transmitted fields as they propagate in the positive z direction. The evaluations exhibited below in Figs. 3.8, 3.9, 3.10 and 3.11 include angles of incidence $\theta_0 = 0^\circ, 30^\circ, 60^\circ$ and 80° , respectively; and each figure has constituent pairs of 3D and density plots for each of the diffraction zones [near ($z = 50 R$), middle ($z = 250 R$) and far ($z = 1000 R$) zones].

Furthermore, the field distributions $\left| \frac{U^+(x, y, z; t)}{U_0} \right|^2$ of the transmitted waves have been evaluated as functions of distance from the 2D perforated screen (z). It is readily seen that for normal incidence ($\theta_0 = 0^\circ$), the amplitude of the transmitted field distribution decreases as distance from the aperture in the $(x - y)$ plane, $r_{\parallel}$, increases, and it also decreases as a function of increasing z . This distribution flattens uniformly for large z . These features are illustrated in the 3D and density plots of the field as functions of x and z for $y = 0$ in Fig. 3.12 below for $\theta_0 = 0^\circ$, revealing the sharp decay of the field as a function of z while spreading in the x -direction.

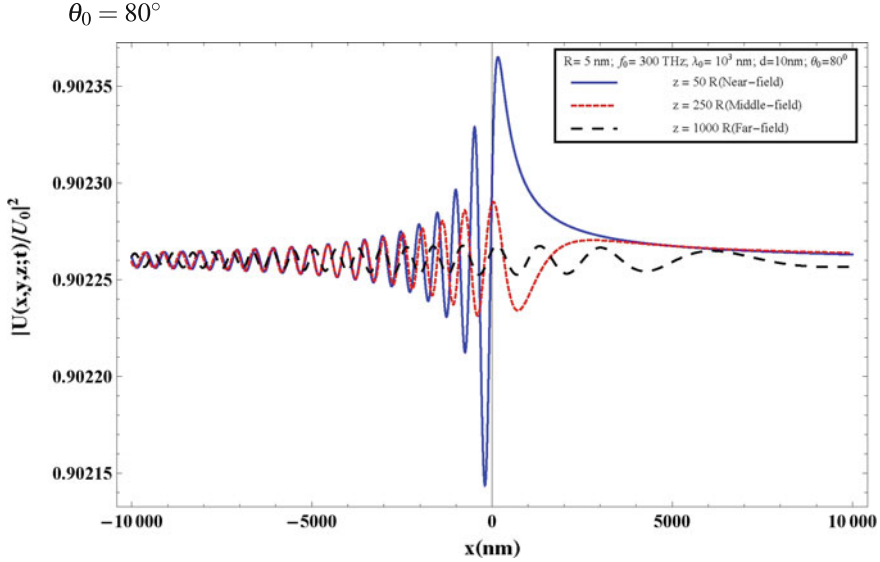


Fig. 3.7 $\theta_0 = 80^\circ$ —Transmitted field square moduli as functions of x produced by a perforated 2D plasmonic layer of GaAs of thickness $d = 10$ nm, effective mass $m^* = 0.067 m_0$ (where m_0 is the free-electron mass, $n_{2D} = 4 \times 10^{15}/\text{cm}^2$ with dielectric constant of the host medium $\epsilon_b^{(3D)} = 1$ in the presence of a nano-hole of radius $R = 5$ nm) for angle of incidence $\theta_0 = 80^\circ$, with incident wavelength $\lambda_0 = 1000$ nm and frequency $f_0 = 300$ THz for $z = 50 R$ [near-field (Blue)], $z = 250 R$ [middle-field (Red)] and $z = 1000 R$ [far-field (Black)] regimes

Moreover, we have also calculated the field distributions $\left| \frac{U^+(x,y,z;t)}{U_0} \right|^2$ of the transmitted wave as functions of the incident frequency $f = f_0$ (THz) at a fixed finite x -value ($x = 200 R$) with $y = 0$ for incident angles $\theta_0 = 0^\circ$, $\theta_0 = 30^\circ$, $\theta_0 = 60^\circ$, and $\theta_0 = 80^\circ$, for the near-field ($z = 50 R$), middle-field ($z = 250 R$) and far-field ($z = 1000 R$) zones of diffraction. These results are shown in Figs. 3.13, 3.14, 3.15 and 3.16:

Furthermore, the 3D and density plot patterns of the transmitted field distributions $\left| \frac{U^+(x,y,z;t)}{U_0} \right|^2$ are exhibited in Figs. 3.17, 3.18, 3.19 and 3.20 as functions of position x and incoming frequency $f = f_0$ for incident angles $\theta_0 = 0^\circ, 30^\circ, 60^\circ, 80^\circ$ with values of $z = 50 R, 250 R$, and $1000 R$ for near, middle and far zones, respectively, ($y \equiv 0$). At normal incidence, symmetry is observed in the intensity distributions in the 3D plots in position space, while the distributions decay as functions of frequency.

In addition to the field distributions of transmitted waves as functions of position and frequency f (THz), we have also determined the field distributions $\left| \frac{U^+(x,y,z;t)}{U_0} \right|^2$ of transmitted waves as functions of incident angle θ_0 for near-field, middle-field and far-field zones of diffraction for $z = 50 R, 250 R$ and $1000 R$, respectively, and $x = 500 R$ with $y = 0$. These results are presented in Fig. 3.21 ($R = 5$ nm and $x = 500 R$).

For $\theta_0 = 0^\circ$ and $R = 5 \text{ nm}$, $\lambda_0 = 1000 \text{ nm}$, $f_0 = 300 \text{ THz}$

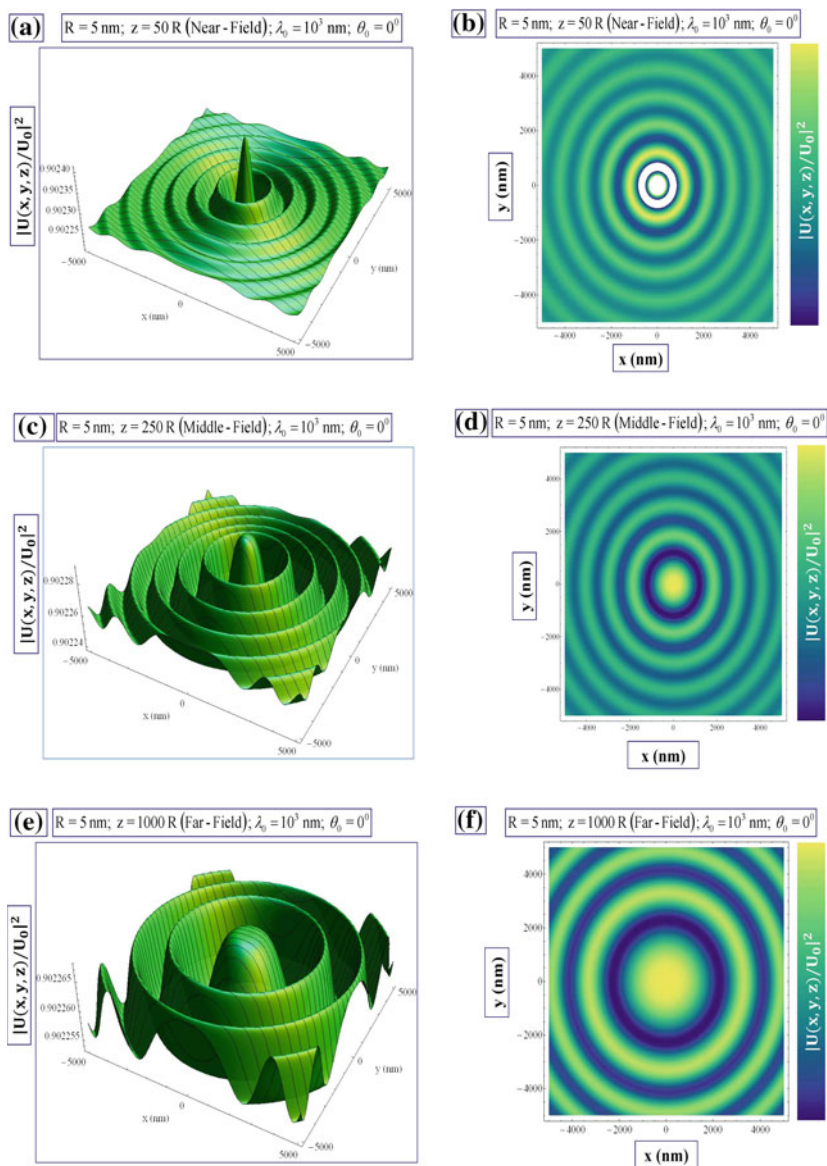


Fig. 3.8 $\theta_0 = 0^\circ$ —Transmitted field distribution of a perforated GaAs layer in terms of 3D (a, c, e) and density (b, d, f) plots

For $\theta_0 = 30^\circ$ and $R = 5 \text{ nm}$, $\lambda_0 = 1000 \text{ nm}$, $f_0 = 300 \text{ THz}$

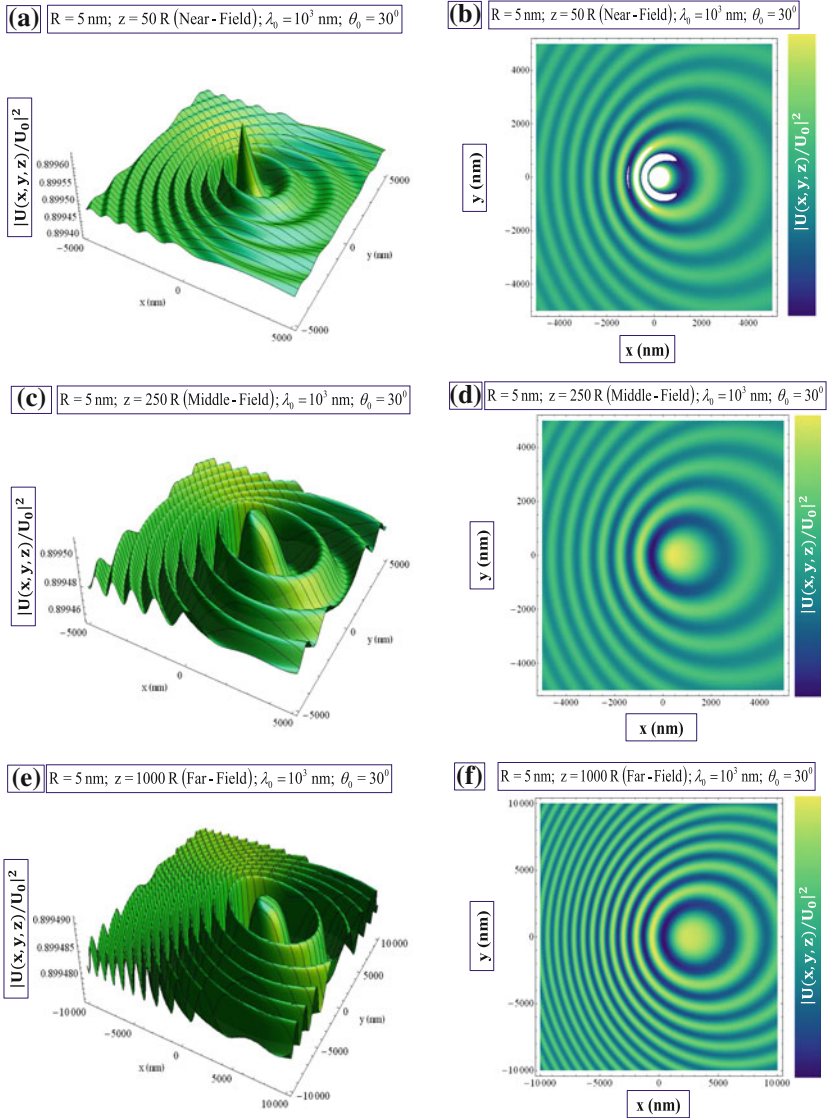


Fig. 3.9 $\theta_0 = 30^\circ$ —Transmitted field distribution of a perforated GaAs layer in terms of 3D (a, c, e) and density (b, d, f) plots

For $\theta_0 = 60^\circ$ and $R = 5 \text{ nm}$, $\lambda_0 = 1000 \text{ nm}$, $f_0 = 300 \text{ THz}$

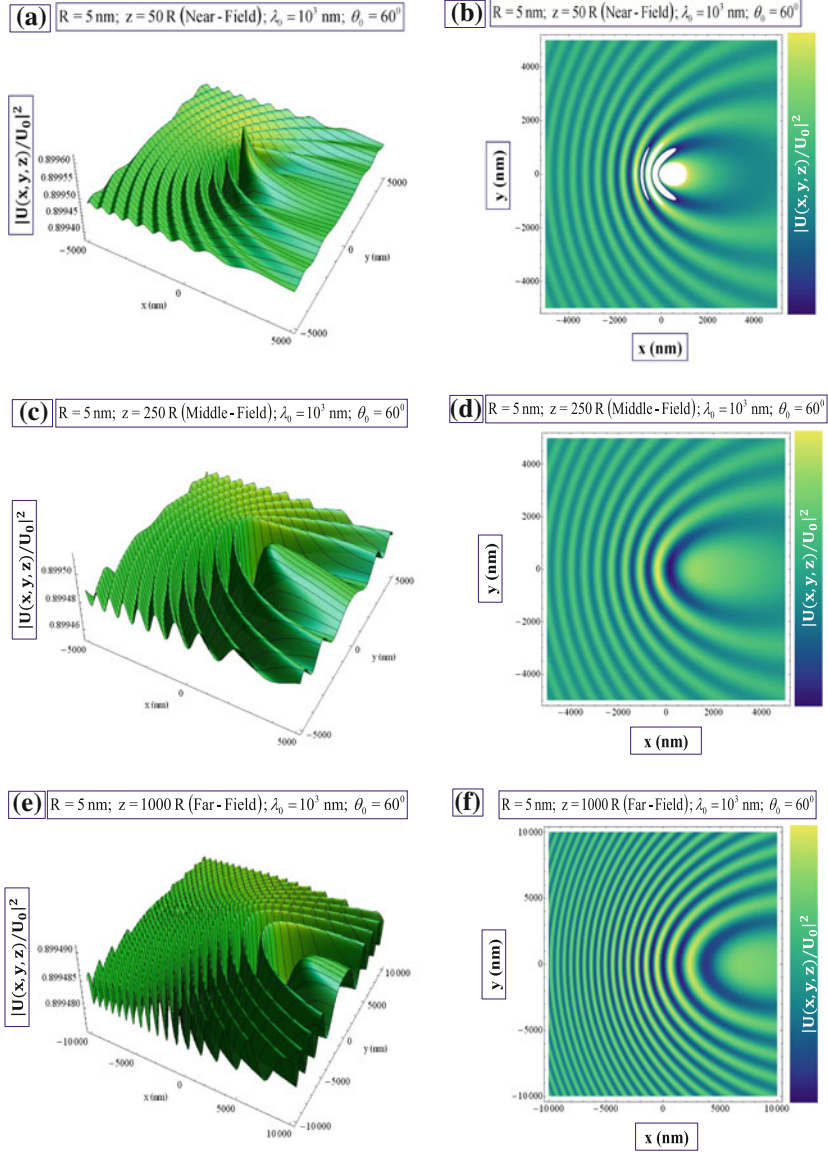


Fig. 3.10 $\theta_0 = 60^\circ$ —Transmitted field distribution of a perforated GaAs layer in terms of 3D (a, c, e) and density (b, d, f) plots

For $\theta_0 = 80^\circ$ and $R = 5 \text{ nm}$, $\lambda_0 = 1000 \text{ nm}$, $f_0 = 300 \text{ THz}$

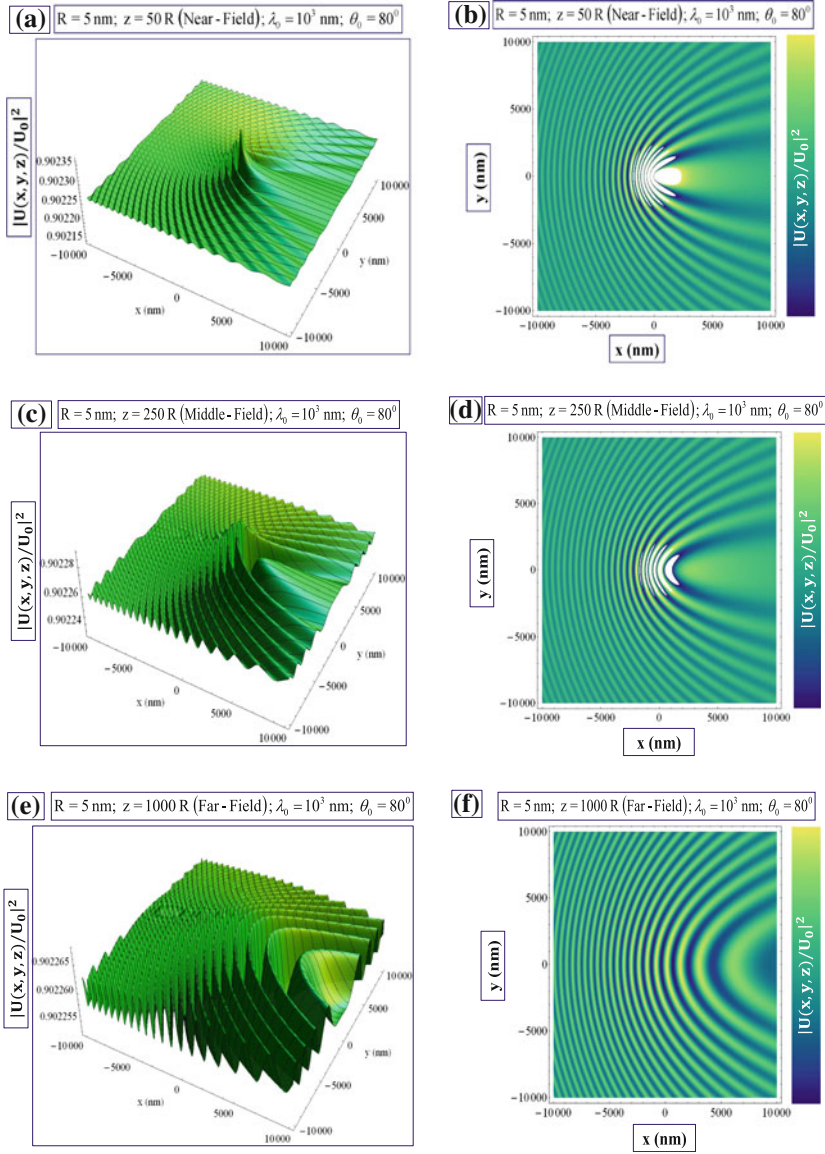


Fig. 3.11 $\theta_0 = 80^\circ$ —Transmitted field distribution of a perforated GaAs layer in terms of 3D (a, c, e) and density (b, d, f) plots

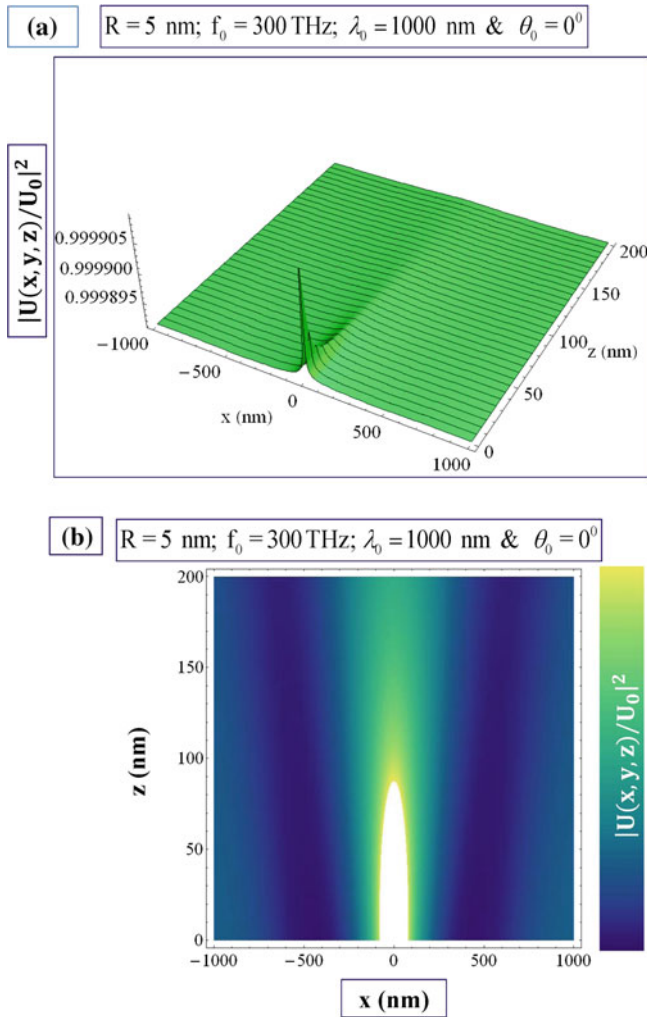


Fig. 3.12 Transmitted field distribution as a function of x and z for $y \equiv 0$ (**a** 3D and **b** density plots) of the 2D plasmonic layer of GaAs of thickness $d = 10 \text{ nm}$ in the presence of a nanohole of $R = 5 \text{ nm}$ at angle $\theta_0 = 0^\circ$, with incident wavelength $\lambda_0 = 1000 \text{ nm}$ and frequency $f = f_0 = 300 \text{ THz}$

For $\theta_0 = 0^\circ$ and $x = 200 R$ ($R = 5 \text{ nm}$)

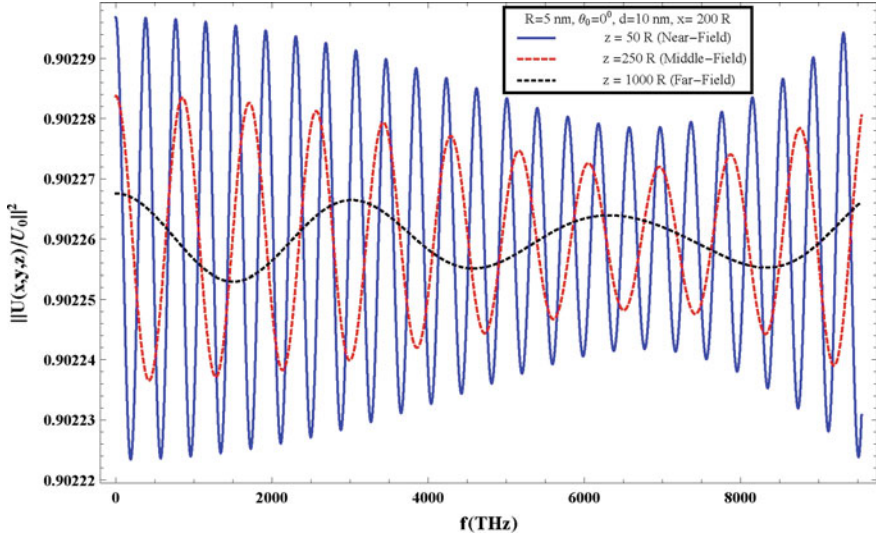


Fig. 3.13 $\theta_0 = 0^\circ$ and $x = 200 R$ ($y \equiv 0$): Transmitted field distributions as functions of frequency f produced by a perforated 2D plasmonic layer of GaAs of thickness $d = 10 \text{ nm}$, effective mass $m^* = 0.067 m_0$ (where m_0 is the free-electron mass, $n_{2D} = 4 \times 10^{15} / \text{cm}^2$ with dielectric constant of the host medium $\epsilon_b^{(3D)} = 1$ in the presence of the nano-hole of $R = 5 \text{ nm}$); for $z = 50 R$ (Blue), $z = 250 R$ (Red) and $z = 1000 R$ (Black)

For $\theta_0 = 30^\circ$ and $x = 200 R$ ($R = 5 \text{ nm}$)

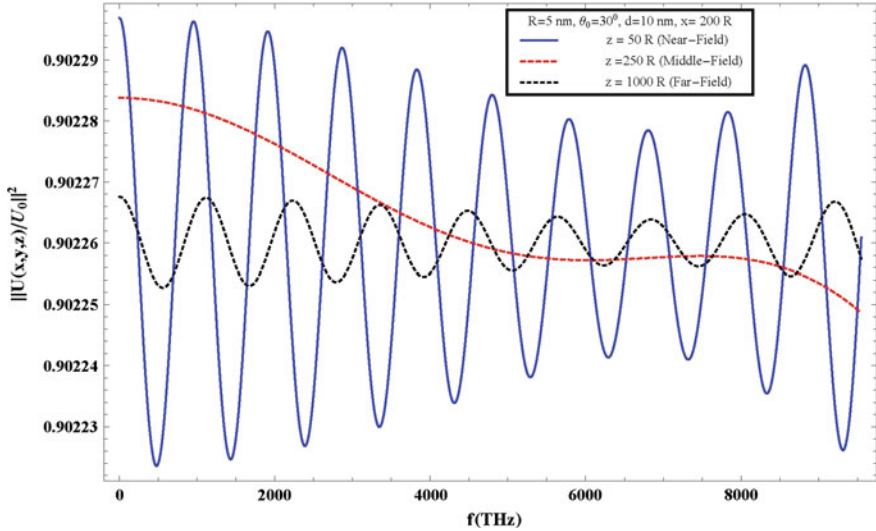


Fig. 3.14 $\theta_0 = 30^\circ$ and $x = 200 R$ ($y \equiv 0$): Transmitted field distributions as functions of frequency f produced by a perforated 2D plasmonic layer of GaAs of thickness $d = 10 \text{ nm}$, effective mass $m^* = 0.067 m_0$ (where m_0 is the free-electron mass, $n_{2D} = 4 \times 10^{15} / \text{cm}^2$ with dielectric constant of the host medium $\epsilon_b^{(3D)} = 1$ in the presence of the nano-hole of $R = 5 \text{ nm}$); for $z = 50 R$ (Blue), $z = 250 R$ (Red) and $z = 1000 R$ (Black)

For $\theta_0 = 60^\circ$ and $x = 200R$ ($R = 5\text{ nm}$)

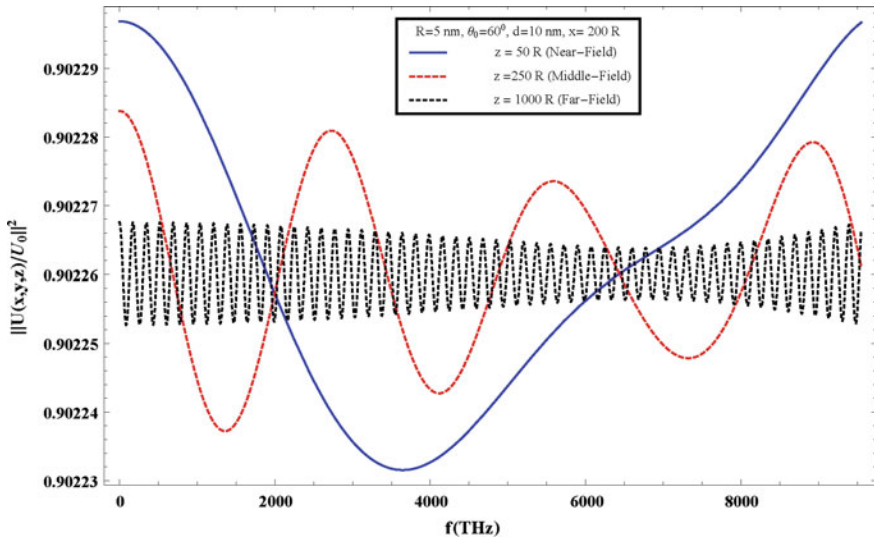


Fig. 3.15 $\theta_0 = 60^\circ$ and $x = 200R$ ($y \equiv 0$): Transmitted field distributions as functions of frequency f produced by a perforated 2D plasmonic layer of GaAs of thickness $d = 10\text{ nm}$, effective mass $m^* = 0.067 m_0$ (where m_0 is the free-electron mass, $n_{2D} = 4 \times 10^{15}/\text{cm}^2$ with dielectric constant of the host medium $\epsilon_b^{(3D)} = 1$ in the presence of the nano-hole of $R = 5\text{ nm}$); for $z = 50R$ (Blue), $z = 250R$ (Red) and $z = 1000R$ (Black)

For $\theta_0 = 80^\circ$ and $x = 200R$ ($R = 5\text{ nm}$)

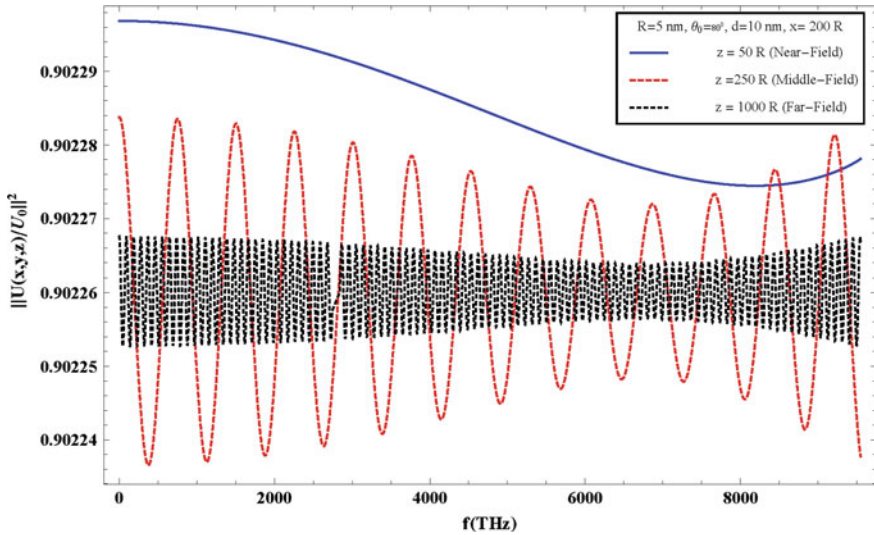


Fig. 3.16 $\theta_0 = 80^\circ$ and $x = 200R$ ($y \equiv 0$): Transmitted field distributions as functions of frequency f produced by a perforated 2D plasmonic layer of GaAs of thickness $d = 10\text{ nm}$, effective mass $m^* = 0.067 m_0$ (where m_0 is the free-electron mass, $n_{2D} = 4 \times 10^{15}/\text{cm}^2$ with dielectric constant of the host medium $\epsilon_b^{(3D)} = 1$ in the presence of the nano-hole of $R = 5\text{ nm}$); for $z = 50R$ (Blue), $z = 250R$ (Red) and $z = 1000R$ (Black)

For $\theta_0 = 0^\circ$ ($R = 5 \text{ nm}$)

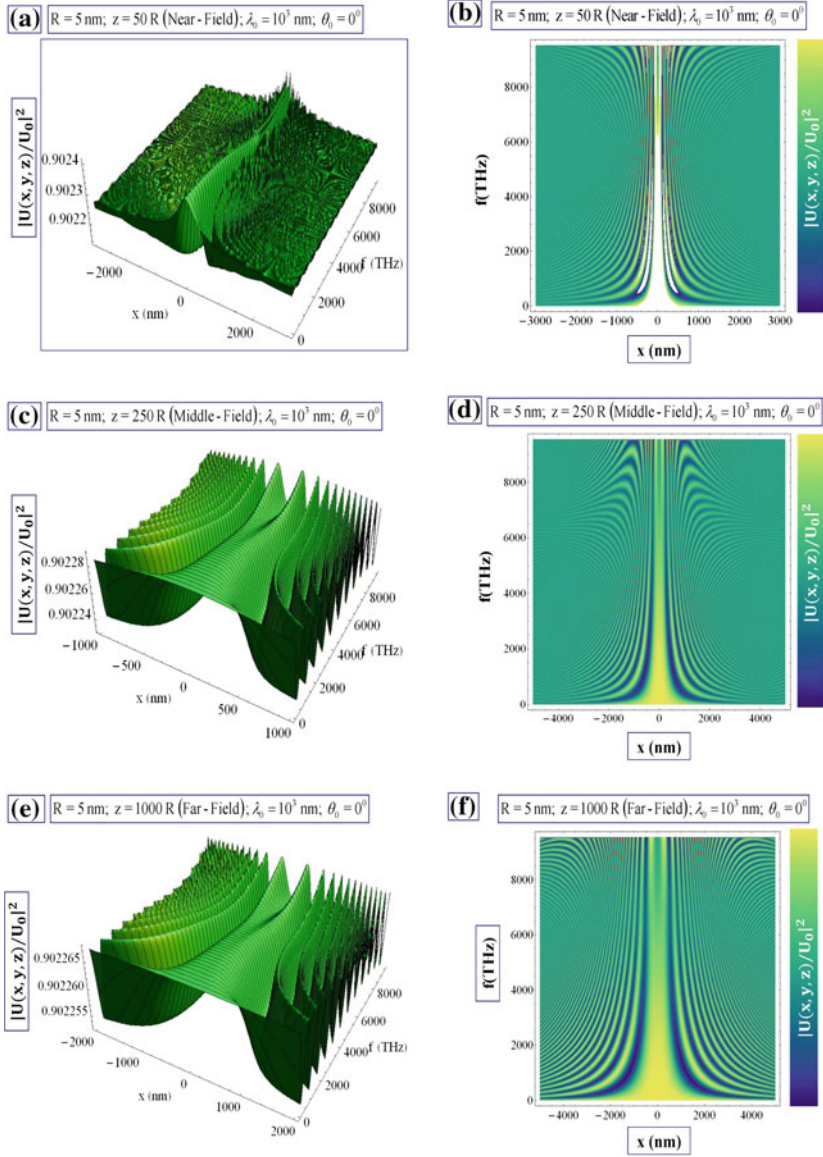


Fig. 3.17 Field distributions as functions of x and $f = f_0$ in terms of 3D (a, c, e) and density (b, d, f) plots

For $\theta_0 = 30^\circ$ ($R = 5 \text{ nm}$)

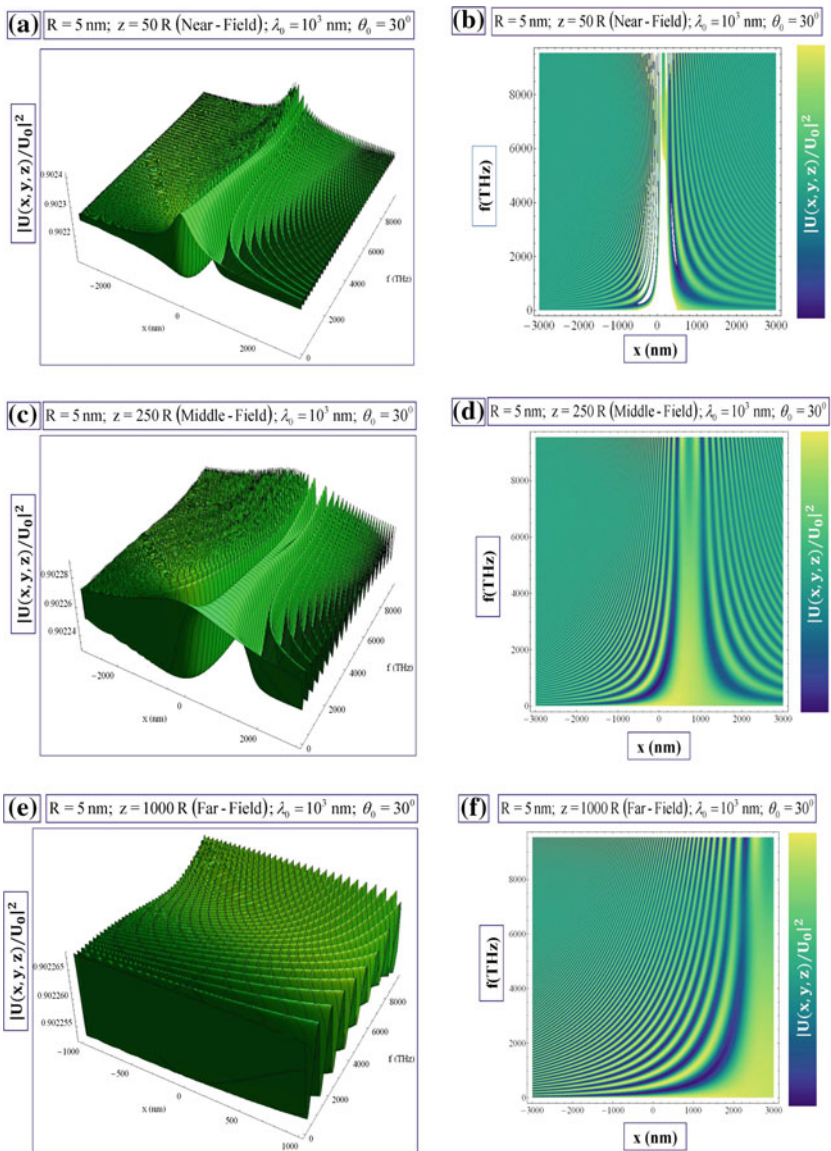


Fig. 3.18 Field distributions as functions of x and $f = f_0$ in terms of 3D (a, c, e) and density (b, d, f) plots

For $\theta_0 = 60^\circ$ ($R = 5 \text{ nm}$)

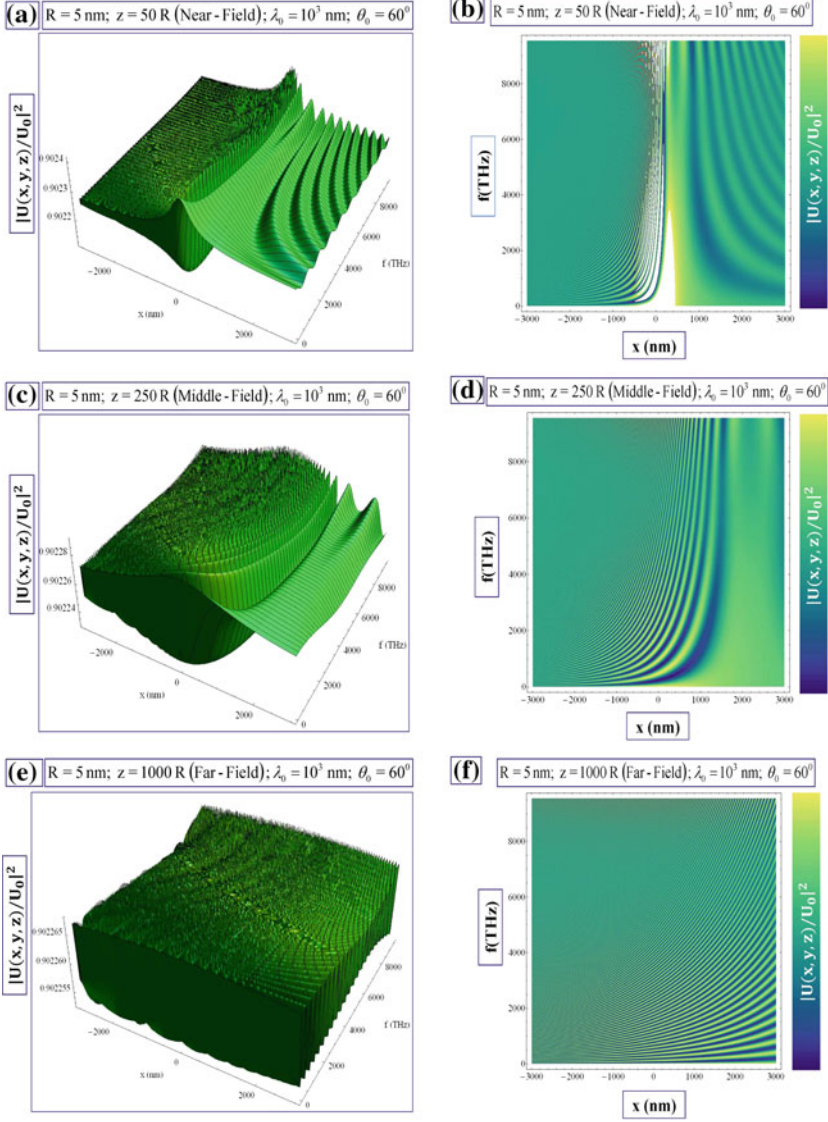


Fig. 3.19 Field distributions as functions of x and $f = f_0$ in terms of 3D (a, c, e) and density (b, d, f) plots

It is also of interest to analyze the field distributions $\left| \frac{U^+(x, y, z; f)}{U_0} \right|^2$ of the transmitted waves as functions of the nano-hole radius R for a fixed value of z with $x = y = 0$. In this case, the acceptable radius range for a subwavelength nano-hole is given by $R < 1/q_{\omega_0}$ as

For $\theta_0 = 80^\circ$ ($R = 5 \text{ nm}$)

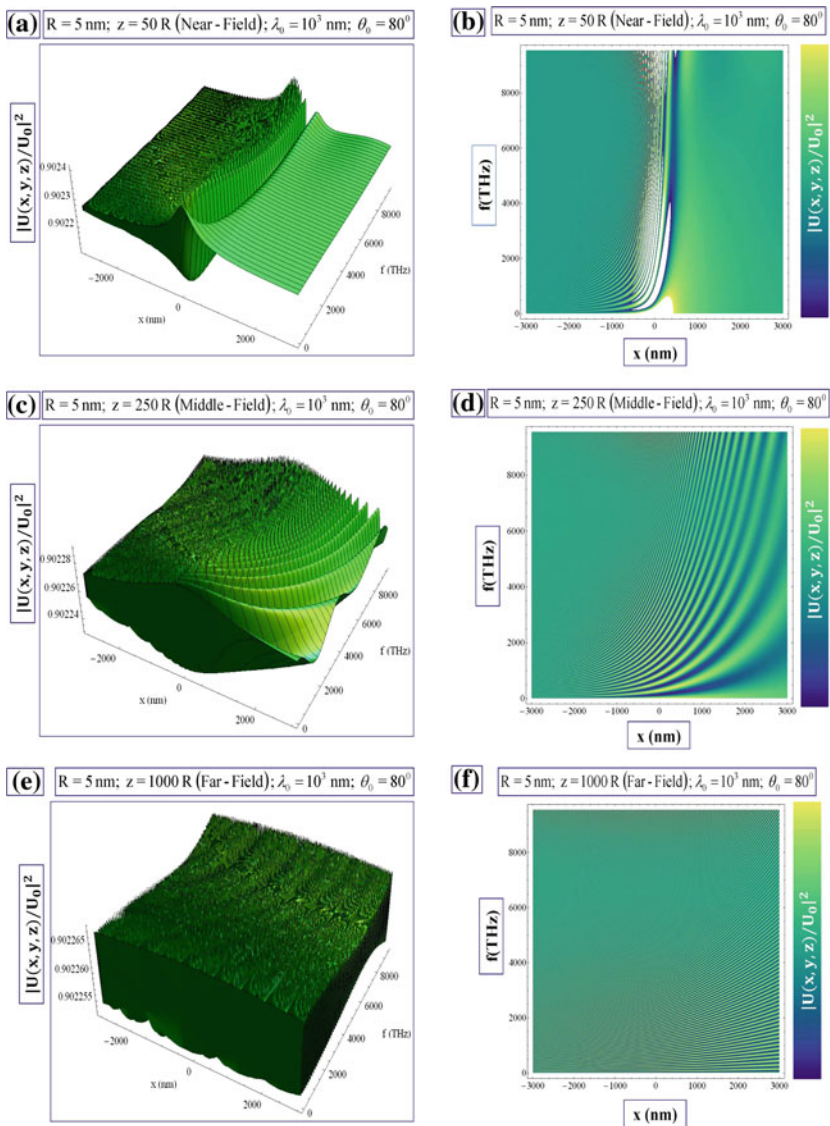


Fig. 3.20 Field distributions as functions of x and $f = f_0$ in terms of 3D (a, c, e) and density (b, d, f) plots

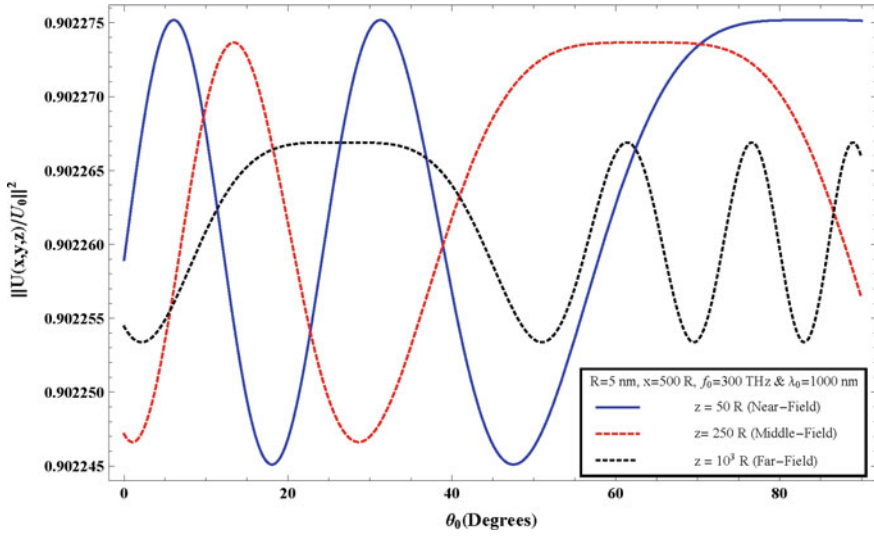


Fig. 3.21 Transmitted field distributions produced by the 2D plasmonic layer of GaAs of thickness $d = 10$ nm, effective mass $m^* = 0.067m_0$ where m_0 is the free-electron mass, $n_{2D} = 4 \times 10^{15}/\text{cm}^2$ with dielectric constant of the host medium $\epsilon_b^{(3D)} = 1$ in the presence of the nano-hole of $R = 5$ nm, with incident wavelength $\lambda_0 = 1000$ nm and linear frequency $f_0 = 300$ THz as functions of θ_0 for $z = 50 R$ (near-field) (Blue), $z = 250 R$ (middle-field) (Red) and $z = 1000 R$ (far-field) (Black)

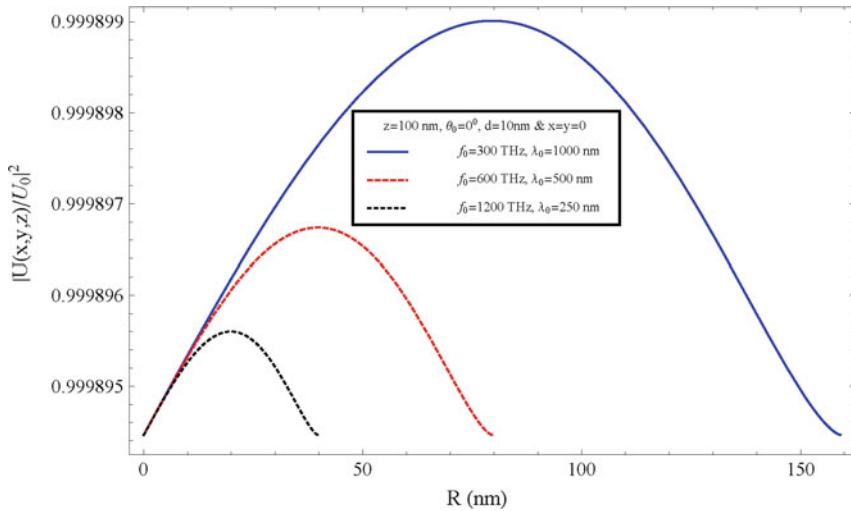


Fig. 3.22 Transmitted field distributions produced by a perforated 2D plasmonic layer (GaAs of thickness $d = 10$ nm) as functions of the hole radius R (nm) with incident wavelength and frequency: $\lambda_0 = 1000$ nm and $f_0 = 300$ THz (Blue), $\lambda_0 = 500$ nm and $f_0 = 600$ THz (Red), and $\lambda_0 = 250$ nm and $f_0 = 1200$ THz (Black), for $z = 100$ nm with $x = y = 0$ and normal incidence

$$R \leq \frac{c}{\omega_0 \sqrt{\varepsilon_b^{(3D)}}} \equiv \tilde{R}. \quad (3.63)$$

The fixed point for which the transmitted field distributions are exhibited as functions of R is $z = 100$ nm, $x = y = 0$. This is illustrated in Fig. 3.22 for normal incidence, $\theta_0 = 0^\circ$, for $f_0 = 300, 600, 1200$ THz.

3.5 Summary

The subject of wave propagation, transmission and diffraction through a subwavelength nanohole aperture embedded in a thin plasmonic layer/screen has been reviewed in this article in terms of a prototypical scalar wave. This analytic approach involving closed form solutions of successive integral equations has facilitated relatively simple numerical computations exhibited in Sect. 3.4, and is not in any way restricted to a metallic screen. Moreover, this formulation, which is based on the use of successive *integral* equations for the associated Helmholtz Green's function involved, automatically incorporates the boundary conditions, which would otherwise need to be addressed explicitly. It also incorporates the role of the thin plasmonic layer, which is smeared by its lateral wavenumber dependence.

In the case of normal incidence, the calculated results shown in the figures of Sect. 3.4 exhibit a positional broadening of the central transmission maximum as z increases from the near to the far zones. This is to be expected as a simple geometric spreading of the radiation emanating from the nano-hole itself, but it is also interesting to note that the transmitted radiation is well supported at large lateral distances from the nano-hole by transmission *directly* through the plasmonic sheet. Moreover, the figures also exhibit interference fringes as a function of lateral distance from the nano-hole in presence of transmission through the sheet as well as through the hole; and as the lateral distance from the hole increases, this interference diminishes and the transmission tends to flatten to that emerging through the sheet alone, in the absence of the distant nano-hole. As the angle of incidence increases, the figures of Sect. 3.4 show that the axis of the central maximum follows it, accompanied by a spatial compression of interference fringe maxima forward of the axis and a spatial thinning of the fringe maxima behind it. Furthermore, the figures also exhibit the dependence of the transmitted/diffracted field intensity on incident frequency $f \equiv f_0$ and nano-hole size (Fig. 3.22).

Finally, the interested reader will find an account of the corresponding electromagnetic vector field propagation/transmission/diffraction through a subwavelength nano-hole aperture in a thin plasmonic screen in research articles by Horing et al. [15], and Miesse et al. [16].

Acknowledgments We gratefully acknowledge support from the NSF-AGEP program.

References

1. H. Cao, A. Nahata, Influence of aperture shape on the transmission properties of a periodic array of subwavelength apertures. *Opt. Express* **12**, 16 (2004)
2. L. Chen, J.T. Robinson, M. Lipson, Role of radiation and surface plasmon polaritons in the optical interactions between a nano-slit and a nano-groove on a metal surface. *Opt. Express* **14**, 26 (2006)
3. C. Genet, T. Ebbesen, Light in tiny holes. *Nature* **445**, 39–46 (2007)
4. R. Gordon, L.K.S. Kumar, A.G. Brolo, Resonant light transmission through a nanohole in a metal film. *IEEE Trans. Nanotechnol.* **5**(3), 291–294 (2006)
5. A. Hajiaboli, M. Kahrizi, V. Truong, Optical behaviour of thick gold and silver films with periodic circular nanohole arrays. *J. Phys. D: Appl. Phys.* **45**, 485105 (2012)
6. S.V. Kukhlevsky, M. Mechler, O. Samek, K. Janssens, Analytical model of the enhanced light transmission through subwavelength metal slits: green's function formalism versus Rayleigh's expansion. *Appl. Phys. B: Lasers Optics* **84**, 19–24 (2006)
7. H. Leong, J. Guo, A surface plasmon resonance spectrometer using a super-period nanohole array. *Opt. Express* **20**(19), 21318–21323 (2012)
8. P.N. Melentiev, A.E. Afansiev, A.A. Kuzin, A.V. Zablotskiy, A.S. Baturin, V.I. Balykin, Single nanohole and photonic crystal: wavelength selective enhanced transmission of light. *Opt. Express* **19**, 23 (2011)
9. O.M. Mendez, M. Cadilhac, R. Petit, Diffraction of a two-dimensional electromagnetic beam wave by a thick slit pierced in a perfectly conducting screen. *J. Opt. Soc. Am.* **73**, 3 (1983)
10. J.M. McMahon, J. Henzie, T.W. Odom, G.C. Schatz, S.K. Gray, Tailoring the sensing capabilities of nanohole arrays in gold films with Rayleigh anomaly-surface plasmon polaritons. *Opt. Express* **15**(26), 18119–18129 (2007)
11. V.A.G. Rivera, in *Plasmonics—Principles and Applications*, chapter 7, ed. by Kim, Kiyoungh (InTech 2012), p.157
12. Y. Wang, Y. Qin, Z. Zhang, *IEEE Photon. J.* **6**(4) (2014)
13. N.J.M. Horing, T.Y. Bagaeva, V.V. Popov, Excitation of radiative polaritons in a two-dimensional excitonic layer by a light pulse. *J. Opt. Soc. Am. B* **24**, 2428–2435 (2007)
14. N.J.M. Horing, Radiative plasmonic-polariton dispersion relation for a thin metallic foil with interband damping transitions. *IEEE Sens. J.* **8**, 771–774 (2008)
15. N.J.M. Horing, D. Miessein, G. Gumbs, Electromagnetic wave transmission through a subwavelength nano-hole in a two-dimensional plasmonic layer. *J. Opt. Soc. Am. A* **32**(6), 1184–1198 (2015)
16. D. Miessein, N.J.M. Horing, H. Lenzing, Incident-angle dependence of electromagnetic wave transmission through a nano-hole in a thin plasmonic semiconductor layer, in preparation for publication

Chapter 4

The Challenge to Develop Metrology at the Nanoscale

R. Ince

Abstract Since nanotechnology goods are manufactured and utilised by the community, legal metrology, human safety, and the environment demand traceable measurement techniques. This is the business of the international network of measurements called metrology. After 2017 the realisation of the international metre will be through the lattice parameter of silicon or another suitable crystalline material. Many NMI's have developed traceable instrumentation systems primarily for AFM, but only partly for SEM and optical instrumentation. None of the existing techniques is able to meet the present requirements for reliable metrology of nanomaterials. However suitable reference materials are being developed alongside standardised sample preparation methods. Present second generation nanostructures are complex requiring multiparameter and ensemble measurements that AFM or SEM cannot offer. Ensemble techniques reveal the sub-nanometre detail required in the healthcare and electronics industry. The next generation of manufacturing in these industries will be three-dimensional complex sub-nm architectures, and nanometrology is currently being driven there.

4.1 Metrology

Metrology is the science of measurement; its mission and implementation is international and ensures that each of the seven international SI units [1] is reliable at all the levels shown in Fig. 4.1. This is achieved via traceability to the representation of each unit held in national laboratories through a chain of unbroken documented calibrations. Representation of each unit is the experiment that realises the unit in national standards laboratories. The national representations must therefore be traceable to the international values (some still held at the international

R. Ince (✉)

Department of Physics, Yeditepe University,
Kayisdagi Caddesi, Kadikoy, 34755 Istanbul, Turkey
e-mail: rince@yeditepe.edu.tr

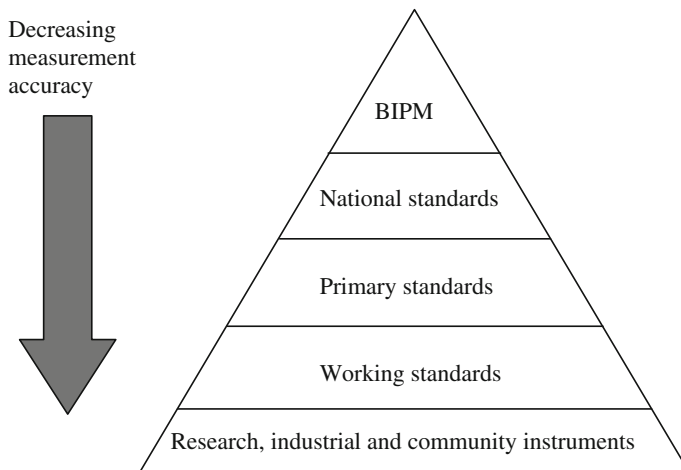


Fig. 4.1 Traceability pyramid showing laboratory hierarchy. Absolute accuracy refers to how close a measurement is to a traceable standard held in the national standards laboratory

bureau of weights and measures, BIPM in Paris) by international comparisons which is described by an uncertainty (a value expressing how near to the real value the measurement is). This creates a highly correlated set of international base units. Scientists themselves are often unaware that such a vast network of perpetual calibrations exists at so many levels (traceability pyramid) and underscores all global measurements for science, trade, and research. Indeed the SI is just as important a part of the national infrastructure as roads, buildings and bridges; the traceability pyramids in Fig. 4.1 illustrate the quintessence of the organization.

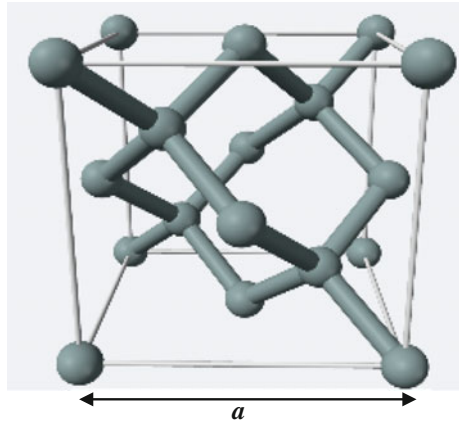
Thus, metrological validation is achieved through comparisons and peer review resulting in a number that expresses the uncertainty called the combined uncertainty, U_c together with the level of confidence (usually 95 %). The value U_c includes in it all the uncertainties that contribute to the measurements and all those which may affect the measurement as the root sum of the squares of each: Pythagorus' theory taken to the n th dimension.

$$U_c = \sqrt{u_1^2 + u_2^2 + \cdots u_n^2} \quad (4.1)$$

At the beginning of the SI system in 1869 the experiments involved transfer artefacts; macro sized objects such as a platinum-iridium cylinder (1 kg) for mass or a bar (1 m) for length. These have very well characterised values. These artefacts are calibrated against others via metrological instruments.

However, even in 1870 James Clerk Maxwell pointed out that artefacts were subject to local conditions, thus presented a barrier to imperishable standards. He predicted their replacement with the parameters of microscopic, universal standards. Over the twentieth century this artefact barrier has been removed and most

Fig. 4.2 The Silicon unit cell, a cubic packing arrangement of 8 atoms. Its volume is determined from the lattice parameter a , the length of one its sides



standards have long been replaced by fundamental constants of nature. The last artefact the unit of mass is due to be replaced in 2017 with a definition from Planck's constant, h [2].

The method by which this has been achieved in the international Avogadro project (2003) [3] is testament to the singular importance of optical and spectroscopic methods in determining dimensions of nanostructures. The lattice constant a , volume, and density ρ of a 1 kg single-crystal sphere, highly enriched with the ^{28}Si were all determined with unprecedented accuracy via optical and spectroscopic methods. The structure of the silicon unit cell is shown in Fig. 4.2.

This enabled determination of the Avogadro number N_A by the equation

$$N_A = nM/\rho a^3 \quad (4.2)$$

where n is the number of atoms per unit cell and M is molar mass.

The Avogadro number is directly linked to Planck's constant via three other SI microscopic representations as:

$$h = \frac{c\alpha^2 A_r M}{2R_\infty N_A} \quad (4.3)$$

where α is the hyperfine splitting constant, R_∞ is the Rydberg gas constant, A_r is relative atomic mass.

The SI definition of length is based on the speed of light, c , in a vacuum. This is realised by measurement of stabilised laser wavelengths using highly sophisticated interferometry of fixed length metal block artefacts (gauge blocks) to accuracies of about 3×10^{-8} under highly controlled conditions; here the wavelength itself must be sub-divided to the nanometre and sub-nm scales required. However at primary level it is achieved by comparator calibration against calibration artefacts called gauge blocks (0.5–100 mm) to an accuracy of 2×10^{-6} under highly controlled conditions [4]. These macroscale artefacts can be picked up and placed in position

for comparison; imagine the challenge of extending the traceability to nanoscale artefacts that cannot even be seen by the naked eye. What does a straight line or a flat surface mean on the nanoscale?

4.2 Nanotechnology Generations, Definitions, and Visions

Technology has innovated away from the 1st generation of passive nanostructures which performed single tasks to the 2nd generation of nanotechnology ‘Active nanostructures’ which multitask e.g. targeted delivery of drugs, the third generation will be nanosystems with thousands of interacting components and several years into the future further innovation to integrated nanosystems will form the 4th generation of systems within systems e.g. molecular manufacturing of genes inside the DNA of targeted cells, and nanosurgery for healing wounds on the cellular level. The aim is to achieve nanoscale control of matter by processes that construct precise three-dimensional structures through molecular manufacturing. This is challenging because quantum physics becomes more important as structures shrink to nanoscale. However, although we have an incomplete knowledge of quantum physics, nature has not; it has built these machines for billions of years as self-assembled viruses and motors, springs, tracks and actuators within cells [5].

In 2011 the European the Commission adopted the definition of nanotechnology as the study of phenomena and fine-tuning of materials at atomic, molecular and macromolecular scales, where properties differ significantly to those at a larger scale. The US National Nanotechnology Initiative defines nanotechnology as the understanding and control of matter engineering of nanosystems and manufacturing them. The measurement of length or size, force, mass, electrical and other properties is included in nanometrology.

Nanotechnology has two primary research thrusts

- Nanoscale electronics (nanowires, connections, memory systems, displays, nanosensors, NEMS, and the protein computer)
- Medicine (drug delivery via functionalization, improved imaging, materials to repair blood vessels and cartilage, nanorobots [6])

Nanotechnology also has two secondary research thrusts

- Nanoenergy (Longer range electric cars with super batteries, strong lightweight materials via nanocomposites, cheaper energy)
- Environment (purifying sea water, cleaning water supplies, removing contaminants)

Nanotechnology also has three *ivory tower* projects under consideration:

- The Space elevator [7–10],
- The protein-based computer [11],

- Molecular fabricators [12–14],
- Nanomachines for cellular repair [12, 15–18].

4.3 Nanotechnology Research Drives

From ancient times, metrology has always been driven by trade and industry; today half of all R&D spending is in healthcare and semiconductors, nanometrology is driven by these industries. Healthcare depends on molecular metrology (the bottom-up approach), whilst the semiconductor industry depends on engineering metrology (top-down approach).

Although research in molecular nanotechnology has only been boosted in the past 10 years, a research rollercoaster began in 1948 in the semiconductor industry, within a few years the world's first transistor computer was built at the University of Manchester in 1953 and commercialized within 3 years [19]. Burgeoning demand for integration into more and more products, and for ever denser chips (currently 995 million transistors per chip at Intel) with better performance and speeds, innovated dimensional metrology.

Measurements of nanosystems are similar to those of the macroscopic world: distance, width, height, form, texture/roughness, volume and physical and chemical characteristics. However, unlike macroscale metrology, the dimensions of nanostructures play a major role in almost all their physical and chemical properties (functionality). For example, when crystal size is smaller than the mean free path of an electron, the conductivity of the crystal is modified; the colour of nanogold goes from the near infra-red through the full visible spectrum to blue, dependent on both size 2–100 nm and shape; seemingly inert substances become catalysts at nanodimensions, stresses in nanosystems become discretized. The latter (as mentioned in the previous section) directly affect the semiconductor and healthcare industries. In anticipation of this profound link, the most common way to classify nanostructures is via their dimensions. Thus, dimensional metrology is the natural basis for all nanometrology; but the link between nanostructure dimension (nanodimensional metrology) and functionality must be strengthened.

4.3.1 *The Semiconductor Industry*

Metrology reduces the cost of manufacturing by bringing in more robust processes, preventing scrap, and helping in ramping and maintaining yield [20].

Metrology tools are required to measure and monitor film thickness, refractive index, resistivity, and stress; metrology systems for defect inspection are essential. For high performance the manufacturing parameters must be controlled: IC line widths and edge profiles must fall within tight tolerances. The minimum feature size

Table 4.1 Metrology systems utilised for measurements in the semiconductor industry

Process step	Measurement	Metrology system
Silicon manufacturing/epitaxy	Resistivity (dependent on thickness), thickness	Four-point probe, FTIR
Wafer inspection	Defects	Defect inspection system
Dielectric deposition	Thickness, refractive index, stress, dielectric constant	Ellipsometer, stress gauge
Etching	Pattern defects, uniformity, etch profile and selectivity	AFM, SEM, reflectometer
Lithography	Critical dimension, pattern defects	SEM
Yield monitoring	Correlation of metrology and inspection results to yield	Data management system

in microchips today is 32 nm wide, yet research aims to decrease this to 20 nm using extreme UV nanolithography. Film thickness must be accurate and uniform and the devices must be free of defects that affect yield: defects as small as a third of the linewidth cause electrical failure. Another important factor is in monitoring of Nano dimensions, since semiconductor wafers tend to change thickness over time due to environmental factors.

As film thickness decreases, it approaches that of the interlayer itself, making it extremely difficult to manufacture or measure with assured accuracy [21]. This is due to limits on instrumentation, data analysis procedures, and the fact that for process control, measurements need to be made to a level below atomic dimensions (Table 4.1).

4.3.2 The Healthcare Industry

Molecular nanometrology aims to advance the little explored and difficult field of measuring nanometre distances/structures in situ, which lies at the heart of molecular biology and medicine. This will lead to improved targeting of drugs to diseased organs and cells reduce side effects and improve efficacy of drugs, improved medical imaging technologies (due to utilisation of magnetic nanoparticles iron oxide) will diagnose diseases earlier on, enabling precise and effective intervention resulting in lower costs.

Important biomedical application areas involve trapping viruses, [22] bacteria, living cells, repairing bones, implant of soft tissues, functionalised nanoparticles for in vitro diagnostics. Here greater control on the nanoscale is required.

Many new drugs have poor solubility or poor permeability, however, through nanotechnology scientists have realised that size distributions less than 100 nm increase surface area, solubility and bioavailability. Thus, ‘active pharmaceutical ingredients’ is a hot topic [23, 24].

Table 4.2 Metrology systems utilised for measurements in the healthcare industry

Process	Measurement	Metrology system
Investigate nanostructures internally	Dimension	Transmission electron microscopy (TEM)
Trap: Dielectric spheres, viruses, bacteria, living cells, organelles, small metal particles, DNA strands. Cell sorting tracking bacteria	Force, dimension, motion tracking	
Imaging	Difficult in biology due to low conductivity	AFM
Imaging	Surface topography, correlation with various physical properties	Optical Tweezers creates an 'optical trap' to hold particles < 1 nm at its centre
Process	Measurement	Scanning tunnelling microscope (STM)

Instruments utilised in the healthcare sciences for nanometrology are Electron Beam Techniques, Scanning Probe Techniques, Optical Tweezers Table 4.2.

4.4 Nanotechnology Artefacts

First we will discuss nanosystems as possible nanometrology artefacts for future nanometrology based on nanosystems presently available due to research, and then we will discuss present nanometrology instruments available for their calibration.

There are 13 main elements and compounds in the literature from which nanostructures are produced. The references in the Table 4.3 show their major area of application in the past 3 years only.

Nanoparticles of the materials in Table 4.3 are mainly produced when their solid particles are vaporised whilst moving on a carrier gas and then allowed to cool in a chamber. They may be vaporised by plasma, laser ablation, or high voltage radio frequency pulses across electrodes.

Presently, there are five overarching nanostructures broadly implemented in the above research thrusts in order: films, composites, nanowires, crystals, rods, fullerenes, tubes, and quantum dots. There is also a menagerie of flourishing secondary nanostructures that demonstrate potential for innovation of any of the major research thrusts. The references attached in the Table 4.4 are noteworthy as the most cited research papers on nanostructures (not including review papers) in each of the past 3 years (Table 4.5):

Table 4.3 Most widely utilised nano materials

Material	Application	References	Times cited
<i>Carbon</i>			
Nanoparticles	EST	[60]	111
Buckyballs [61]	Nanobot components	[62]	25
Nanotubes [63]	E	[64]	111
Graphene [65]	E	[66]	204
Diamondoid [6, 67, 68]	Mechanosynthesis	[69]	11
Platinum	Enhanced catalysis	[70]	182
Gold	Chemo-sensor	[71]	120
Silica	TDD	[72]	102
Boron nitride	electronic components	[73]	89
Boron	TDD	[74]	15
Silicon	EST, nanoantennae/metamaterials	[75, 76]	86, 85
Silver	Light amplification/photonics	[77]	68
Iron oxide	TDD	[78]	63
Iron	Enhanced catalysis	[79, 80]	28, 28
Titanium dioxide	Enhanced electric biosensor	[81]	59
Palladium	Enhanced catalysis	[82]	40
Neodymium	Sub tissue optical probes	[83]	37

ECT/EST energy conversion/storage technology, *E* electronic devices and components, *TDD* targeted drug delivery

4.5 Nanomaterials Measurement Challenges

Physical dimensions of nanosystems must be measured with traceability and novel chemical and physical phenomena accounted for in order to engineer and manufacture nanosystems for precise applications. The definition of a nanomaterial includes measurements of the size and size distributions of Nano-objects embedded in another material [25]. During fabrication these systems require measurement down to accuracies of 0.1 nm. However, robust, easy to use measurement methods for their characterisation are not present. This means a revolutionary ‘holistic’ approach is required for nanoscale metrology. This must involve morphing some of the techniques below or development of Innovative techniques such as X-ray interference, X-ray diffraction, small angle X-ray spectroscopy (SAXS) and non-linear optical microscopy and spectroscopy (e.g. Raman).

Table 4.4 Overarching nanostructures

Overarching nanostructures		Topic	Times cited	References
Films	2013	SC	901	[84]
	2014	Chemo sensor	8	[85]
	2015	Dental adhesives	1	[86]
Composites	2013	SC	373	[87]
	2014	EST	80	[88]
	2015	Bio sensor, TDD, tissue engineering	6, 5, 5	[89, 90]
Wires	2013	SC	244	[91]
	2014	EST	60	[92]
	2015	PC	5	[93]
Crystals	2013	EST	135	[94]
	2014	PC	37	[95]
	2015	SC	9	[96]
Rods	2013	SC	134	[97]
	2014	PC	36	[98]
	2015	EST/ECT, artificial retinas	3,1	[99, 100]
Fullerenes	2013	SC	130	[101]
	2014	SC	30	[102]
	2015	SC	5	[103]
Tubes	2013	Carbon computer	111	[64]
	2014	EST	60	[92]
	2015	Bio sensor	6	[104]
Quantum dots	2013	Optical imaging agents	109	[105]
	2014	SC	58	[106]
	2015	PC	4	[107]

SC solar cells, PC Photocatalysis for transportation fuels

4.6 Nanomaterials Instruments

Spectroscopic ellipsometers, scanning probe microscopes and AFM and SEM are widely used in nanotechnology, however, legal metrology, human safety and the environment demand traceable measurement techniques—nanometrology. This requires experience in the determination of uncertainties and metrological validation through comparisons and peer review. The metrological problems associated with these devices are discussed below.

Table 4.5 Secondary nanostructures

Secondary nanostructures		Topic	Times cited	References
Fibers	2013	Functional materials/E	901	[108]
	2014	EST	8	[109]
	2015	PC/ECT	1	[110]
Nanopores	2013	EST	373	[111]
	2014	Molecular sieve	80	[112]
	2015	E	6,5,5	[113]
Cages	2013	EST	244	[114]
	2014	TDD	60	[115]
	2015	Catalytic enhancement	5	[116]
Pillars	2013	EST	135	[117]
	2014	Enhanced ferromagnetism	37	[118]
	2015	Reversible adhesives, chemo sensor	9	[119, 120]
Needles	2013	EST	134	[121]
	2014	ECT	36	[122]
	2015	EST	3,1	[123]
Meshes	2013	Catalytic enhancement	130	[124]
	2014	Foldable photoelectronics/ muscle-like transducers	30	[125]
	2015	Energy conservation	5	[126]
Foams	2013	Efficient reaction medium	111	[127]
	2014	Nanoscaffolding/architecture, efficient reaction medium	60	[128, 129]
	2015	Super insulation, radiation protection		[130, 131]

4.6.1 Spectroscopic Ellipsometry (SE)

This is an indirect, non-imaging method for dimensional measurement of thin film thickness but also extricates refractive index. They are a non-contact/non-destructive method for solid/air or liquid/air interfaces and are used at every stage of semiconductor processing. Fresnel's equations are utilised to calculate theoretical spectra to be compared with the measured spectra. Regression analysis is then performed by varying the parameters of interest until a best fit is obtained [26]. However, some materials have a refractive index which varies as a function of position and more elaborate models are required to analyse them; this directly affects accuracy (Table 4.6).

The preferred approach for establishing and maintaining accuracy of measurement tool performance is through use of transfer artefacts materials [27–29].

Table 4.6 Typical commercial SE

Spectral range/nm	Resolution/nm	Measurement time/s	Accuracy (NMI)/nm	Repeatability (NMI)/nm	Possible traceability
440–1000	2	Typical 5	$d = 100 \text{ nm} \pm 0.04$	$d = 15 \text{ nm} \pm 0.02$	Wafer transfer artefacts from NMI ^a

^aOnly for very well-characterised materials systems (e.g. Silicon oxide on silicon)

4.6.1.1 Main Measurement Uncertainty Contributions

- Choice of starting assumptions during data analysis, e.g., different dielectric functions available in the literature
- Correlation of variables when extracting multiple parameters in the thin film regime increases the uncertainty for evaluated parameters.

These latter two issues may not be of concern in a particular manufacturing plant because a fixed choice for the silicon dielectric function can be locally standardized on, and the main concern will be detecting small changes from baseline values. However, these issues strongly impact the certification of reference materials. It is not yet possible with SE measurements alone to extract values of all the parameters for a thin film with its interface layer and to provide certified values with 1 % uncertainty at a 95 % confidence level.

4.6.2 Metrological AFM

This is a nanotechnology analogy of the coordinate measuring machine (CMM), the present system for calibrating gauge blocks. It realises real three-dimensional nanometrology, and like the CMM the scanning method over small areas is time consuming. However, a need exists to modify existing, commercial nanoscale measurement systems into routinely available instruments which provide traceable measurements for nanometre scale calibration [24, 30].

In order to ensure traceability at a higher level, several NMIs have developed metrological AFMs. Presently about 17 NMIs worldwide either have metrological AFM in use or in development [31–34]. The largest NMIs have already developed second or third generation metrological AFMs [30, 35–39]. Several international comparisons have been organized in nanometrology [40–44]. One of the main standards is step height, h (0.3–2000 nm).

A step height standard in amorphous silicon dioxide is defined in analogy to the international standard organisation’s (ISO) 5436 definition [43] (Fig. 4.3).

Fig. 4.3 A step-height standard [43]

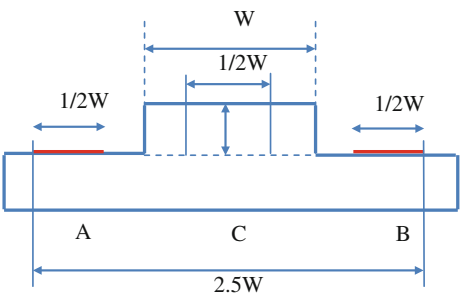


Table 4.7 Metrological AFM

Uncertainty/nm	Measurement time/s	Expanded Uncertainty U_c /nm	Repeatability (NMI)/nm	Traceability
0.1 [132]	15 $\mu\text{m} \times 12 \mu\text{m}$. >5 min	SHN 43.00 1.06 [133]	−0.8	NMI via laser standards
0.1		SHN = 1012.50 1.14 [133]	−0.26	
0.18		SHN = 7.0 0.43 [134]	0.1	
0.18		SHN 1000 nm 0.89 [134]	0.4	

SHN Step height nominal value, U_c at 95 % confidence

A coordinated and optimised approach to the development of traceability at the nanometre scale is pursued in Europe by the European Association of National Metrology Institutes (EURAMET) projects to ensure (Tables 4.7):

- Development of ready-to-use methodologies for calibration of
 - step height
 - line width
 - grating pitch
 - surface roughness parameters
- Development of a methodology for general shape measurements of the various nanostructures
- Study of the effects of data post-processing
- Elaboration of good practice guides to minimise operator influence on the results
- Preparation and realisation of an inter-laboratory comparison with all interested European NMIs on
 - step height
 - line width (Fig. 4.4)

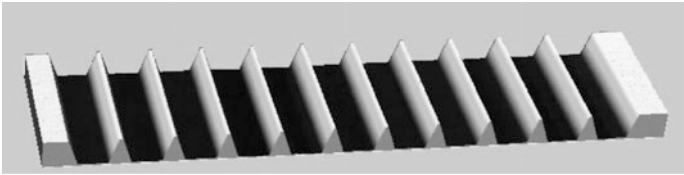


Fig. 4.4 Transfer standard image in AFM [58]

4.6.2.1 Main Measurement Uncertainty Contributions

Laser calibration, step height roughness, sample temperature, imperfect shape of tip.

4.6.3 Scanning Electron Microscopy (SEM)

Projects a two-dimensional image of a three-dimensional object, so is not appropriate for measurements of three-dimensional profiles. The resultant image is processed to realise traceable measurements.

The interaction of secondary electrons with nanostructures being measured changes the dimension. Modelling is required to de-convolute this from the image [45].

SEM can inspect, characterise, and measure huge quantities of nanomaterials (such as carbon nanotubes or cellulose nanocrystals) and over a much greater area of the sample and at higher speed than AFM. However, SEM has poorer resolution than AFM.

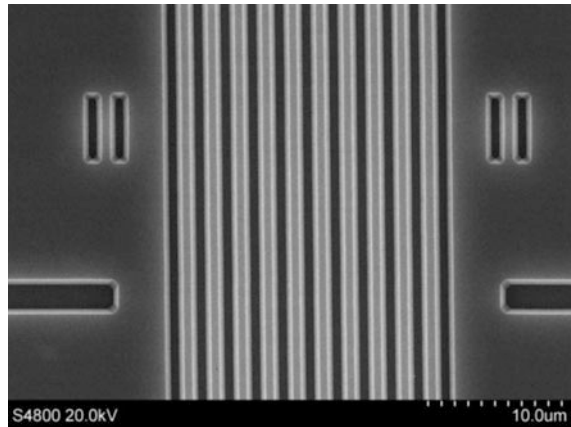
Measurements impact the healthcare (nanobio) and semiconductor industries. A nanostructure can be calibrated by a metrological AFM and can then be used as a 3D calibration artefact for other types of SPMs or SEMs. Gold nanoparticle reference materials of size 10 nm and over are provided by some NMI’s [46] (Table 4.8).

Table 4.8 Metrological SEM

Resolution/nm [46]	Measurement time/s	Expanded uncertainty U _c /nm [135]	Repeatability/nm [136]	Traceability [137]
Voltage 30 keV 0.4 Voltage 1 keV 1.6	60 mm × 130 mm < 60	100–1500 1 %	0.3	Metrological AFM

U_c at 95 % confidence

Fig. 4.5 The same transfer standard from Fig. 4.4 projected as a two-dimensional SEM image at 20.0 kV $10\ \mu\text{m} \times 10\ \mu\text{m}$ [58]



4.6.3.1 Main Measurement Uncertainty Contributions

Sample vibration, Proper instrument calibration, modelling of SEM image, choice of operating parameters/understanding trade-offs, Beam-induced specimen contamination [47] (Fig. 4.5).

4.7 Other Methods for Dimensional Nanometrology

The metrological systems described above must be further refined since they cannot be used in deep holes or for sub-nm precision. For example, the feasibility of utilizing CNT as probe tips is under investigation, since their high aspect ratio will reproduce deep trenches, whilst maintaining high precision due to their nanoscale diameters. Diamondoid tips are necessary both for sub-nm precision and for avoiding wear which occurs on silicon tips [24]. Furthermore, there is the unique challenge of gaining structural information on large-scale nanosystems with random density inhomogeneities such as nanoparticles in suspension and dilute biological samples; ensemble methods are required for this.

Thus the nanostructures in Tables 4.2 and 4.3 can be utilized either as parts of nano-characterisation instruments or as reference materials to validate their performances. Additionally the repeatability and uncertainties of the instruments must be decreased. The following are alternate novel techniques that are currently under investigation for future nanometrology:

4.7.1 X-ray Interferometry (XRI)

This is currently being developed in the European Metrology Research Programme (EMRP) as a technique for disseminating traceable measurements of sub-nanometre step heights. Silicon reference samples have mono-atomic step terraces arranged in an amphitheatre-like structure.

X-ray and visible light interferometry can be combined for better precision. In visible light subdivision of an optical fringe may give rise to errors that are significant at the sub-nanometre level. The X-ray interferometer is used to subdivide the optical fringes. Traceability to the meter is achieved via traceable calibrations of the lattice parameter of silicon and of the laser frequency. Each X-ray fringe corresponds to a displacement equal to the lattice parameter of silicon (~ 0.19 nm) for the (220) lattice planes.

The X-ray interferometer's translation stage must be moved in steps of discrete numbers of X-ray fringes, so servo positioning capability at the picometer level has been developed and includes quadrature counting of X-ray fringes with a resolution of 24 pm [48] (Fig. 4.6).

X-rays from a Mo K_α source are split by the first crystal then recombined. When the analyser is moved in the direction orthogonal to the diffracting planes, the periodic variation in the transmitted and diffracted X-ray intensities (caused by the diffracting-plane spacing) is observed. The analyser displacement δx is measured by

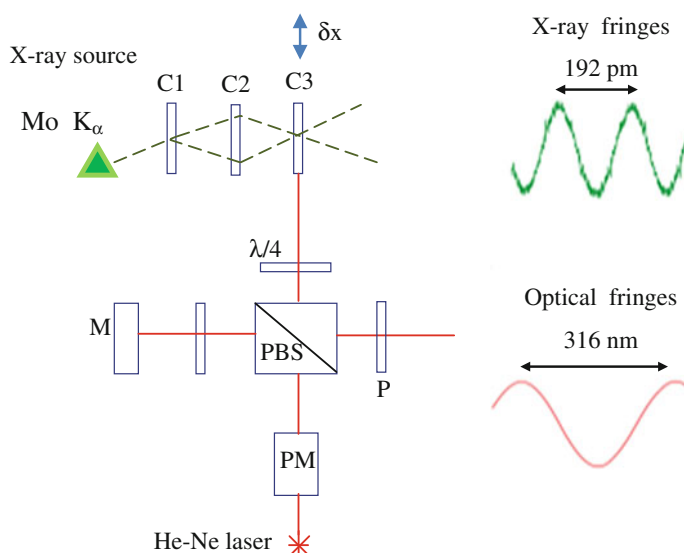


Fig. 4.6 C1, C2, C3 Si crystals, M = mirror, PBS polarising beam splitter, PM phase modulator, P = polariser [59]

the optical interferometer. Polarization encoding and phase modulation enable picometre and nanoradian resolutions to be achieved. Uncertainties at 95 % confidence level are ± 30 pm [49].

4.7.2 *Small Angle X-ray Scattering Diffractometer (SAXS)*

This is deflection of collimated radiation away from a straight trajectory after interacting with structures much larger than the wavelength of the radiation; these deflections are small (0.1 – 10°). It has unique potential for analysing inner structures of disordered systems and gains structural information on large-scales systems (1 nm–hundreds of nm) with random density inhomogeneities. For nanoparticles in suspension and dilute biological samples, this ensemble method is essential for dimensional metrology.

It is an indirect method of obtaining dimensional information of sample structure and requiring intense, monochromatic X-rays of low divergence. This is because biological molecules and nanostructures normally scatter weakly and the measured solutions are dilute. Such directional, high intensity light can only be acquired from synchrotron radiation beam lines of typical size (100×100 μm) to provide detailed data for 10 – 500 nm structures placed with high uniformity in pitch.

By selection of suitable monochromator, X-ray wavelengths from several nanometres down to 0.1 nm at very low angles (typically 0.1 – 10°) can be utilised to determine dimensional measurements of samples from a plot of the scattered electron intensity (I) against momentum transfer(s) at very low angles. The particle diameter and the size distribution are obtained from the scattering pattern as well as structural information: shape and size of molecules, characteristic distances of partially ordered materials, pore sizes, and other data [50, 51].

The solvent intensity is subtracted to obtain a scattering curve representing infinite dilution. Thus, the isotropic intensity distribution remaining is due to random positions and orientations of particles which, for monodisperse non-interacting particles depend on the scattering from a single particle averaged over all orientations (Fig. 4.7).

As for SE, analysis is based on complex fitting to multi-parameter geometric models; for example, a diffraction grating is modelled as a stack of trapezoids, so uniqueness of fit is not guaranteed since it cannot represent the complexity of the sample. A new paradigm is required to elevate scattering and diffraction measurements to quantitative dimensional metrology; this can be achieved by developing a materials model. The uniqueness of the sample structure and shape would then depend on physical parameters (thermodynamics or statistical mechanics) of the model, not the fitting procedure [52].

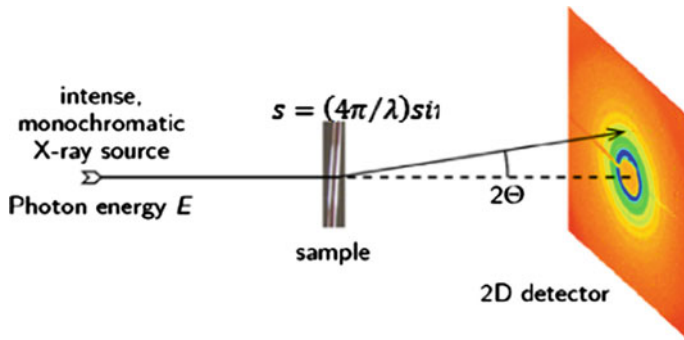


Fig. 4.7 Schematic of a small angle X-ray scattering diffractometer for analysis of dilute, biological and nanostructured samples

4.7.3 Electron/X-ray Diffraction

Suitable for transparent or opaque structures and offers non-destructive, fast collection of data. High-resolution diffraction can provide dimensional information with a relatively simple scan; however, the data have inherent complex error function [53]. Parallel beam diffractometers (traceable to XRI) have self-calibrating goniometers (stable, high-resolution encoders that have been calibrated using the circle closure method) incorporated within them.

To ensure traceability to System Internationale (SI) based units, certification measurements must be performed on this type of goniometer. This can be achieved via Standard Reference Materials (SRM's).

It has been developed to an optical ruler imaging system based on cross-correlation of a feature-rich diffraction pattern (a robust ruler) and silicon sample wafers. The diffraction pattern is projected onto the sample. This produces metrology with a spatial precision of 17.2 nm across silicon wafers [54].

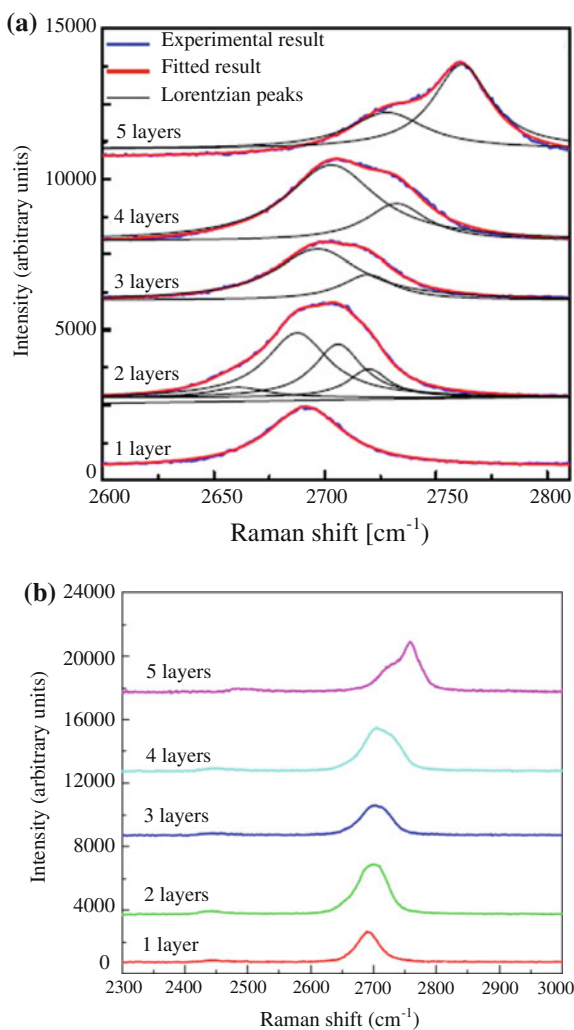
4.7.4 Raman Spectroscopy as a Nanometrology Tool

This is a non-linear optical spectroscopy technique that is capable of providing dimensional information at the nanoscale.

Two-dimensional band features of graphene are highly reproducible and, together with the G-peak position, can be used to count the number of graphene layers.

Raman features remain the same when you place graphene on arbitrary substrate [55, 56] (Fig. 4.8).

Fig. 4.8 Raman spectra of graphene at 300 K, excitation wavelength = 488 nm **a** shows broadening and blue shift **b** shows highly reproducible 2D-band features, and together with the G peak position can be used to count the number of graphene layers



4.8 The Challenge for Nanometrology

The challenge is to develop globally accepted standards and measurement techniques which are reliable, stable, precise and accurate. This will involve the development of innovative systems and standards to meet the needs of the next-generation of manufacturing. These new systems will involve imaging, modelling and manipulating matter at the nanometre length scale.

Various techniques have been discussed above; recently a few standards have been published for terminology, characterization of nanomaterials and nanoparticles

based on some physical phenomena observed on particle interactions or forces at nanoscale using SPM measurement tools (AFM, SEM, Interferometers) [57].

To meet the requirements of future manufacturing and healthcare industries suitable reference materials must be developed and standardised sample preparation methods must be developed. Dimensional metrology must be correlated with functionality by the creation of an international database. Monte Carlo modelling software can be utilised to generate a library of waveforms (for microscopies and spectroscopies) for three-dimensional structures with shape parameters (e.g. widths, angles, radii of curvature) spanning a range of values around that expected together with functional data. The task is then to identify three-dimensional shapes with modelled images that fit the one acquired alongside expected functional behaviour.

As described in Tables 4.4 and 4.5 present second generation nanostructures are more complex than the first generation passive nanotechnology products. They are now functional structures requiring more complex measurement tasks to be performed: shape, size distribution, characteristic distances of partially ordered materials, pore sizes, and other data. The need for more accurate, precise and fast ensemble techniques will only grow as nanotechnology stretches into the third and fourth generation: integrated nanosystems.

Traceability is one of the fundamental requirements for manufacturing and assembly of products, but it is not possible to ensure traceability at nanoscale yet. There are suitable calibration artefacts for achieving traceability, but the measurement equipment required to ensure traceability has not been developed. National reference standards are not always available and uncertainty estimation models have yet to be developed.

4.9 Conclusion

Sub-division of optical laser wavelengths is naturally limited, so research must pursue the realisation of the metre through the lattice parameter of silicon or another suitable crystalline material. Reference samples will be arranged as mono-atomic steps on silicon crystal surfaces or two-dimensional and three-dimensional pyramids and cubes to calibrate nanotechnology instruments. XRI must be developed to replace visible light interferometry. This will involve improving its hardware, optics and data processing and provide a route to traceability to the lattice constant ' a ' as the metre realisation.

High resolution microscopes and integral non-imaging methods (ellipsometry, scatterometry) are widely used for R&D purposes. Many NMI's have developed traceable instrumentation systems primarily for AFM, but only partly for SEM and optical instrumentation due to lack of understanding of the probe-sample interaction. This can be improved by advanced modelling techniques, such as Molecular Dynamic Modelling (MDC), Monte-Carlo-Simulations.

Nanomaterials can be utilised for the validation of the measurement performance of Nano-characterisation instruments. However, suitable reference materials and

methods must be developed for this purpose. Bridging the gap between well characterised reference nanomaterials and manufactured nanomaterials embedded in complex matrix materials is a challenge. There is a great variety of different nanomaterials, so a suitable categorisation approach is required to limit the characterisation efforts to a meaningful manageable amount. The Tables 4.1, 4.2 and 4.3 describe thirteen different main materials and fifteen nano architectures in common use, these are a platform for such reference materials.

The success of nanoelectronics is a precondition for any technological innovation; expected future key challenges for the metrology and standards of to be developed are well described in the ITRS roadmap document. Next generation nanoelectronics technology will develop into three-dimensional complex sub-nm architectures such as FinFET's, and nanometrology is currently being driven there.

None of the existing techniques is able to meet the requirements for reliable metrology of nanomaterials. To meet these requirements suitable reference materials must be developed and standardised sample preparation methods must be developed. Dimensional metrology must be correlated with analytical metrology and functionality, possibly by the creation of an international database.

Innovative techniques such as non-linear optical microscopy and spectroscopy, X-ray scattering and FRET have to be developed further and intercompared with presently employed nanometrology methods such as AFM and SEM to validate accuracy. These innovative techniques have greater capability than present systems for sensing the more complex structures and samples that are of keen interest to the healthcare industry. For nanometrology to move into the biological environment these techniques must be developed.

In addition they have great potential of realising the correlation gap between analytical and dimensional metrology. These are important goals and incorporation of these innovative techniques into nanometrology per se will provide the healthcare and semiconductor industries with remarkable tools.

Acknowledgments I would like to thank Seda Aytekin for her assistance in finding the relevant references.

References

1. SI Brochure www.bipm.org
2. B. Andreas et al., Determination of the Avogadro constant by counting the atoms in a ^{28}Si crystal. Phys. Rev. Lett. **106**, 030801 (2011)
3. <http://www.bipm.org/en/bipm/mass/avogadro>
4. Yeditepe University dimensional calibration services (YUKAL) handbook (Istanbul, Turkey, 2010)
5. E.L. Wolf, *Nanophysics and nanotechnology* (Wiley-VCH Weinheim, Germany, 2005), pp. 28–37
6. Freitas R.A. Jr, Meeting the challenge of building diamondoid medical nanorobots. Int. J. Robot. Res. **28**, 548–557 (2009)
7. A.C. Clarke, *The Fountains of Paradise* Ballantine (Harcourt, New York, 1978)

8. S. Knudsen, L. Golubovic, Rotating space elevators: physics of celestial scale spinning strings. *Eur. Phys. J. Plus* **129**, 1–24 (2014)
9. N.M. Pugno, Towards the Artsutanov's dream of the space elevator: The ultimate design of a 35 GPa strong tether thanks to graphene. *Acta Astronaut.* **82**, 221–224 (2013)
10. P.K. Aravind, The physics of the space elevator. *Am. J. Phys.* **75**, 125–130 (2007)
11. R.R. Birge, Protein-based computers. *Sci. Am.* **272**, 90–95 (1995)
12. K.E. Drexler, *Engines of creation: the coming era of nanotechnology* (Anchor books edition, 1986), pp. 7–12
13. K.E. Drexler, *Nanosystems: Molecular Machinery Manufacturing and Computation* (Wiley-interscience, 1992), pp. 191–248
14. K.E. Drexler, *Engines of Creation 2.0: The Coming Era of Nanotechnology*, 20th anniversary e-book edition, pp. 517–535 (2007)
15. R.D. Kamm, R. Bashir, Creating living cellular machines. *Ann. Biomed. Eng.* **42**, 445–459 (2014)
16. P. Ball, Natural strategies for the molecular engineer. *Nanotechnology* **13**, R15–R28 (2002)
17. A.A. Barhorst, O.P. Harrison, G.D. Bachand, Modelling elasto-mechanical phenomena involved in the motor-driven assembly of nanomaterials, in *ASME 19th International Conference on Design Theory and Methodology/1st International Conference on Micro & Nano Systems*, vol. 3, pp. 669–677 (2008)
18. G.A. Ozin, I. Manners, S. Fournier-Bidoz et al., Dream nanomachines. *Adv. Mat.* **17**, 3011 (2005)
19. D.P. Anderson, T. Kilburn, A pioneer of computer design. *IEEE Ann. Hist. Comp.* **31**, 84 (2009)
20. A.C. Diebold, K. Monahan, Next-generation metrology must meet challenges. *Solid State Technol.* **41**, 50 (1998)
21. K. Seshan (ed.), *The Handbook of Thin Film Deposition: Processes and Technologies*, 6, 241 (2002)
22. M.A. Jamakhani, M.R. Jadhav, G.S. Kamble, V.R. Gambhire, Nanometrology in biological and medical sciences. *Int. J. Adv. Biotechnol. Res.* **2**, 213–223 (2011)
23. L. Peltonen, J. Hirvonen, Pharmaceutical nanocrystals by nanomilling: critical process parameters particle fracturing and stabilization method. *Eur. J. Pharm. Biopharm.* **81**, 214–222 (2012)
24. Euromet (2014) The strategic research agenda of EURAMET
25. JRC Reference Report: Requirements on measurements for the implementation of the European Commission definition of the term nanomaterial
26. J.A. Woollam, B. Johs, C.M. Herzinger et al., Overview of variable angle spectroscopic ellipsometry, in *Critical reviews of optical science and technology: Optical Metrology CR72*, ed. by A-J. Ghanim (1999), pp. 3–28
27. R.K. Leach, C. Giusca, Results from a comparison of optical thin film thickness measurement. NPL Report UK (2009)
28. M.P. Seah, S.J. Spencer, F. Bensebaa et al., Critical review of the current status of thickness measurements for ultra-thin SiO₂ on Si: Part V results of a CCQM pilot study. *Surf. Interfaces Anal.* **36**, 1269–1303 (2004)
29. B. Belzer, K. Eberhardt, D. Chandler-Horowitz et al., Thin film reference materials development final report for CRADA CN-1364. NIST Gaithersburg 2012 (1998)
30. Euromet EMPIR Call, Traceable AFM measurement capability. Industry and Research Potential Selected Research Topic: SRT-r **05**, 1–3 (2014)
31. V. Korpelainen, J. Seppä, A. Lassila, Design and characterization of MIKES metrological atomic force microscope. *Prec. Eng.* **34**, 735–744 (2010)
32. S. Ducourtieux, B. Poyet, J. David, Development of a metrological atomic force microscope with minimized Abbe error and differential interferometer based real-time position control. *Measur. Sci. Technol.* **22**, 094010 (2011)
33. B. Babic, C.H. Freund, J. Herrmann et al., Metrological scanning probe microscope based on a quartz tuning fork detector. *J. Micro/Nanolith MEMS MOEMS* **11**, 011003 (2012)

34. P. Klapetek, M. Valtr, M. Matula, A long-range scanning probe microscope for automotive reflector optical quality inspection. *Measur. Sci. Technol.* **22**, 094011 (2011)
35. G. Dai, F. Pohlenz, X. Min et al., Metrological large range scanning probe microscope. *Rev. Sci. Instrum.* **75**, 962–969 (2004)
36. I. Misumi, S. Gonda, Q. Huang et al., Sub-hundred nanometre pitch measurement using an AFM with differential laser interferometers for designing usable lateral scales. *Measur. Sci. Technol.* **16**, 2080–2090 (2005)
37. J.A. Kramar, Nanometre resolution metrology with the molecular measuring machine. *Measur. Sci. Technol.* **16**, 2121–2128 (2005)
38. J. Haycocks, K. Jackson, Traceable calibration of transfer standards for scanning probe microscopy. *Prec. Eng.* **29**, 168–175 (2005)
39. V. Korpelainen, A. Lassila, Calibration of a commercial AFM: traceability for a coordinate system. *Measur. Sci. Technol.* **18**, 395–403 (2007)
40. F. Meli, NANO4-1D Gratings. Final report (2000)
41. J. Garnaes, K. Dirscherl, NANO5—2D Grating—Final report CCL-S4. *Metrologia. Tech. Suppl.* **45**, 04003 (2008)
42. L. Koenders, R. Bergmans, J. Garnaes et al., Comparison on nanometrology: nano 2—step height. *Metrologia* **40**, 04001 (2003)
43. L. Klapetek, P. Koenders, F. Meli, G.B. Picotto, Comparison on step height measurements in the nano and micrometre range by scanning force microscopes. *Metrologia* **43**, 04001 (2006)
44. G. Hamann, Positioning in sub-nanometer resolution High-Accuracy and Traceable AFM Measurements on Semiconductors and Step Standards. Whitepaper Physik Instrumente GmbH (2013)
45. M.T. Postek, A.E. Vladar, Modelling for accurate dimensional scanning electron microscope metrology: then and now. *Scanning Special Iss.* **33**, 111–125 (2011)
46. M.T. Postek, A.E. Vladar, J.S. Villarrubiar, *Nanometrology solutions using an ultra-high resolution in-lens SEM* (NIST White Paper, 2011)
47. J. Cazaux, Errors in nanometrology by SEM. *Nanotechnol* **15**, 1195–1199 (2004)
48. A. Yacoot, U. Kuetgens, Sub-atomic dimensional metrology: developments in the control of x-ray interferometers. *Measur. Sci. Technol.* **23**, 074003 (2012)
49. G. Basile, P. Becker, A. Bergamin et al., Combined optical and X-ray interferometry for high-precision dimensional metrology. *Proc. Royal Soc. A-Math. Phys. Eng. Sci.* **456**, 701–729 (2000)
50. L. Koenders, *Defining a European Strategy for Nanometrology- Dimensional and Mechanical Aspects* (Co-nanomet Physikalisch-Technische Bundesanstalt, Budapest, 2011)
51. M. Krumrey, G. Gleber, F. Scholze, J. Wernecke, Synchrotron radiation-based x-ray reflection and scattering techniques for dimensional nanometrology. *Measur. Sci. Technol.* **22**, 094032 (2011)
52. NIST, Dimensional metrology for nanofabrication. Materials Science and Engineering Division White Paper (2009)
53. J.P. Cline, D.A. Windover, Materials structure characterization. Diffraction Metrology and Standards NIST White Paper (2014)
54. N. Yoshimizu, A. Lal, C.R. Pollock, Nanometrology optical ruler imaging system using diffraction from a quasiperiodic structure. *Opt. Express* **18**, 20827–20838 (2010)
55. I. Calizo, S. Ghosh, D. Teweldebrhan et al., Raman nanometrology of graphene on arbitrary substrates and at variable temperature. *Proc. SPIE* **7037**, 70371B (2008)
56. A.A. Balandin, Thermal transport in graphene and graphene multi-layers. Seminar Report Nano-Device Laboratory, Univ California-Riverside (2008)
57. A. Vladár, PML Scientists Develop 3D SEM Metrology for 10 nm Structures. Semiconductor and dimensional division NIST White Paper (2014)
58. P. Todua, Nanometrology nanoscale reference materials for high-resolution microscopy. national scientific metrology center/center for surface and vacuum research (CSVr) Moscow Institute of Physics and Technology (MIPT). Vienna, Austria (2011)

59. P. Becker, M. Jentschel, G. Mana, G. Zosi, Towards an atomic realisation of the kilogram: the measurements of N_A and $N_{A,h}$, in *Metrology and Fundamental Constants*, ed. by T. Haensch, S. Leschiutta, A. Wallard (IOS Press, Italy, 2007), pp. 519–530
60. C. He, S. Wu, N. Zhao et al., Carbon-encapsulated Fe_3O_4 nanoparticles as a high-rate lithium ion battery anode material. *ACS Nano* **7**, 4459–4469 (2013)
61. H.W. Kroto, J.R. Heath, S.C. O'Brien, R.F. Curl, R.E. Smalley, C₆₀: Buckminsterfullerene. *Nature* **318**, 162–163 (1985)
62. H. Isobe, S. Hitosugi, T. Yamasaki et al., Molecular bearings of finite carbon nanotubes and fullerenes in ensemble rolling motion. *Chem. Sci.* **4**, 1293–1297 (2013)
63. S. Iijima, T. Ichihashi, Single-shell carbon nanotubes of 1-nm diameter. *Nature* **363**, 603–605 (1993)
64. M.M. Shulaker, G. Hills, N. Patil et al., Carbon nanotube computer. *Nature* **501**, 526 (2013)
65. A.K. Geim, K.S. Novoselov, The rise of graphene. *Nat. Mater.* **6**, 183–191 (2007)
66. T. Georgiou, R. Jalil, B.D. Belle et al., Vertical field-effect transistor based on graphene- WS_2 heterostructures for flexible and transparent electronics. *Nat. Nanotechnol.* **8**, 100–103 (2013)
67. R.C. Merkle, R.A. Jr Freitas, Theoretical analysis of a carbon-carbon dimer placement tool for diamond mechanosynthesis. *J. Nanosci. Nanotechnol.* **3**, 319–324 (2003)
68. D.G. Allis, K.E. Drexler, Design and analysis of a molecular tool for carbon transfer in mechanosynthesis. *J. Comp. Theor. Nanosci.* **2**, 45–55 (2005)
69. A. Yamamoto, T. Hamada, I. Hisaki et al., Dynamically deformable cube-like hydrogen-bonding networks in water-responsive diamondoid porous organic salts. *Angew. Chem. Int. Ed.* **52**, 1709–1712 (2013)
70. D. Wang, H.L. Xin, R. Hovden et al., Structurally ordered intermetallic platinum-cobalt core-shell nanoparticles with enhanced activity and stability as oxygen reduction electrocatalysts. *Nat. Mater.* **12**, 81–87 (2013)
71. H. Li, D.X. Chen, Y.L. Sun et al., Viologen-mediated assembly of and sensing with carboxylatopillar[5]arene-modified gold nanoparticles. *J. Am. Chem. Soc.* **135**, 1570–1576 (2013)
72. J. Liu, W. Bu, L. Pan et al., NIR-triggered anticancer drug delivery by upconverting nanoparticles with integrated azobenzene-modified mesoporous silica. *Angew. Chem. Int. Ed.* **52**, 4375–4379 (2013)
73. L. Wang, I. Meric, P.I. Huang et al., One-dimensional electrical contact to a two-dimensional material. *Science* **342**, 614–617 (2013)
74. P.J. Kueffer, C.A. Maitz, A.A. Khan et al., Boron neutron capture therapy demonstrated in mice bearing EMT6 tumors following selective delivery of boron by rationally designed liposomes. *Proc. Nat. Acad. Sci. USA* **110**, 6512–6517 (2013)
75. H. Wu, G. Yu, K. Pan et al., Stable Li-ion battery anodes by in-situ polymerization of conducting hydrogel to conformally coat silicon nanoparticles. *Nat. Commun.* **4**, 1943 (2013)
76. Y.H. Fu, A.I. Kuznetsov, A.E. Miroshnichenko et al., Directional visible light scattering by silicon nanoparticles. *Nat. Commun.* **4**, 1527 (2013)
77. H. Choi, S.J. Ko, Y. Choi et al., Versatile surface plasmon resonance of carbon-dot-supported silver nanoparticles in polymer optoelectronic devices. *Nat. Photonics* **7**, 732–738 (2013)
78. C. Wang, H. Xu, C. Liang et al., Iron oxide @ polypyrrole nanoparticles as a multifunctional drug carrier for remotely controlled cancer therapy with synergistic antitumor effect. *ACS Nano* **7**, 6782–6795 (2013)
79. G. Chen, Y. Zhao, G. Fu et al., Interfacial effects in iron-nickel hydroxide-platinum nanoparticles enhance catalytic oxidation. *Science* **344**, 495–499 (2014)
80. Y. Hu, J.O. Jensen, W. Zhang et al., Hollow spheres of iron carbide nanoparticles encased in graphitic layers as oxygen reduction catalysts. *Angew. Chem. Int.* **53**, 3675–3679 (2014)
81. B.J. Sanghavi, A.K. Srivastava, Adsorptive stripping voltammetric determination of imipramine trimipramine and desipramine employing titanium dioxide nanoparticles and

- an Amberlite XAD-2 modified glassy carbon paste electrode. *Analyst* **138**, 1395–1404 (2013)
82. S. Sarina, H. Zhu, E. Jaatinen et al., Enhancing catalytic performance of palladium in gold and palladium alloy nanoparticles for organic synthesis reactions through visible light irradiation at ambient temperatures. *J. Am. Chem. Soc.* **135**, 5793–5801 (2013)
 83. U. Rocha, C. Jacinto, S.W. Ferreira et al., Subtissue thermal sensing based on neodymium-doped LaF₃ nanoparticles. *ACS Nano* **7**, 1188–1199 (2013)
 84. J. Burschka, N. Pellet, S.J. Moon et al., Sequential deposition as a route to high-performance perovskite-sensitized solar cells. *Nature* **499**, 316 (2013)
 85. X. Zhu, X. Niu, H. Zhao et al., Doping ionic liquid into Prussian blue-multiwalled carbon nanotubes modified screen-printed electrode to enhance the nonenzymatic H₂O₂ sensing performance. *Sens. Act. B* **195**, 274–280 (2014)
 86. C.C. Druck, J.L. Pozzobon, G.L. Callegari et al., Adhesion to Y-TZP ceramic: study of silica nanofilm coating on the surface of Y-TZP. *J. Biomed. Mat. Res. B-App. Biomater.* **103**, 143–150 (2015)
 87. J.H. Heo, S.H. Im, J.H. Noh et al., Efficient inorganic-organic hybrid heterojunction solar cells containing perovskite compound and polymeric hole conductors. *Nat Photon* **7**, 487–492 (2013)
 88. T. Stephenson, Z. Li, B. Olsen et al., Lithium ion battery applications of molybdenum disulfide (MoS₂) nanocomposites. *Energy Environ. Sci.* **7**, 209–231 (2014)
 89. T. Choi, S.H. Kim, C.W. Lee et al., Synthesis of carbon nanotube-nickel nanocomposites using atomic layer deposition for high-performance non-enzymatic glucose sensing. *Biosens. Bioelectron.* **63**, 325–330 (2015)
 90. F. Song, X. Li, Q. Wang et al., Nanocomposite hydrogels and their applications in drug delivery and tissue engineering. *J. Biomed. Nanotechnol.* **11**, 40–52 (2015)
 91. J. Wallentin, N. Anttu, D. Asoli et al., InP nanowire array solar cells achieving 13.8 % efficiency by exceeding the ray optics limit. *Science* **339**, 1057–1060 (2013)
 92. P. Yang, Y. Ding, Z. Lin et al., Low-cost high-performance solid-state asymmetric supercapacitors based on MnO₂ nanowires and Fe₂O₃ nanotubes. *Nano Lett.* **14**, 731–736 (2014)
 93. L. Fu, W. Cai, A. Wang et al., Photocatalytic hydrogenation of nitrobenzene to aniline over tungsten oxide-silver nanowires. *Mat. Lett.* **142**, 201–203 (2015)
 94. X. Zhou, L.J. Wan, Y.G. Guo, Binding SnO₂ nanocrystals in nitrogen-doped graphene sheets as anode materials for lithium-ion batteries. *Adv. Mater.* **25**, 2152–2157 (2013)
 95. Z. Bian, Zhang T.P. Tachikawa et al., Au/TiO₂ superstructure-based plasmonic photocatalysts exhibiting efficient charge separation and unprecedented activity. *J. Am. Chem. Soc.* **136**, 458–465 (2014)
 96. C.J. Hages, S. Levchenko, C.K. Miskin et al., Improved performance of Ge-alloyed CZTGeSSe thin-film solar cells through control of elemental losses. *Prog. Photovolt.* **23**, 376–384 (2015)
 97. H.S. Kim, J.W. Lee, N. Yantara et al., High efficiency solid-state sensitized solar cell-based on submicrometer rutile TiO₂ nanorod and CH₃NH₃PbI₃ perovskite sensitizer. *Nano Lett.* **13**, 2412–2417 (2013)
 98. J. Yu, J. Jin, B. Cheng et al., A noble metal-free reduced graphene oxide-CdS nanorod composite for the enhanced visible-light photocatalytic reduction of CO₂ to solar fuel. *J. Mat. Chem. A* **2**, 3407–3416 (2014)
 99. S. Danwittayakul, M. Jaisai, J. Dutta, Efficient solar photocatalytic degradation of textile wastewater using ZnO/ZTO composites. *App. Cat. B-Environ.* **163**, 1–8 (2015)
 100. L. Bareket, N. Waiskopf, D. Rand, G. Lubin, M. David-Pur et al., Semiconductor nanorod-carbon nanotube biomimetic films for wire-free photostimulation of blind retinas. *Nano Lett.* **14**, 6685–6692 (2014)
 101. J.Y. Jeng, Y.F. Chiang, M.H. Lee et al., CH₃NH₃PbI₃ perovskite/fullerene planar-heterojunction hybrid solar cells. *Adv. Mater.* **25**, 3727–3732 (2013)

102. Y. Liu, J. Zhao, Z. Li et al., Aggregation and morphology control enables multiple cases of high-efficiency polymer solar cells. *Nat. Commun.* **5**, 5293 (2014)
103. Y. Lin, J. Wang, Z.G. Zhang et al., An electron acceptor challenging fullerenes for efficient polymer solar cells. *Adv. Mater.* **27**, 1170–1174 (2015)
104. T. Choi, S.H. Kim, C.W. Lee et al., Synthesis of carbon nanotube-nickel nanocomposites using atomic layer deposition for high-performance non-enzymatic glucose sensing. *Biosens. Bioelectron.* **63**, 325–330 (2015)
105. O. Chen, J. Zhao, V.P. Chauhan et al., Compact high-quality CdSe-CdS core-shell nanocrystals with narrow emission linewidths and suppressed blinking. *Nat. Mater.* **12**, 445–451 (2013)
106. C. Clavero, Plasmon-induced hot-electron generation at nanoparticle/metal-oxide interfaces for photovoltaic and photocatalytic devices. *Nat. Photon.* **8**, 95–103 (2014)
107. S.G. Kumar, K.S.R. Koteswara, Zinc oxide based photocatalysis: tailoring surface-bulk structure and related interfacial charge carrier dynamics for better environmental applications. *RSC Adv.* **5**, 3306–3351 (2015)
108. S. Agarwal, A. Greiner, J.H. Wendorff, Functional materials by electrospinning of polymers. *Prog. Polym. Sci.* **38**, 963–991 (2013)
109. C. Zhu, X. Mu, P.A. van Aken et al., Single-layered ultrasmall nanoplates of MoS₂ embedded in carbon nanofibers with excellent electrochemical performance for lithium and sodium storage. *Angew. Chem. Int. Ed.* **53**, 2152–2156 (2014)
110. C. Liu, L. Wang, Y. Tang et al., Vertical single or few-layer MoS₂ nanosheets rooting into TiO₂ nanofibers for highly efficient photocatalytic hydrogen evolution. *App. Cat. B-Environ.* **164**, 1–9 (2015)
111. S.I. Kim, J.S. Lee, H.J. Ahn et al., Facile route to an efficient NiO supercapacitor with a three-dimensional nanonetwork morphology. *ACS Appl. Mater. Interfaces* **5**, 1596–1603 (2013)
112. R.K. Joshi, P. Carbone, F.C. Wang et al., Precise and ultrafast molecular sieving through graphene oxide membranes. *Science* **343**, 752–754 (2014)
113. H. Zhang, X. Hou, Z. Yang et al., Bio-inspired smart single asymmetric hourglass nanochannels for continuous shape and ion transport control. *Small* **11**, 786–791 (2015)
114. M.H. Oh, T. Yu, S.H. Yu et al., Galvanic replacement reactions in metal oxide nanocrystals. *Science* **340**, 964–968 (2013)
115. S. Dasgupta, T. Auth, G. Gompper, Shape and orientation matter for the cellular uptake of nonspherical particles. *Nano Lett.* **14**, 687–693 (2014)
116. J. Min, F. Wang, Y. Cai et al., Azeotropic distillation assisted fabrication of silver nanocages and their catalytic property for reduction of 4-nitrophenol. *Chem. Commun.* **51**, 761–764 (2015)
117. Z. Yu, B. Duong, D. Abbitt et al., Highly ordered MnO₂ nanopillars for enhanced supercapacitor performance. *Adv. Mater.* **25**, 3302–3306 (2013)
118. D.D. Wang, G.Z. Xing, F. Yan et al., Ferromagnetic (Mn N)-codoped ZnO nanopillars array: experimental and computational insights. *Appl. Phys. Lett.* **104**, 022412 (2014)
119. D.M. Drotlef, E. Appel, H. Peisker et al., Morphological studies of the toe pads of the rock frog *Staurois parvus* (family: Ranidae) and their relevance to the development of new biomimetically inspired reversible adhesives. *Int. Focus* **5**, 0036 (2015)
120. K. Wu, T. Rindzevicius, M.S. Schmidt et al., Wafer-scale leaning silver nanopillars for molecular detection at ultra-low concentrations. *J. Phys. Chem. C* **119**, 2053–2062 (2015)
121. W. Liu, Y.Q. Feng, X.B. Yan et al., Superior micro-supercapacitors based on graphene quantum dots. *Adv. Funct. Mater.* **23**, 4111–4122 (2013)
122. X. Liu, Z. Chang, L. Luo et al., Hierarchical Zn_xCo_{3-x}O₄ Nanoarrays with High Activity for Electrocatalytic Oxygen Evolution. *Chem of Mat* **26**, 1889–1895 (2014)
123. M. Yu, J. Chen, J. Liu et al., Mesoporous NiCo₂O₄ nanoneedles grown on 3D graphene-nickel foam for supercapacitor and methanol electro-oxidation. *Electrochim. Acta* **151**, 99–108 (2015)

124. X. Wang, L. Jiao, K. Sheng et al., Solution-processable graphene nanomeshes with controlled pore structures. *Sci. Rep.* **3**, 1996 (2013)
125. C.F. Guo, T. Sun, Q. Liu et al., Highly stretchable and transparent nanomesh electrodes made by grain boundary lithography. *Nat. Commun.* **5**, 3121 (2014)
126. B. Piccione, D.S. Gianola, Tunable thermoelectric transport in nanomeshes via elastic strain engineering. *Appl. Phys. Lett.* **106**, 113101 (2015)
127. S. Costeux, L. Zhu, Low density thermoplastic nanofoams nucleated by nanoparticles. *Polymer* **54**, 2785–2795 (2013)
128. A. Wittstock, M. Baeumer, Catalysis by unsupported skeletal gold catalysts. *Acc. Chem. Res.* **47**, 731–739 (2014)
129. S. Sen, D. Liu, G.T.R. Palmore, Electrochemical reduction of CO₂ at copper nanofoams. *ACS Catalysis* **4**, 3091–3095 (2014)
130. S. Costeux, I. Khan, S.P. Bunker et al., Experimental study and modelling of nanofoams formation from single phase acrylic copolymers. *J. Cell. Plastic.* **51**, 197–221 (2015)
131. C. Anders, E.M. Bringa, H.M. Urbassek, Sputtering of a metal nanofoam by Au ions. *Nucl. Instrum. Methods. Phys. Res. Sect. B-Beam Interact. Mater. Atom* **342**, 234–239 (2015)
132. R. Köning, J. Flügge, D. Hüser et al., Dimensional micro- and nanometrology at PTB, in *2011 Proceedings of 8th International Conference on Measurement*, Smolenice, Slovakia 95:106 (2011)
133. L. Koenders, EUROMET project 707 step height standards. PTB Final Report (2005)
134. V. Korpelainen, J. Seppä, A. Lassila, Measurement strategies and uncertainty estimations for pitch and step height calibrations by metrological AFM. *Proc. SPIE* **8036**, 80360Q1 (2011)
135. J.B. Warren, B.J. Panessa-Warren, A comparison of nanometrology using TEM and SEM. *Microsc. Microanal.* **11**(2), 1932–1933 (2005)
136. T. Ikegami, A. Yamaguchi, M. Tanaka et al., Evolution and future of critical dimension measurement system for semiconductor processes. Hitachi high Technologies Inc White paper 203–209 (2011)
137. Nanosensors, *Transfer Standards for Scanning Probe Microscopy* (Nano World AG Data Sheet, 2013)

Chapter 5

Terahertz Devices and Systems for the Spectroscopic Analysis of Biomolecules—“Complexity Great and Small”

Anthony John Vickers and David Crompton

Abstract The world has an ever-growing need to understand the complexity of biomolecules from fields as far ranging as drug design Hajduk and Geer, *Nat Rev* 6 (3):211, 2007, [1] and Congreve et al., *Prog Med Chem* 53(1), 2014, [2], crop characterization Ge et al., 15(6):12560, 2015, [3] and organic electronics Gao et al., *Phys Rev Lett* 114(12):128701, 2015, [4]. There are many scientific techniques for studying bio-molecules namely mass spectroscopy (MS), nuclear magnetic resonance (NMR), droplet single biomolecule studies, spectroscopy, biocalorimetry, chromatography, crystallography, electrophoresis, and bioinformatics. In this chapter the focus will be on the study of biomolecules using the relatively new technique of terahertz spectroscopy. This chapter will draw on the THz spectroscopy work of the authors and others to demonstrate the need to understand both the biomolecule and its water environment, which provide both great (biomolecule) and small (water) complexity.

5.1 Introduction

Terahertz spectroscopy has developed over the past two decades and has started to make an impact on the understanding of biomolecules. One of the key issues regarding the study of biomolecules is to study them in their natural aqueous environment as this offers the greatest possibility of understanding their natural behavior. This provides both a challenge and an opportunity in THz spectroscopy as water strongly absorbs in the THz frequency range [5]. This strong water THz absorption could be seen as a disadvantage making observation of the biomolecule absorption difficult if not impossible. However, as we shall discuss in this chapter, it can be argued that the modeling and study of biomolecules cannot be undertaken in

A.J. Vickers (✉) · D. Crompton
School of Computer Science and Electronic Engineering, University of Essex,
Wivenhoe Park, Colchester, Essex CO4 3SQ, UK
e-mail: vicka@essex.ac.uk

isolation from its aqueous environment. Taking the proposition to another level it is perhaps possible to argue that biomolecules have some water molecules so strongly associated with them that they are truly a part of the protein. This is the basis of a new drug design software package, WaterMap [6].

5.2 THz Time Domain Spectroscopy

Terahertz time domain spectroscopy (THz-TDS) describes a technique that has become widely adopted in the field of THz spectroscopy. In THz-TDS, picosecond THz pulses are generated from a photoconductive antenna pumped by a femtosecond laser. High frequency sampling of the generated THz pulse is achieved by delaying a probe laser, in time, to sample the THz EM field. By changing the time interval of the probe laser, the THz EM field is sampled at each time interval until a complete THz pulse has been sampled. The resulting THz pulse contains all the frequency information of the THz beam in the time domain, which can be extracted by a Fourier transform algorithm. A typical frequency range for photoconductive antennas is 0.1–4 THz. Detection of the THz pulse is achieved using techniques such as photodetectors or electro-optic sampling. In the case of a THz photodetector the THz pulse creates an electric field across a pair of electrodes deposited on the photodetector material, for example GaAs. The THz pulse biases the detector, a signal being generated by the delayed optical femtosecond pulse. The signal, for a given time delay, is then averaged using a lock-in amplifier. Modulation of the lock-in amplifier is either achieved through the modulation of the THz emitter bias or through the optical modulation of the femtosecond laser pulse train. THz photodetectors have characteristic resonant frequencies that effect the efficiency as a function of frequency. Electro-optic (EO) sampling is a technique often used in place of the photoconductive detector in THz-TDS spectroscopy. Unlike the photoconductive antennas, EO sampling does not have resonant frequencies, so has the same efficiency in THz detection across the frequency range. The principle of the EO technique is based on an EO medium, often constructed from a ZnTe crystal, which introduces birefringence when under the influence of an external THz EM field. Typically for a THz-TDS system, the EO medium induces an elliptical birefringence to the probe laser beam, dependent on the strength of the instantaneous THz EM field. A Wollaston prism splits the now elliptically polarized probe laser into component p and s polarizations, which are subsequently detected by two balanced photodiodes. The difference in p and s magnitude is then directly proportional to the strength of the THz EM field. As with the photoconductive antenna, the probe laser is delayed in time to detect the entire THz pulse. Without the THz EM field, the probe laser will still undergo some equilibrium birefringence resulting in a non-zero detection. To compensate for this effect, a quarter-wave plate is placed between the Wollaston prism and the EO medium. The quarter-wave plate is then rotated till the closest value of 0 V is achieved on the balanced photo detector. The p and s polarizations should now have equal magnitude at the detector for the null result.

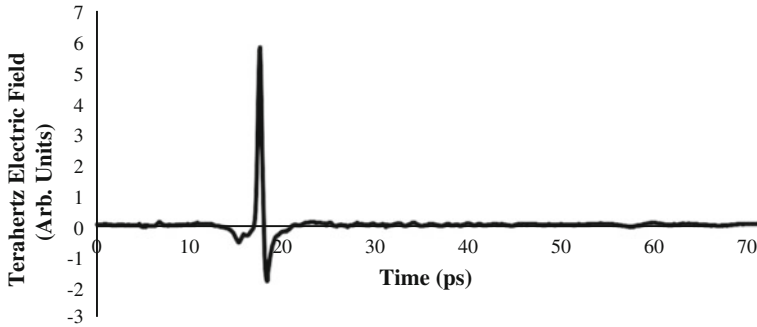


Fig. 5.1 A typical THz pulse recorded using a time domain system

The overall system consisting of a THz emitter, pumped by a femtosecond laser, which also provides the variable time delayed probe pulse for the THz detector, is referred to as a THz Time Domain Spectrometer (TDS) and provides a sampling rate beyond current sampling oscilloscope technology [7]. A THz pulse sampled using the above method is displayed in Fig. 5.1.

The choice of THz emitter and detector for a time domain system influences the overall design of the spectrometer. For detectors and emitters without any associated focusing element, a four parabolic mirror system is usually adopted. The first parabolic mirror gathers the THz emission from the emitter and generates a parallel beam. The second parabolic mirror focuses this parallel beam onto the sample cell. A third parabolic mirror gathers the THz beam as it expands after the sample cell, again producing a parallel THz beam. The fourth parabolic mirror then focuses the THz beam onto the detector. An alternative design, in the case where both THz emitter and detector have associated focusing elements, requires only two parabolic mirrors [8]. This design is sketched in Fig. 5.2. In this case a photoconductive emitter and detector are used. Both emitter and detector have Si lenses positioned at the point of emission and detection. The Si lenses can be manipulated in x–y to center the lens on the emitter and detector to achieve optimum signal generation and detection.

In the spectrometer shown in Fig. 5.2 the sample cell is an Attenuated Total Reflection (ATR) cell, although in this design the ATR sample cell can be replaced by a transmission cell. Attenuated total reflectance cells work on the principal of the evanescent wave effect observed at the point of total internal reflection. This evanescent wave extends out in an exponentially decaying wave that interacts with a sample.

This method then avoids direct transmission through the sample. To achieve total internal reflection of the THz beam, a general design utilises a dove prism placed within a THz beam path, shown in Fig. 5.2. In biological systems, this is a key advantage over transmission systems, which require short sample path lengths. The evanescent tail is also highly sensitive to changes in sample concentrations making ATR ideal for complex biological molecules that are difficult to purify in

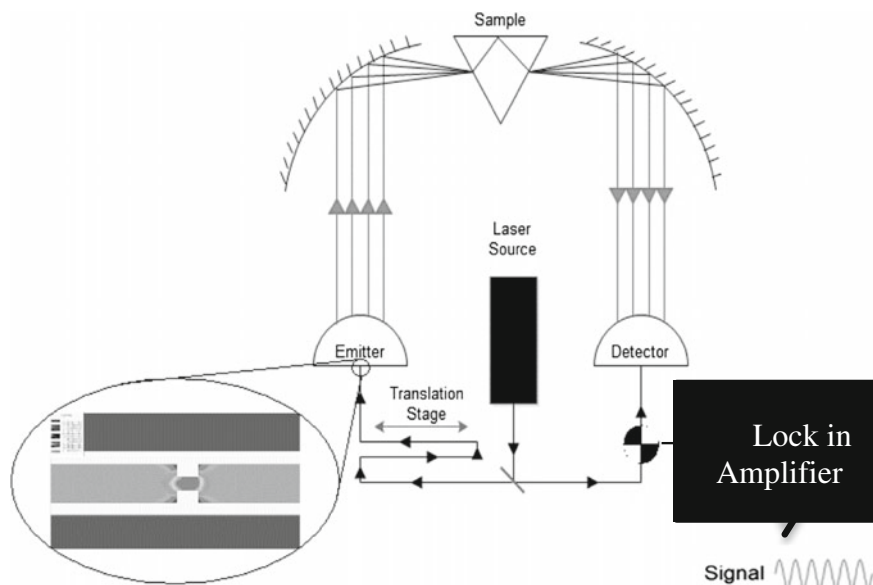


Fig. 5.2 A typical THz time domain spectrometer showing a photoconductive emitter and detector, and an Attenuated Total Reflection (ATR) sample stage [8]

high concentrations. Techniques like high throughput protein-drug candidate screening require accurate results with minimum protein sample usage. The concentrations required for protein-drug binding in THz transmittance systems are too high to make this a viable option. This is one example where the sensitivity of the ATR unit could be a dominating feature in its adoption in the THz frequency range.

5.3 Small Complexity—Simple Polar Liquids

One of the key areas of interest for THz spectroscopy is its ability to probe the low frequency dynamics of water and other polar liquids. These dynamics are linked to the continuous breaking and reformation of the hydrogen bonds that universally bind polar liquids together across a hydrogen bond network. As water is a solvent present in almost all biological processes, this network will almost certainly interact with these processes through its dynamics. In the following section we explore the composition of the hydrogen bond dynamics and how these dynamics underpin the low frequencies in the THz regime.

ATR THz studies of simple polar liquids have been undertaken by a number of groups including ourselves [9, 10]. We prepared an aqueous solution series (0, 10, 20 ..., 100 %) of propan-2-ol. In turn, a drop of each solution placed on the ATR sample cell, creating a sample with a depth of approximately 0.5 mm. The average

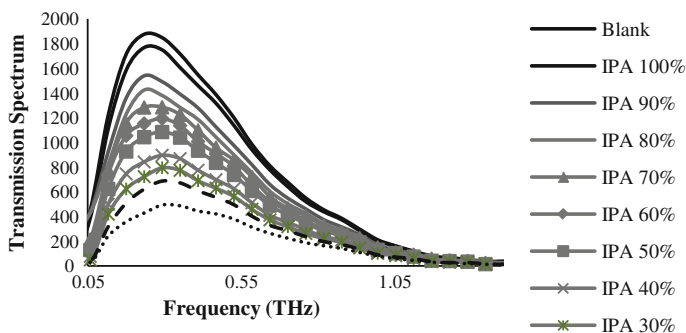


Fig. 5.3 The frequency spectrum of varying IPA concentrations was calculated from the FFT of THz pulse response. A high detection indicates less THz radiation has been absorbed by the sample as more radiation has reached the detector. IPA has significantly less THz absorbance compared to water with the max peak showing a linear response to the addition of IPA. The standard error is approximately 8 units

spectra of each solution was calculated from three consecutive scans with 10 averages per time step. The peltier cooler was set to maintain the sample cell temperature at 21 °C, and the spectrometer was dry nitrogen purged to prevent water absorption by the THz beam on its passage through the spectrometer. Figure 5.3 shows the THz frequency spectrum for the propan-2-ol concentration series. The frequency spectrum, at each concentration, is defined by singular large, broad peak. Broad peaks are a ubiquitous feature in THz spectroscopy and are seen in THz studies using aqueous solutions [11]. The broad peaks are caused by molecular vibrations associated with THz absorption not occupying discrete energy levels and instead harmonically fluctuate at frequencies across the THz regime. It is therefore difficult to assign specific vibrational modes to frequencies as many vibrational modes can overlap in the same frequency space.

The large disparity between water and propan-2-ol THz absorption is best explained by the physical properties of the molecules that cause THz absorbance. Molecules have two key characteristics that absorb THz radiation. The first is the collective motions of large molecules, typically manifested in the torsional motion of hundreds of atoms along a molecule's backbone. Propan-2-ol is a relatively small molecule and so these collective motions are unlikely to contribute to propan-2-ol THz absorption. The second characteristic that leads to THz absorption is the existence of an extended dynamical hydrogen bond network. Propan-2-ol has a functional hydroxyl group located at the center of the carbon chain that is able to form hydrogen bonds. So intuitively propan-2-ol must be able to form a hydrogen bond network that absorbs THz radiation. The lack of THz absorbance then indicates these hydrogen bonds are significantly different to those found in water.

The geometry of water allows the acceptance and donation of up to four hydrogen bonds freely arrange from any angle on the molecule, whereas propan-2-ol has a carbon chain that is likely to sterically hinder the formation of

Table 5.1 Recalculation of atomic partial charges

Propan-2-ol molecule constituents	Atomic point charges	
	Jorgensen charges	New propan-2-ol charges
q(O)	-0.683	-0.620
q(HO)	0.418	0.388
q(C)	-0.18	-0.456
q(C)	-0.12	-0.444
q(CO)	0.145	0.148
q(H)	0.06	0.141

hydrogen bonds. Propan-2-ol is also a moderately larger molecule compared to water, creating a lower hydrogen bond density per volume of solution. It can be concluded the observed propan-2-ol absorbed THz spectrum is caused by the lack of hydrogen bond density and the relative small size of the molecule.

To help clarify the underlying hydrogen bond dynamics of the polar liquid concentration series, molecular dynamics are used to view the hydrogen bond formation events associated with the THz frequency range. This is followed by an analysis of the hydrogen bond network to determine the dynamical effects that are not seen solely in the THz-TDS spectrum.

For this study, the OPLS-All Atom force field [12] was selected for its accuracy in liquid simulation dynamics. Water was modeled by the TIP4P water model as it works well with OPLS-All Atom and accurately models the tetrahedral shape of the water molecule. The propan-2-ol topology parameters were based on the topology of propan-1-ol, also built by Jorgensen et al. [13], using GAUSSIAN [14] to calculate the charge distribution for propan-2-ol (Table 5.1).

The initial system for the simulation consisted of 1000 alcohol molecules with a XYZ box size of 3.8 nm for ethanol and 4.3 nm for propan-2-ol. To generate the different solution concentrations, a percentage volume of alcohol molecules were removed at random from the system and then solvated with the TIP4P water. The appropriate number of waters to add was selected by the GROMACS solvation algorithm, with a van der Waals tolerance of 0.105 and a density of 1. The number of water molecules added to each alcohol concentration is shown in Table 5.2. The true percentage of alcohol to water molecules after solvation is also shown in the table, which differs slightly from the alcohol volume to water ratio.

Each concentration solution was simulated for 1 ns to allow adjustment of the total size of its system box to equilibrate the pressure. This had the effect of slightly increasing the box size at higher concentrations of alcohol due to their lower density than water. From the production runs the hydrogen bond auto correlation, the total number of hydrogen bonds and hydrogen bond length were taken from twenty averages of 10 ps simulations. The temperature was maintained at 300 K for these simulations with an NPT ensemble and a 2 fs step between each time frame output.

From the simulation, the number of water molecules added to the system does not increase linearly with decreasing alcohol concentration. Instead a smooth gradual increase in the number of water molecules and hydrogen bonds is seen as

Table 5.2 Polar liquid system parameters

Alcohol volume %	Alcohol molecules in box	Water molecules in 3.8 nm ethanol box	Water molecules in 4.3 nm propan-2-ol box
0	0	1728	2617
5	50	1555	2397
10	100	1389	2170
30	300	821	1395
50	500	400	845
70	700	165	412
80	800	74	200
90	900	30	64
95	950	10	39
100	1000	0	0

the alcohol concentration changes. The THz absorption aqueous solution series therefore cannot be explained by a simple linear model of increasing water hydrogen bonds.

The non linear increase in water molecules can be explained by the efficiencies in packing alcohol-alcohol, alcohol-water and water-water bonds in the system. Water, with its optimised geometry can efficiently place hydrogen bonds, and in the simulation, pure water forms 3.34 bonds per water molecule. As alcohol is added to the system, alcohols are sterically hindered from forming hydrogen bonds by their carbon chains, causing inefficient packing of water molecules and a decrease in the relative density of the system.

The reason why we observe a linear increase THz absorption in the experimental results, but not equally linear increasing number of hydrogen bonds can be explained in the terms of the dynamics of the hydrogen bonds. Strong and long-lived hydrogen bonds have been previously correlated with higher THz absorption [15]. In two systems with equal hydrogen bond density, the stronger and longer-lived hydrogen bonds will absorb more THz radiation than those with weaker and shorter lifetimes.

The strength of hydrogen bonds can be assumed to directly relate to the lifetime of the hydrogen bond. The strength-lifetime relationship is calculated from the molecular dynamic simulations using the Hydrogen Bond Auto Correlation function [16]. The existence of hydrogen bonds is calculated for any bond forming pair within 35 nm. The HBAC function used in this experiment allows for the re-formation of the same hydrogen bond, that is a donor and acceptor may move away from the hydrogen bond geometry and subsequently move back. This is then counted as one continuous hydrogen bond.

HBAC is calculated for 1, 2.5 and 5 picoseconds for the propan-2-ol concentration series. A high HBAC can indicate strong and long-lived hydrogen bonds relative to those with a lower HBAC [15]. Below an HBAC value of 0.5, half of the original hydrogen bonds have been broken. At high propan-2-ol concentrations,

extreme hydrogen bond lifetimes are seen, suggesting these hydrogen bonds are locked in position by steric constraints of the alcohol.

The high HBAC values in the alcohol systems then seem to suggest alcohol forms strong hydrogen bonds, despite the perceived inefficient packing of the water molecules. The HBAC results are then counter intuitive, as alcohols are expected to weaken hydrogen bonds through a weak hydroxyl group and carbon steric effects. As a further test of hydrogen bond strength the average distance distribution of the hydrogen bonds was calculated for the propan-2-ol concentration series. Hydrogen bond length can be directly correlated to the strength of hydrogen bonds, with shorter bonds having a greater strength than the equivalent longer bonds [17]. The result of the distance distribution indicates the bonds in high alcohol concentrations are stretched relative to those found in water by approximately 0.01 Å between pure propan-2-ol and water. These results then disagree with the assumption that HBAC predicts the hydrogen bonds in the alcohol solutions are stronger. From the distance distribution of hydrogen bonds and the HBAC it can be concluded the hydrogen bonds in the alcohol solutions are both sterically hindered and relatively weak in comparison to bulk water.

The contradiction in results between HBAC and hydrogen bond distance distribution suggests the THz absorption of hydrogen bonds cannot be simply explained in terms of the hydrogen bond strength. Instead it appears it is the hydrogen bond dynamics, controlled by both the strength of the bond and steric hindrance from the alcohol, that dominate the THz absorbance at the low hydrogen bond density seen in the high propan-2-ol concentrations.

This leads to the conclusion that the alcohol solutions have weaker and sterically hindered hydrogen bonds, which have increased THz absorption properties compared to non-sterically hindered and strong hydrogen bonds of bulk water. After 50 % water concentration the weaker, but more abundant water-water hydrogen bonds begin to dominate the system and increase in THz absorption through their efficient packing of hydrogen bonds at high water concentrations. It is then the combine effects of sterically hindered hydrogen bonds at high propan-2-ol concentrations and the sheer number of packed hydrogen bonds in pure water that cause the linear increase in THz absorption observed in the experimental results. A similar result has been observed when water is exposed to a hydrophobic region of a protein [18]. The hydrophobic regions of the protein act in the same manner as the alcohol solutions, actively hindering the dynamics of hydrogen bonding, by repelling water away from the protein surface. This led to an observed increase in THz absorption, which this report attributes to the increase of sterically hindered waters.

5.4 Great Complexity—Proteins

Low frequency dynamics are not exclusive to hydrogen bond networks, but also exist as large collective motions across proteins. These collective motions change as the protein undergoes biological processes and often underpin the activation or deactivation of the protein. THz radiation is able to probe these collective motions

and in this section we use THz radiation to monitor the activation of G-Protein coupled receptors.

G-Protein coupled receptors (GPCRs) are one of the largest and most diverse group of membrane proteins found within the animal kingdom [19]. GPCRs are characterized by a common structure of seven transmembrane helices (TM1, TM2,...,TM7) that allow them to form cytoplasmic and extracellular domains across the cellular lipid bilayer [20]. Their principal role is to act as a transmembrane messenger and regulator of many biological functions, which has made them a highly prized drug target, with 40–60 % of all modern medicines believed to target a GPCR [21, 22]. Due to their significance in drug candidate research, GPCRs are also one of the most widely studied classes of proteins.

The typical modes of action for GPCR drugs or ligands can be broadly classed in two groups of agonists and antagonists, where the drug promotes or inhibits the GPCR activation [23]. By binding to the GPCR, the drug or ligand can alter the flow of activation energy across the protein by altering the molecular constraints of the protein. This flow of activation energy is strongly linked to the low frequency normal modes of the molecule, which sample the local spatial geometry [24, 25]. The sampling of the local geometry can lead to conformational changes, which ultimately activate or deactivate the protein. The activated and deactivated low frequency modes appear to be conserved between GPCRs [26] as the low frequency modes seem to be an inherent property of the secondary structure of the protein [27]. The THz frequency regime is able to directly probe these low frequency modes and therefore sample the activation and conformational state of the protein.

In this section, the low frequency dynamics of GPCR Bovine Rhodopsin are studied in an aqueous environment in both a membrane and non-membrane solution. Bovine rhodopsin is a key protein in monochromatic vision in cattle. Within the GPCR super family rhodopsin is a type A, the largest class of GPCRs. The proteins in this family are called opsins (visual purple), light sensitive proteins, which are non-homologous, but have evolved to perform a similar function and share a similar structure. Opsins are further sub divided into two groups, type I Prokaryotic [28] and type II Animal [29].

Rhodopsin being a photoreactive protein makes it an ideal GPCR to study, as it can be simply activated via light illumination without affecting chemical composition of the solution. ATR THz-TDS spectroscopy is then used to observe any change in the low frequency dynamics of the protein in both the inactive dark and active light state. Any shift in low frequency dynamics should indicate a conformational change upon protein activation. By testing the rhodopsin in membrane and non-membrane solution, the impact of the membrane on the low frequency dynamics can also be established.

Further to ATR THz-TDS spectroscopy of rhodopsin, elastic networks are computationally calculated from crystallised rhodopsin proteins in both activation states. From these elastic models, the normal modes of the protein are calculated, which are synonymous with the low frequency modes. The normal modes of the elastic models are then used to help interpret data in the experimentally observed THz spectrum.

Bovine rhodopsin and all opsin type proteins are activated through photoisomerisation of the chromophore, where specific wavelengths of the electromagnetic spectrum act as the primary stimulus [30]. The rhodopsin chromophore retinal is linked to the protein via a protonated Schiff base located at Lys296 on the TM7 helix [20]. This location places the positively charged Schiff base into a hydrophobic area neutralized by a neighboring Glu113 residue. This structure helps stabilize the inactive rhodopsin molecule until it is activated [31]. The retinal, in the inactive state, is initially in the *cis* conformation. When illuminated, photoisomerisation of the retinal occurs, resulting in a *cis* to all-*trans* formation between the carbon-11 and carbon-12 double bonds [20]. The initial *cis* to all-*trans* takes just a few femtoseconds once photo excited [32]. This extension of the chromophore leads to the series of conformational changes that ultimately lead to the activation of the rhodopsin molecule within a few milliseconds [33].

Inside the core of rhodopsin exists what is believed to be a highly conserved water network that interacts with the protein through hydrogen bonding [34]. The hydrogen bond network is thought to help mediate conformational changes across the rhodopsin protein and assist the color tuning of the rhodopsin chromophore that helps stabilize the ground state of the GPCR [35]. The sensitivity of THz radiation to the hydrogen bond dynamics and the importance of hydrogen bonds to protein structure potentially makes THz radiation highly sensitive to the activation state of the protein.

Activation-critical hydrogen bonding is also observed between residues within the hydrophobic core of the protein. During activation the β -ionone ring moves toward TM5, this movement rearranges the hydrogen bond network leading to a weakening of the hydrogen bond to the Glu122 residue. A change in Glu122 absorption occurs between the Meta Ila and the Meta I Ib states suggesting that in the Meta Ila state the hydrogen bond network is still in the inactive conformation [36]. The disruption of the Schiff base upon activation also causes a rearrangement of the core hydrogen bond network around TM1, TM2, and TM7 with the help of Asp83 [37].

Previous THz-TDS studies of rhodopsin have used thin film dried samples [26] and small aqueous sample transmission path lengths (15 μm) [38] to prevent high water absorptions. For the experimental method in this chapter, THz-TDS is combined with an ATR sample unit to achieve high sensitivity despite the aqueous solvent. The theoretical method uses NMA to calculate the low frequency modes of the protein by using an elastic model based on the atomic coordinates from x-ray crystallography of protein structures. From each low frequency mode, the fluctuations of the protein residues can be seen.

For the THz-TDS experiment, rhodopsin is activated by white light illumination, so any change in the THz spectrum between activation states should be due to the low frequency dynamics and hydrogen bonding within the protein. Rhodopsin is tested both in a lipid bilayer membrane and non-membrane solution to observe if the membrane has any noticeable impact on the protein's low frequency dynamics. Any change in the low frequency dynamics between inactive, active, membrane and non-membrane results should be revealed by a change in the THz spectrum and the level of residue fluctuations in the NMA.

Bovine rhodopsin rod outer segments were prepared from bovine retina using sucrose gradient centrifugation. The samples were urea washed and suspended in an aqueous buffer (20 mM Tris-HCl [pH 7.2], 2 mM EDTA, 1 mM 2-mercaptoethanol, 0.1 mM PMSF). The resulting concentration was approximately 2 mg/ml [38]. The rhodopsin samples were prepared in both membranes (BRM) and non-membrane (BR). A dim red light is used for lighting during preparation to prevent accidental activation of the protein.

The THz-TDS with ATR unit, described in Sect. 5.2, was used to resolve the THz spectrum of each bovine rhodopsin sample. The experiments were carried out in a controlled environment with a temperature of 20 °C and a dry nitrogen purged atmosphere. The samples were kept in the dark to avoid accidental activation and were pipetted in turn onto the sample facet of the dove prism. A total of four scans for each sample was taken, with an average of 30 scans of each step of the THz-TDS translation stage. This resulted in four sets of data, the inactive bovine rhodopsin membrane dark state (BRMDark) and bovine rhodopsin non-membrane dark state (BRDark). Followed by after 1 min illumination with a white LED, the active bovine rhodopsin membrane light state (BRMLight) and bovine rhodopsin non-membrane light state (BRLight).

Figure 5.4 shows, the peak terahertz electric field response of BRMDark, BRMLight, BRDark and BRLight samples. The total amplitude of the THz electric field appears to be unaffected by the protein's activation state or its membrane environment. The phase of the THz pulse, however, varies between each of the samples. For the membrane and non-membrane solutions BRMDark-BRDark and BRMLight-BRLight respective samples, the phase change appears as a +0.1 ps time shift. This is not unexpected as the membranes in the bovine rhodopsin solution will displace some of the water, which has a high refractive index in the THz region that will affect the phase of the THz pulse [15].

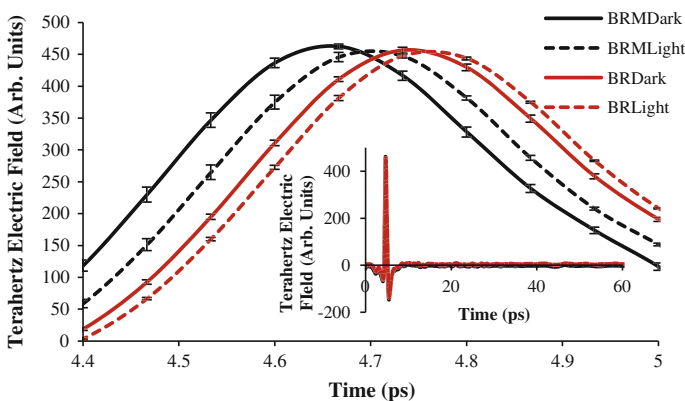


Fig. 5.4 THz-TDS of bovine rhodopsin: key differences between the samples THz pulses can be seen at the 4.7 ps peak. *Error bars* show the standard deviation between data averages. *Inset* The entire THz pulse recorded for each sample

Upon protein activation, a phase change is also observed between the inactive dark state and active light state. For membrane and non-membrane solutions, this phase change from dark to light state is $+0.63$ and $+0.6$ ps respectively. As the chemical composition of the sample from dark to light is unchanged, this suggests the rhodopsin protein has undergone a conformational change. The more pronounced phase change in the membrane solution may indicate the membrane has a non-trivial influence on the conformational change of the protein.

From the THz electric field response, a fast Fourier transform algorithm is used to generate the THz transmission frequency spectrum of the sample. Background correction is applied to the sample spectrum by division by the reference transmission spectrum. In this instance, the reference spectrum is the ATR prism with no sample. The result of this calculation is shown in Fig. 5.5.

There are three key features of the rhodopsin absorption spectrum that can be observed in these results. First it appears rhodopsin in both aqueous and membrane systems absorb more radiation at lower THz frequencies. The low frequency modes of the protein are expected to significantly contribute the absorptions in this region at these frequencies. Secondly, the rhodopsin in membrane absorbs less THz radiation than the rhodopsin in aqueous solution. The membranes displacement of highly absorbing waters may account for the decrease in THz absorption between the rhodopsin samples, as previous reports have shown membranes absorb significantly less THz radiation than water [39]. Although the membranes do not significantly absorb THz radiation, they are likely to affect the THz absorbing low frequency dynamics of the protein through steric or electrostatic interaction due to their close proximity. The potential of this interaction is further investigate in the difference spectrum below. The third feature is the light activated state is observed to absorb more THz radiation than the dark inactive state for both aqueous and membrane solutions. This indicates the unlocking of low frequency modes upon protein activation, with the increased low frequency modes absorbing more THz radiation.

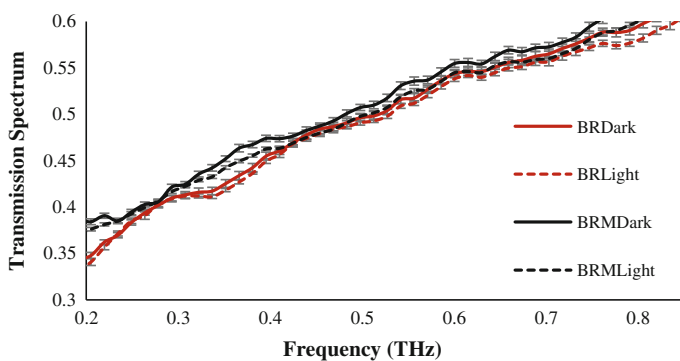


Fig. 5.5 THz frequency spectrum of bovine rhodopsin: the results have been corrected for spectrometer absorptions and the standard error is calculated from the average of four sample scans

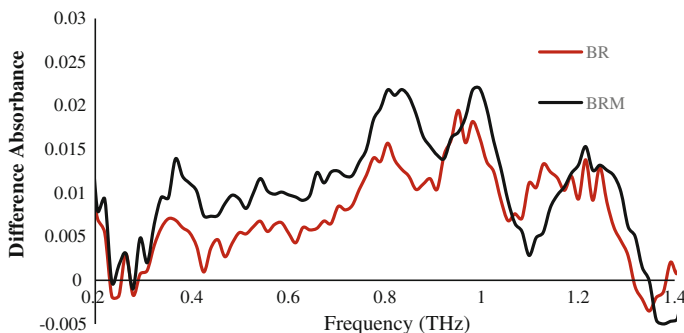


Fig. 5.6 Difference absorbance of rhodopsin in aqueous and membrane solutions

To help elucidate the THz characteristics of the activation states of rhodopsin, the difference spectrum between BRMDark-BRMLight and BRDark-BRLight is calculated. This effectively removes the THz characteristics of the water and membrane molecules, leaving the contributions of the low frequency modes from the protein inactive-active conformational change. The results of the difference spectrum are shown in Fig. 5.6.

There appears to be significant differences between the inactive state and active state between 0.2 and 1.4 THz for both membrane and non-membrane systems. This difference is more pronounced in the membrane solution with an average difference of 0.005 (Arb.Units) higher than the rhodopsin in aqueous solution. The membrane appears to have a significant effect on the low frequency dynamics of rhodopsin, making a more defined response between the two activation states. This suggests the membrane may stabilise the low frequency dynamics between inactive and active rhodopsin, where the non-membrane protein low frequency modes are less restricted. After 1.4 THz, the difference spectrum remains at approximately zero (Arb. Units), indicating the majority of the activation low frequency modes lie between 0 and 1.4 THz.

For the purpose of molecular dynamic simulations, there are effectively four samples in the rhodopsin THz-TDS experiments, inactive and active rhodopsin both in- and not-in membrane solution. Inactive rhodopsin was provided by the protein data bank (PDB) file 1U19 [40], where the retinal remains chemically bound to the protein. Active rhodopsin was modelled by the PDB file 2 × 72 [41], which has been crystallised after the dissociation of the retinal ligand. Topologies for the protein was generated using GROMACS. A pre-equilibrated 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) membrane lipid bilayer, by Andreas Kukol, was used to imbed the membrane proteins [42].

To generate the non-membrane samples, each of the inactive and active state proteins was solvated with SPC water with a 20 Å buffer away from the protein surface. The membranes systems were made by superimposing the membrane onto the protein in VMD, which is then followed by a 20 Å SPC buffer solvation. GROMACS is then used to “grow” the protein in the membrane removing any

steric clashes [43]. This results in four model systems; inactive no-membrane, inactive membrane, active no-membrane and active membrane. Each system is then minimised to 1 kJ mol^{-1} , followed by 200 ps of NVT ensemble and 1 ns of NPT ensemble. Standard Berendsen temperature and pressure coupling were used to maintain the system at 1 ATM and 300 K, with periodic boundary conditions.

The normal mode analysis for each model rhodopsin system was calculated using an anisotropic network model (ANM). The calculation was performed on all backbone carbons excluding the initial 33 residues and all residues up to residue 334. These excluded residues are the C and N terminus of the protein, which are loose chains that may mask the low frequencies of the protein. The first six modes of the calculation are trivial, representing the rotational movement of the protein, so are also ignored.

The results of this calculation, for the first three non-trivial modes, are shown in Fig. 5.7, as the mean square fluctuation of each residue. From Fig. 5.7, there can be seen significant differences between each active state, typically by a change in peak amplitudes at residues 100, 240 and 370 in non-membrane and residues 145 and 240 in membrane solution. The residues that form these peaks are prominently found in the loop regions of the protein, so a relatively high mobility is expected. In membrane there are significantly fewer peaks in the square fluctuation across the protein's residues, suggesting the membrane is suppressing some of the mean square fluctuations. A change is also observed in the frequency of the normal modes between each of the active states. In non-membrane solution the frequency appears to undergo a red shift, for the first and third modes, when the protein is activated. In membrane, only a small red shift is observed for the first mode and large blue shifts for the second and third mode.

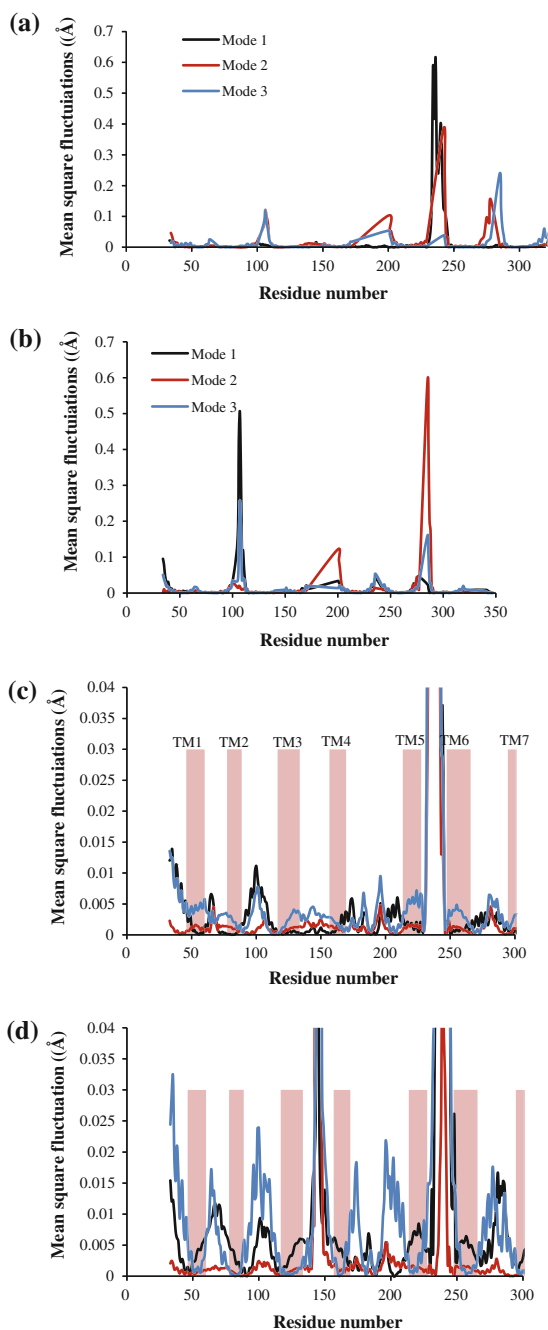
Figure 5.7c, d shows in more detail the square fluctuations of α -helical structure in the membrane solution, which are highlighted in red. A clear increase in square fluctuation can be seen in modes 1 and 3, with mode 2 remaining relatively the same except at loop residue 150. This same distinct increase in α -helical structure square fluctuations, between activation states, is ill-defined in the non-membrane systems.

In conclusion, in membrane, there is a greater observed difference spectrum between the low frequency modes of the inactive and active states compared to the non-membrane solutions, but membranes themselves have minimal THz absorption [39]. NMA of the protein shows the membrane suppresses the low frequency modes in the α -helical and loop structures of the protein. Despite the suppression, the membrane still allows for an increase in the low frequency modes upon protein activation.

Without membrane, the low frequency modes of the α -helices and loops appear less defined, combined with significantly larger square fluctuations across the protein for both activation states. Suggesting the membranes may act as a stabilising agent, helping the protein maintain the inactive and active conformation.

It seems likely that the protein has evolved to predominantly exist in either inactive or active state when in its natural membrane environment. Taken out of this environment has allowed the protein to sample more conformational states, shown

Fig. 5.7 **a** Mean square fluctuation as a function of residue number for the non-membrane inactive bovine rhodopsin (1U19). **b** Mean square fluctuation as a function of residue number for the non-membrane active bovine rhodopsin (2×72). **c** Mean square fluctuation as a function of residue number for the membrane inactive bovine rhodopsin (2×72). **d** Mean square fluctuation as a function of residue number for the membrane active bovine rhodopsin (2×72)



by the increased NMA square fluctuations in the non-membrane solutions. Despite being able to sample more conformations, many of these may be inactive or partially active. The membrane, however, stabilises the inactive state and may force the protein towards a full activation by restricting these square fluctuations in the protein. The increased difference, seen in the THz difference spectrum, between the membrane activation states may then be explained by the membrane interaction with the low frequency modes.

In the discussion above, the changes in the THz spectrum of the protein were linked to the low frequency dynamics of the protein's α -helical structure. Also the water and hydrogen bonding plays an active role in the function and structure of rhodopsin. Other reports have shown that the dynamics of the hydrogen bonding are directly influenced by the solute and the low frequency dynamics of biomolecules can be coupled to the solvating water [40, 42]. This suggests the change in the THz between activation states may be partly influenced by the waters coupled to the low frequency dynamics of the protein.

GLU122 and surrounding residues have been identified as key residues on forming protein activation dependent hydrogen bonds [36]. To assess the hydrogen bonds around this residue in the model systems, HBAC analysis and the total number of the hydrogen bonds are calculated 1 nm around GLU122. The HBAC is shown to be higher in the inactive form than the active form, indicating sterically hindered bonds prior to protein activation. This is in agreement with the observed weakening of the hydrogen bonds during activation of the protein [36].

Normally an increased restriction and density of hydrogen bonds would indicate stronger THz absorbing properties of the system, which has been observed in the previous section on polar liquids. Ignoring the low frequency mode contributions, this would suggest the inactive state around GLU122 absorbs more THz radiation than the active state of the protein. But the THz absorption spectrum of the activation states reveals the reverse is true. This indicates the low frequency dynamics of the protein are then the most significant contributor to the THz characteristics of the system. The low frequency modes, however, may actively contribute to the weakening of the hydrogen bond network, allowing the ligand to fully dissociate from the system.

Future molecular dynamics simulations may reveal if the low frequency dynamics are directly coupled to the weakening of the hydrogen bonds at the ligand binding site. Further THz-TDS experiments may also explore the intermediary states of rhodopsin activation, where the research in this chapter is solely focused on the inactive and active states of rhodopsin. Of particular interest is the change from the Meta I state to the Meta II state, where the TM6 helix shifts releasing activation energy. The sensitivity of the ATR THz-TDS should be able to detect the TM6 shift as an unlocking of low frequency modes. It may then be possible to link the shift as a critical step to protein activation, by measuring its contribution to the active state low frequency modes.

5.5 Conclusions

THz TD spectroscopy has been demonstrated as tool to study biomolecules in an aqueous environment. The use of an attenuated total reflection sample cell enables rapid sample change, avoids the need for standardized volume cells, and does not require the spectrometer to be repeatedly nitrogen purged when changing samples. We have show through the experimental results and the simulations that water plays a very important role in the overall dynamics of biomolecules. This work and the work of others has provided some initial steps towards a long term goal of applying biomeolecular engineering to the rapid design of drugs, an aim that would provide significant benefit to world health.

References

1. P.J. Hajduk, J. Geer, *Nat. Rev. Drug Discov.* **6**(3), 211 (2007)
2. M. Congreve, J.M. Dias, F.H. Marshall, *Prog. Med. Chem.* **53**, 1 (2014)
3. H. Ge, Y. Jiang, F. Lian, Y. Zhang, S. Xia, *Sensors* **15**(6), 12560 (2015)
4. F. Gao, W. Tress, J. Wang, O. Inganas, *Phys. Rev. Lett.* **114**(12), 128701 (2015)
5. J. Xu, W. Plaxco, S.J. Allen, *Protein Sci.* **15**, 1175 (2006)
6. Nimbus Therapeutics, Our technologies, [nimbustx.com. http://nimbustx.com/our-approach/our-technologies](http://nimbustx.com/our-approach/our-technologies). Accessed 15 July 2015
7. P. Thoma, V. Judin, *Optik & Photonik* **8**, 58 (2013)
8. D. Crompton, A.J. Vickers, in *13th Biennial Baltic Semiconductor Conference*, vol. 13 (2012)
9. D. Crompton A.J. Vickers, in *Computer Science and Electronic Engineering Conference*, vol. 47 (2012)
10. H. Naito, Y. Ogawa, K. Shiraga, N. Kondo, I. Osaka, A. Kubota, in *International Symposium on Systems Integration (SII)*, vol. 192 (2011)
11. D.H. Choi, H. Son, S. Jung, J. Park, W.Y. Park, O.S. Kwon, G.S. Park, in *Infrared. Millimeter, and THz Waves*, 1 (2012)
12. W.L. Jorgensen, in *Encyclopedia of Computational Chemistry* (2002)
13. W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **118**, 11225 (1996)
14. M. Frisch, G. Trucks, H.B. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. Petersson, Inc., Wallingford, CT, **270**, 271 (2009)
15. M. Heyden, *Fakultat fur Chemie und Biochemie* (Ruhr-Universitat Bochum, Bochum, 2010)
16. D. van der Spoel, E. Lindahl, B. Hess, A. Van Buuren, E. Apol, P. Meulenhoff, D. Tieleman, A. Sijbers, K. Feenstra, R. van Drunen, *Gromacs User manual version 3.3* (2008)
17. S. Goryainov, *Physica B* **407**, 4233 (2012)
18. T.Q. Luong, P.K. Verma, R.K. Mitra, M. Havenith, *Biophys. J.* **101**, 925 (2011)
19. K. Palczewski, T. Kumasaka, T. Hori, C.A. Behnke, H. Motoshima, B.A. Fox, I.L. Trong, D. C. Teller, T. Okada, R.E. Stenkamp, *Science Signaling* **289**, 739 (2000)
20. D.M. Rosenbaum, S.G. Rasmussen, B.K. Kobilka, *Nature* **459**, 356 (2009)
21. J.P. Overington, B. Al-Lazikani, A.L. Hopkins, *Nat. Rev. Drug Discovery* **5**, 993 (2006)
22. Y. Arinaminpathy, E. Khurana, D.M. Engelman, M.B. Gerstein, *Drug Discovery Today* **14**, 1130 (2009)
23. B. Bosier, E. Hermans, *Trends Pharmacol. Sci.* **28**, 438 (2007)
24. K.C. Chou, *Biophys. J.* **48**, 289 (1985)
25. K.C. Chou, *Biophys. Chem.* **30**, 3 (1988)

26. R. Balu, H. Zhang, E. Zukowski, J.-Y. Chen, A. Markelz, S. Gregurick, *Biophys. J.* **94**, 3217 (2008)
27. K. Yamamoto, K. Tominaga, H. Sasakawa, A. Tamura, H. Murakami, H. Ohtake, N. Sarukura, *Biophys. J.* **89**, L22 (2005)
28. M.X. Ruiz-Gonzalez, I. Marín, *J. Mol. Evol.* **58**, 348 (2004)
29. Y. Shichida, *Philos. Trans. R. Soc. B: Biol. Sci.* **364**, 2881 (2009)
30. J. Buczyłko, J.C. Saari, R.K. Crouch, K. Palczewski, *J. Biol. Chem.* **271**, 20621 (1996)
31. G.B. Cohen, D.D. Oprian, P.R. Robinson, *Biochemistry* **31**, 12592 (1992)
32. R.S. Liu, G.S. Hammond, *Photochem. Photobiol. Sci.* **2**, 835 (2003)
33. J.-P. Vilardaga, M. Bünemann, C. Krasel, M. Castro, M.J. Lohse, *Nat. Biotechnol.* **21**, 807 (2003)
34. R. Nygaard, T.M. Frimurer, B. Holst, M.M. Rosenkilde, T.W. Schwartz, *Trends Pharmacol. Sci.* **30**, 249 (2009)
35. T. Okada, Y. Fujiyoshi, M. Silow, J. Navarro, E.M. Landau, Y. Shichida, *Proc. Natl. Acad. Sci.* **99**, 5982 (2002)
36. H.-W. Choe, Y.J. Kim, J.H. Park, T. Morizumi, E.F. Pai, N. Krauß, K.P. Hofmann, P. Scheerer, O.P. Ernst, *Nature* **471**, 651 (2011)
37. P.T. Ho, J.M. Moran, K.Y. Lo, *Astrophys. J. Lett.* **616**, L1 (2004)
38. A. Vickers, R. Dudley, P. Reeves, C. Reynolds, S. Gadde, T. Nithyanand, Y. Ma, in *Infrared and Millimeter Waves*, vol. 395 (2007)
39. K. Tielrooij, D. Paparo, L. Piatkowski, H. Bakker, M. Bonn, *Biophys. J.* **97**, 2484 (2009)
40. F. Despa, A. Fernández, R.S. Berry, *Phys. Rev. Lett.* **93**, 228104 (2004)
41. J. Standfuss, P.C. Edwards, A. D'Antona, M. Fransen, G. Xie, D.D. Oprian, G.F. Schertler, *Nature* **471**, 656 (2011)
42. P. Fenimore, H. Frauenfelder, B. McMahon, F. Parak, *Proc. Natl. Acad. Sci.* **99**, 16047 (2002)
43. M.G. Wolf, M. Hoeffling, C. Aponte-Santamaria, H. Grubmüller, G. Groenhof, *J. Comput. Chem.* **31**, 2169 (2010)

Chapter 6

Recent Progress in XAFS Study for Semiconducting Thin Films

Takafumi Miyanaga and Takashi Azuhata

Abstract X-ray absorption fine structure (XAFS) is a powerful structural analysis tool for studying the local structure in thin films. In the recent developments of this technology, high-brightness light emitting diodes and laser diodes are applied to semiconducting InGaN nanolayers. In this chapter, we provide a brief review on basic theory and measurements of XAFS, and then present some recent results of XAFS application to *c*-plane (0001) InGaN, *m*-plane (1 $\bar{1}$ 00) InGaN, *m*-plane AlGaN, and *c*-plane MgZnO to reveal the atomic distribution and the interatomic distances in the films.

6.1 Introduction

X-ray absorption fine structure (XAFS) is a well-established structural analysis technique used for the characterization of materials in a variety of forms [1–3]. In particular, XAFS is a powerful tool in investigating local structures of thin layers composed of two or more elements. The characteristics of XAFS are element-specificity and electron-shell-specificity, which make XAFS a useful probe complementary to other techniques, to obtain structural information and electronic structure. XAFS is considered to be established in the 1970s with pioneering work by Lytle, Sayers, and Stern [4] after struggling for many years [5]. XAFS has progressed tremendously in the development of experimental techniques and theoretical understanding. Modern synchrotron light sources have played an important role and have provided a large variety of new applications, such as structural characterization of nanostructures and time-resolved studies for dynamic processes. Due to the progress in theoretical understanding, more exact XAFS functions can be calculated with computer code. XAFS is presently used over many fields; physics, chemistry, material sciences, geology, biology, and environmental science.

T. Miyanaga (✉) · T. Azuhata
Department of Advanced Physics, Hirosaki University, Hirosaki, Aomori 036-8561, Japan
e-mail: takaf@hirosaki-u.ac.jp

Semiconductors are basic materials used in a variety of electronic and photonic devices and are indispensable in today's society. Presently, a new world of semiconductor nanomaterials is being extensively explored. Semiconducting materials also play an important role in sustainable energy supply, which is one of the great global challenges for the 21st century to reduce greenhouse gases such as CO₂, and to provide clean energy especially after the Fukushima nuclear accident. By directly converting sunlight to electricity, photovoltaic devices such as solar cells can provide a versatile and renewable energy source.

The most effective utilization of these semiconductor materials needs knowledge of their structural properties which connect with electrical, optical, and magnetic properties and so can enhance device performance. XAFS has provided unique and valuable structural and electronic information in a large number of semiconductor systems.

GaN is one of the direct wide band gap semiconductors and has potential for high-power devices and radio-frequency applications. Due to the direct band gap it is also attractive for high performance light emitting devices, such as light emitting diodes (LEDs) and laser diodes (LDs). The first high efficient blue LED was realized using GaN and its mixtures with AlN and InN. The advantage of shorter wavelength in optical storage devices is to provide higher storage density. And the combination of the devices that emit red, green, and blue light is utilized in panel displays and white-light emission. In 2014, the Nobel Prize in physics was awarded in this field. On the other hand, ZnO-based semiconductors are also of current research interest due to the direct wide gap and piezo- and pyroelectric properties as materials competitive to GaN-based ones.

Such devices consist of single-quantum-well (SQW) or multi-quantum-well (MQW) structures to control electronic and, consequently, optical properties. The electronic structures in SQWs and MQWs are different substantially from those of bulk materials, because electrons are confined in one space direction. In quantum wells (QWs), the electron-hole recombination rate is improved and the luminescence efficiency is enhanced. Naturally, the structure and atomic distribution in QWs is different from bulk, so it is important to study them by XAFS. The latter can be considered a hybrid technique because the information for both local structure and electronic structure is included, which makes XAFS a unique analytical technique to study semiconducting thin films.

6.2 XAFS Spectroscopy: Basic Theory and Measurements

6.2.1 Basic Theory of XAFS

XAFS is the fine structure on X-ray absorption coefficient spectra around an absorption edge of a particular element. When monochromatic X-rays are irradiated on a sample, the intensity of the transmitted X-ray, I_t , depends on the photon energy

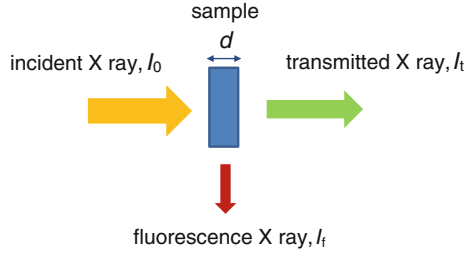


Fig. 6.1 Schematic representation of transmission and fluorescence mode

E and sample thickness d (Fig. 6.1). According to the Beer-Lambert law, I_t is expressed as

$$I_t = I_0 e^{-\mu(E)d}. \quad (6.1)$$

Here, I_0 is the intensity of the incident X-ray and $\mu(E)$ is the energy-dependent X-ray absorption coefficient. Since $\mu(E) \sim a Z^4/m E^3$ where a , Z , and m are respectively the density of the sample, atomic number, and electron mass, the magnitude of $\mu(E)$ monotonically decreases as the photon energy is increased. However, if the photon energy reaches the binding energy of core electrons, the energy of a photon is consumed to excite a core electron and the photon is annihilated to create a photoelectron and a core hole. As the result, a new absorption channel is opened and $\mu(E)$ shows steep increase around the absorption edge. The core hole recombines with an electron at the higher energy state with X-ray fluorescence or the creation of an Auger electron.

According to the quantum mechanical perturbation theory, transition rate is proportional to the square of the corresponding matrix element M [6]. Thus, $\mu(E)$ is

$$\mu \propto |M|^2 = |\langle f | H_I | i \rangle|^2 = |\langle f_0 + \delta f | H_I | i \rangle|^2 \cong |\langle f_0 | H_I | i \rangle|^2 + 2\text{Re}\langle f_0 | H_I | i \rangle^* \langle \delta f | H_I | i \rangle, \quad (6.2)$$

where $|i\rangle$ and $|f\rangle$ are the initial and final states, respectively, and H_I is the interaction Hamiltonian between the X-ray photon and the electron. In the case of K -edge absorption, $|i\rangle$ is the $1s$ electron state. $|f_0\rangle$ is the final state for a hypothetical isolated atom. $|f_0 + \delta f\rangle$ is the final state with interference between outgoing and incoming electron waves after the electron is backscattered by surrounding atoms, which contains the essential information on XAFS. In this equation, the second order term of $|\delta f\rangle$ is neglected. The first term $|\langle f_0 | H_I | i \rangle|^2$ on the right side corresponds to μ_0 and XAFS function $\chi(k)$ is

$$\chi(k) = (\mu - \mu_0)/\mu_0 \propto \langle \delta f | H_I | i \rangle, \quad (6.3)$$

where k is the photoelectron wavenumber converted from E using $k = \sqrt{2m(E - E_0)/\hbar^2}$.

For higher energy region of about 60–1000 eV from the absorption edge, the expression can be simplified by some approximations: Muffin-tin approximation for the atomic potential, one-electron approximation in the absorption process, dipole approximation for the photon-electron interaction, plane-wave approximation for the photoelectron, and single scattering approximation. The fine structure in this energy region is called Extended XAFS (EXAFS) and the expression can be simplified as

$$\chi(k) = \sum_j \frac{N_j^* S_0^2(k)}{kr_j^2} |f_j(k, \pi)| e^{-2(\sigma_j^2 k^2 + r_j/\lambda(k))} \sin(2kr_j + 2\delta(k) + \phi_j(k)), \quad (6.4)$$

where r_j , N_j^* , and σ_j are the interatomic distance, the effective coordination number, and the root-mean-square relative displacement (or Debye-Waller factor), respectively, for atomic pair j . These parameters can be determined by the non-linear least square fitting (so-called curve fitting) method by, e.g., XANADU [7] and Athena and Artemis code [8]. The phase shift $\phi_j(k)$, the backscattering amplitude $f_j(k)$, and the electron mean free path $\lambda(k)$ are generally obtained from theoretical calculation, for example, FEFF code [9].

In principle, the fitting of the theoretical curve to experimental data can be performed over the entire EXAFS spectrum. More frequently, however, the fit is carried out only on a filtered portion of the data which corresponds to the first several coordination shells. Comparison of the theory with experimental data can be performed in the r -space after Fourier transform or in the k -space after back-transform. The maximum number of independent parameters that can be reasonably obtained from one EXAFS spectrum is proposed as the following equation [3]:

$$N_{\text{ind}} \sim 2\Delta k \Delta r / \pi, \quad (6.5)$$

where $\Delta k (=k_{\text{max}} - k_{\text{min}})$ and $\Delta r (=r_{\text{max}} - r_{\text{min}})$ are intervals in each space of the spectrum.

For crystalline and oriented samples, the effective coordination number N_j^* is expressed as

$$N_j^* = 3 \sum_i \cos^2 \theta_i, \quad (6.6)$$

where θ_i is the angle between the direction of i th photoelectron scattering atom from the X-ray absorbing atom and the polarization direction of the incident X-ray. The Debye-Waller factor can be separated into a thermal vibration term and statistical disorder term as

$$\sigma^2 = \sigma_T^2 + \sigma_S^2. \quad (6.7)$$

The thermal vibration process ($\sim 10^{-13}$ s) is much slower than the X-ray absorption process ($\sim 10^{-16}$ s) and the effect reduces the amplitude of the XAFS signal. This process is well understood by the Einstein model within a harmonic approximation and σ_T^2 can be expressed as

$$\sigma_T^2 = \frac{\hbar}{2\mu\omega_E} \coth\left(\frac{\hbar\omega_E}{2k_B T}\right), \quad (6.8)$$

where ω_E is Einstein frequency and μ is reduced mass of the atomic pair, respectively. This equation is quite useful for the harmonic or weak anharmonic systems. For strong anharmonic systems, the cumulant expansion is valuable. The temperature dependent factor in the EXAFS function with relative displacement of Δ_a can be expressed as the following cumulant expansion [10]:

$$\langle e^{2ik\Delta_z} \rangle = \exp \left[\sum_{n=1}^{\infty} \frac{(2ik)^n}{n!} \langle \Delta_z^n \rangle_c \right]. \quad (6.9)$$

The odd terms appear in the phase part and the even terms appear in the amplitude part of the EXAFS function. The higher order terms include the anharmonic contribution in the atomic potential.

On the other hand, the absorption around the absorption edge is mainly attributed to the transition of core electrons to unoccupied bound states. This feature of the absorption coefficient spectrum is referred to as X-ray absorption near edge structure (XANES) or near edge XAFS (NEXAFS). XANES reflects the electronic structure of the X-ray absorbing atoms and the local structure around them. Theoretical calculations in this region are quite complicated and interpretation of the spectrum is still an active research topic.

The intensity of pre-edge peak depends on the symmetry character of the X-ray absorbing atom. For example, in the *K*-edge of Ti, V, Cr, Mn, a $1s \rightarrow 3d$ peak is dipole-forbidden for O_h symmetry but dipole-allowed for T_d or D_{4h} symmetry. One can easily check the symmetry by observing the intensity of the pre-edge peak. The energy position of the threshold of each edge in the absorption spectrum is quite sensitive to the valence state of the X-ray absorbing atom and approximately linearly depends on it. This characteristic is very useful to the mixed valence state and one can follow the change of the valence in battery materials being charged or discharged by in situ XAFS measurements.

6.2.2 Measurements of XAFS

Photon source: Most X-ray absorption spectra are measured using synchrotron radiation sources because an X-ray beam with high intensity and a continuous energy spectrum is required. Electrons are forced to go around a storage ring with the aid of a series of bending magnets and emit polarized pulsed X-rays of high intensity and a continuous energy spectrum over a wide energy range. In addition to bending magnets, insertion devices such as a wiggler or undulator are also installed in modern synchrotron facilities. To collimate and focus the X-ray beam mirrors are used, followed by slits to adjust the beam size. To select the required X-ray energy, a double-crystal monochromator is utilized based on the Bragg diffraction condition

$$n\lambda = 2d \sin\theta, \quad (6.10)$$

where n , λ , d , and θ are an integer, the X-ray wavelength, the lattice spacing of the diffracting crystal, and the angle between the beam and the diffracting crystal, respectively. X-rays of higher harmonics ($n \geq 2$) included in the incident beam can be removed by another mirror. The X-ray absorption coefficient can be determined from the relative intensities of the incident and transmitted X-ray beams, X-ray fluorescence, or Auger electrons (corresponding to transmission mode), fluorescence mode, and electron yield mode, respectively.

Transmission mode: Intensity of incident and transmitted X-ray beams is measured using ionization chambers filled with inert gas and the absorption coefficient μ can be obtained from (6.1). For transmission measurements, appropriately concentrated samples are needed. Furthermore, the samples are required to be highly homogeneous and to have constant thickness without pinholes. Thus, if necessary, samples should be prepared by crushing an appropriate amount of materials and mixing it with a suitable binder such as BN, polyethylene, or cellulose which are transparent to X-rays.

Fluorescence mode: Intensity of the characteristic fluorescent X-ray is usually measured using a simple ionization chamber (Lytle detector) or energy-dispersive detectors such as a silicon drift detector and a germanium solid state detector. The intensity of X-ray fluorescence (I_f in Fig. 6.1) is represented as

$$I_f = \frac{I_0 \mu_i(E) \varepsilon_f (\Omega/4\pi)}{\mu_T(E) + \mu_T(E_f)} \times [1 - \exp[-\{\mu_T(E) + \mu_T(E_f)\}d]] \quad (6.11)$$

where $\mu_T(E)$ is the total absorption coefficient of the sample for incident X-ray, $\mu_T(E_f)$ is the absorption coefficient for the fluorescent X-ray, $\mu_i(E)$ is the absorption coefficient for the element i of interest (essence of XAFS signal), Ω is the solid angle of the detector, ε_f is fluorescence efficiency, and d is the sample thickness. It should be noted that the self-absorption effect must be taken into account for thick or concentrated samples. The intensity of the fluorescence line normalized by that of incident X-rays (I_f/I_0) is proportional to $\mu_i(E)$ under the

conditions: dilute sample as *ca.* 0.01 mol dm^{-1} or thin layer under *ca.* 100 nm. For multi-element samples, the characteristic X-ray must be separated from other X-ray such as lines due to elastic scattering and Compton scattering and other fluorescence lines. Si and Ge detectors are capable of extracting the intensity of the fluorescence X-ray of interest. To improve *S/N* ratio of the signal, a multi-element detector is necessary. Fluorescence mode has a great advantage for highly dilute samples such as thin films deposited on thick substrates.

Electron yield mode: In this mode, the amount of electrons emitted from the sample surface by X-ray irradiation, such as photoelectrons, secondary electrons, and Auger electrons, is measured. Relatively short mean free path of the electrons makes this technique surface sensitive (5–10 nm from the surface) and free from the self-absorption effect. This mode is useful for XAFS in the low energy (soft X-ray) region.

Another unique method is **XEOL** (X-ray Excited Optical Luminescence), in which the intensity of photoluminescence emitted from the sample provides the XAFS function. This method is capable of distinguishing the electronic structure or the local structure around particular chemical species from other species by the difference in their photon energies (see Chap. 11 in [2]).

6.3 Application to Semiconducting Thin Films

6.3.1 *c*-Plane InGaN SQWs

$\text{In}_x\text{Ga}_{1-x}\text{N}$ is a key material in high-brightness blue and green LEDs and purplish-blue LDs [11]. Although such devices have very high densities of threading dislocations, they show high quantum efficiency in contrast to conventional III–V and II–VI semiconductor-based devices [12]. The fluctuation of In atom concentration in InGaN active layers has been proposed as its origin [13]. In this model, injected carriers are localized into fluctuation-induced potential minima and can efficiently contribute to radiative recombination before capture by dislocation-activated non-radiative recombination centers, which is considered to lead to the high quantum efficiency. Therefore, in order to make the emission mechanism clear in InGaN-based light-emitting devices, it is important to clarify the local structure around In atoms in active layers. In this section, we present the results of EXAFS measurements of In *K* edge for high quality *c*-plane (0001) $\text{In}_x\text{Ga}_{1-x}\text{N}$ ($x = 0.145, 0.20, 0.275$) SQW LED wafers by fluorescence detection using a linearly polarized intense synchrotron X-ray beam and a high performance detector. Polarization direction of X-ray electric field was set to be both horizontal and vertical to SQWs. The anisotropy of the local structure around In atoms is discussed [14].

The samples used in this study are $\text{In}_x\text{Ga}_{1-x}\text{N}$ ($x = 0.145, 0.20, 0.275$) SQWs of 3 nm thickness, which were grown by metal organic chemical vapor deposition

(MOCVD) on sapphire (0001) substrates. Each sample consists of a GaN buffer layer, an n -GaN:Si barrier layer (4000 nm), an $\text{In}_x\text{Ga}_{1-x}\text{N}$ SQW (3 nm), a p - $\text{Al}_{0.2}\text{Ga}_{0.8}\text{N}:\text{Mg}$ barrier layer (100 nm), and a p -GaN layer (500 nm). The values of x were determined using the relation between photoluminescence (PL) emission energy and x for 3 nm thick $\text{In}_x\text{Ga}_{1-x}\text{N}$ QWs, the uncertainty of which is estimated to be ± 0.025 [15].

Figure 6.2 shows In K -edge X-ray absorption spectra of the $\text{In}_{0.20}\text{Ga}_{0.80}\text{N}$ SQW for (a) horizontal (electric field of the X-ray, E , is perpendicular to the c -axis of the sample) and (b) vertical direction. The spectra were collected on BL10XU, BL01B1, and BL07B1 at SPring-8, Japan.

The background was subtracted and EXAFS oscillating function $\chi(k)$ was extracted as shown in Fig. 6.3a. In general, $\chi(k)$ is multiplied by k or k^2 , k in this case. X-ray photon energy was converted into wavenumber of photoelectron as discussed previously. Figure 6.3b shows radial distribution functions obtained by Fourier transform of the $k\chi(k)$ functions. The peaks around 1.6 Å and 2.8 Å correspond to In–N and In–In/In–Ga atomic pairs, respectively. After the EXAFS

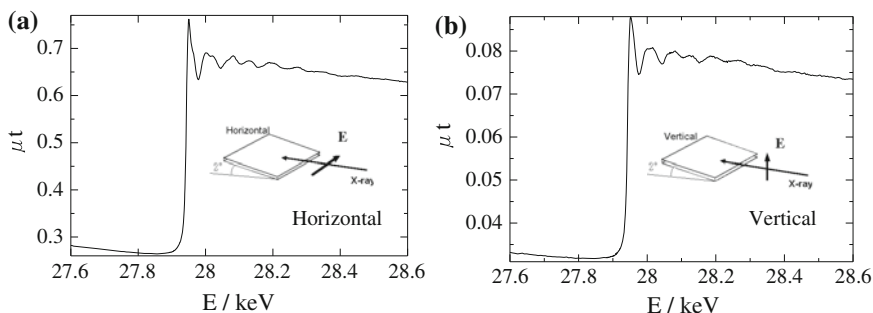


Fig. 6.2 In K -edge X-ray absorption spectra of the $\text{In}_{0.20}\text{Ga}_{0.80}\text{N}$ SQW for **a** horizontal (E is perpendicular to the c -axis) and **b** vertical direction [14]

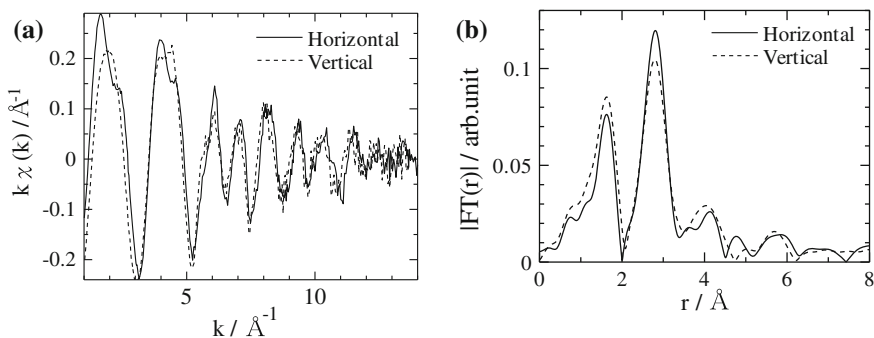


Fig. 6.3 **a** EXAFS $k\chi(k)$ functions and **b** their Fourier transforms for the $\text{In}_{0.20}\text{Ga}_{0.80}\text{N}$ SQW [14]

functions are fitted to the theoretical curves as (6.4), the structure parameters are obtained.

As the results, it was found that both $r_{\text{In-Ga}}$ and $r_{\text{In-In}}$ for six out-of-plane atoms are larger than those for six in-plane atoms, suggesting that the InGaN SQWs are biaxially compressed in the ab plane by the effect of adjacent layers. Because $r_{\text{In-N}}$ is almost constant for both directions, the angle of In–N–Ga/In in the horizontal direction is found to be smaller than that in the vertical direction. It should be noted that the values of $r_{\text{In-In}}$ in our particular samples are almost independent of x and close to that of metallic indium 3.2–3.3 Å which is predicted in zinc-blende InGaN and is also expected in wurtzite InGaN to cause strong coupling of valence states by Kent and Zunger [16].

The local InN molar fraction around In atoms, y , can be expressed using effective coordination numbers as

$$y = N_{\text{In-In}}^* / (N_{\text{In-In}}^* + N_{\text{In-Ga}}^*). \quad (6.12)$$

This is essentially different from the average InN molar fraction x obtained from PL emission energy [15], since EXAFS is sensitive only to the local structure around specific atoms, In in this case. It should be noted that the value of the denominator is 9.0 (12.0) for in-plane (out-of-plane) molar fraction.

In Fig. 6.4, the values of y/x are plotted as a function of x for the three $\text{In}_x\text{Ga}_{1-x}\text{N}$ SQW samples. The In local structure can be classified into three categories: (1) If $y = x$, In atoms are randomly distributed; (2) if $y \leq x$, In atoms are isolated and superlattice structures exist; (3) if $y > x$, In atoms are aggregated with each other and separated from Ga atoms. As can be seen in Fig. 6.4, it is found that the In atom distribution is statistically random in the horizontal direction in all the samples. Note that the accidental aggregation of In atoms is possible even in statistically random alloys which are not necessarily homogeneous. On the other hand, In atoms in the vertical direction for $x = 0.20$ are found to be aggregated with each other and are located at the top and bottom, that is, In atoms tend to be just above and/or just

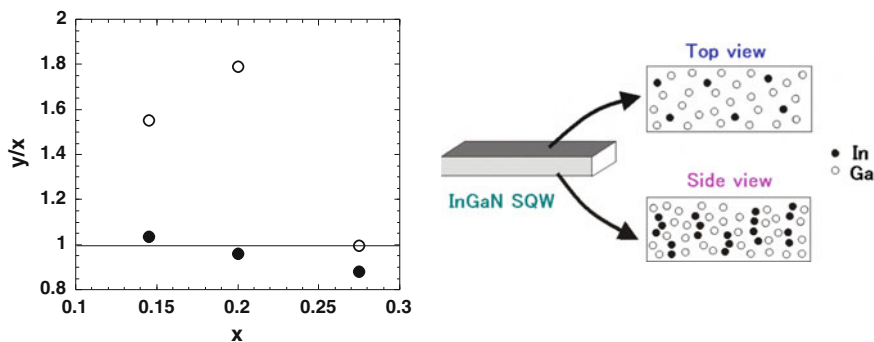


Fig. 6.4 The plot of y/x with x for the three InGaN SQWs. The schematic model for In atom distribution is also shown (see text)

below In atoms. Although higher experimental accuracy is necessary to make a decision, such vertical aggregation is likely to exist also in the $\text{In}_{0.145}\text{Ga}_{0.855}\text{N}$ SQW because the plot is above the solid line corresponding to $y = x$. For $x = 0.275$, y in the vertical direction is almost equal to x , indicating random distribution. The average InN molar fractions $x = 0.145$, 0.20 , and 0.275 correspond to blue, green, and amber LEDs, respectively. According to the emission energy dependence of the external quantum efficiency for InGaN LEDs by Mukai et al. [17], InGaN-based blue to green LEDs show higher external quantum efficiency. Thus, the present EXAFS results suggest that the aggregation of In atoms in the vertical direction in active layers correlates with the higher quantum efficiency of blue to green InGaN LEDs.

6.3.2 *c*-Plane InGaN MQW

The local structure around In atoms in the $\text{In}_x\text{Ga}_{1-x}\text{N}$ ($x = 0.145, 0.20, 0.275$) SQW LED wafers were discussed in the previous section: In atoms are aggregated and located at the top and bottom in the vertical direction of the SQW planes. The In–In in-plane distance is smaller than the out-of-plane distance. This means that the SQWs of InGaN are biaxially compressed in the *ab* plane by the effect of adjacent layers.

On the other hand, MQWs are used in LDs to improve device performance such as threshold current density and optical gain. In this section, we discuss the local structure around In atoms in a *c*-plane $\text{In}_x\text{Ga}_{1-x}\text{N}$ MQW, where the sample surface is set to be horizontal and vertical to the X-ray polarization electric vector [18]. The results are discussed and compared with those for the previous $\text{In}_{0.2}\text{Ga}_{0.8}\text{N}$ SQW study [14].

The sample was an MQW structure of 10 periods of $\text{In}_{0.2}\text{Ga}_{0.8}\text{N}$ (2.5 nm) and $\text{In}_{0.05}\text{Ga}_{0.95}\text{N}$ (7.5 nm), grown by MOCVD on a sapphire (0001) substrate [19]. X-ray absorption measurements were carried out at BL10XU of SPring-8.

Figure 6.5 shows (a) the In *K*-edge EXAFS $k\chi(k)$ functions and (b) Fourier transforms of the $\text{In}_x\text{Ga}_{1-x}\text{N}$ MQW for horizontal (solid line) and vertical (dashed line) polarization directions. The spectral difference between these two directions is relatively small. The EXAFS amplitude for the MQW is larger than that for the SQW. The range for the Fourier transform is $2.7\text{--}13.3 \text{ \AA}^{-1}$. The 1st peak at $1.5 \text{ \AA}$ in Fig. 6.5b corresponds to the In–N atomic pair, while the 2nd peak at $2.8 \text{ \AA}$ corresponds to the In–In and In–Ga atomic pairs. The Debye-Waller factors for the In–In (around $0.04 \text{ \AA}$) and In–Ga (around $0.06 \text{ \AA}$) atomic pairs in the MQW are smaller than those in the SQW ($0.08 \text{ \AA}$ and $0.07 \text{ \AA}$, respectively) [14]. In particular, the difference for the In–In pairs is prominent. This result corresponds to the difference in the intensities of the 2nd peaks in Fig. 6.5b.

In the case of the SQW, the In–In proportion (or coordination number, N) is higher for the vertical direction ($N = 4.0$) than for horizontal direction ($N = 2.1$). This indicates that the In atom is aggregated in the vertical direction [2]. On the

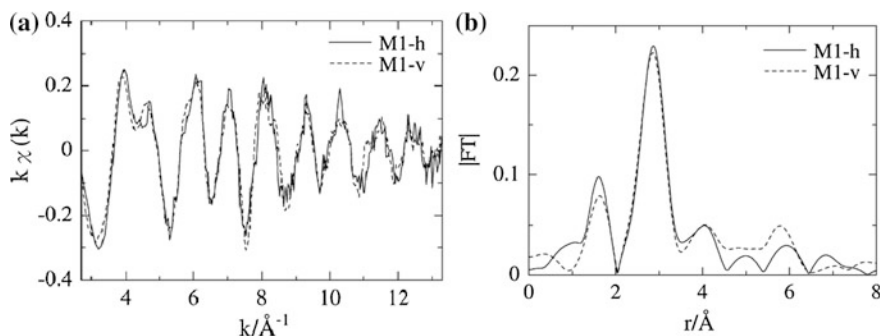


Fig. 6.5 **a** In K -edge $k\chi(k)$ and **b** Fourier transform of the $\text{In}_x\text{Ga}_{1-x}\text{N}$ MQW ($x_{\text{av}} = 0.088$) for horizontal (solid line) and vertical (dashed line) directions [18]

other hand, in the present InGaN MQW, N for In–In is the same for the horizontal and vertical direction ($N = 1.2$). This result indicates that the nearest-neighbor In atom is isotropically distributed around the absorber In atoms in the InGaN MQW. The average In concentration is 0.088 in this InGaN MQW. Therefore, N is expected to be 1.1 if In atoms are randomly distributed and the value of N ($=1.2$) is quite close to this value. The interatomic distances (In–In 3.29 Å and In–Ga 3.27 Å) in the vertical direction is longer than that in the horizontal direction (In–In 3.23 Å and In–Ga 3.21 Å) for the SQW, where the InGaN layer is biaxially compressed in the ab plane by the effect of the adjacent layers. On the other hand, those in both directions are the same for the InGaN MQW (In–In 3.25 Å and In–Ga 3.24 Å). Judging from these results, it is concluded that the strain in the $\text{In}_x\text{Ga}_{1-x}\text{N}$ MQW is reduced compared to that in the SQW [14].

6.3.3 m -Plane InGaN Thin Film

As discussed in the previous section, in the c -plane $\text{In}_x\text{Ga}_{1-x}\text{N}$ QWs, the localized exciton emission exhibits high internal quantum efficiency despite the presence of high-density threading dislocations. The internal quantum efficiency, however, decreases abruptly for the emission wavelength longer than around 520 nm (InN molar fraction $x < 0.2$) due to the quantum-confined Stark effects (QCSEs). The effects are caused by the electrostatic field perpendicular to QWs due to the spontaneous and piezoelectric polarization [13, 15, 20]. In contrast to polar c -plane QWs, nonpolar m -plane ($1\bar{1}00$) and a -plane ($11\bar{2}0$) QWs can avoid the QCSEs because the polar c -axis is on the QW plane. Thus, nonpolar InGaN epilayers are attracting much attention to achieve higher internal quantum efficiency. The structural study for nonpolar InGaN films is important to reveal the emission mechanism. However, there is little research in this area presently. Recently, Chichibu et al. reported m -plane $\text{In}_x\text{Ga}_{1-x}\text{N}$ ($x < 0.14$) thin films grown on

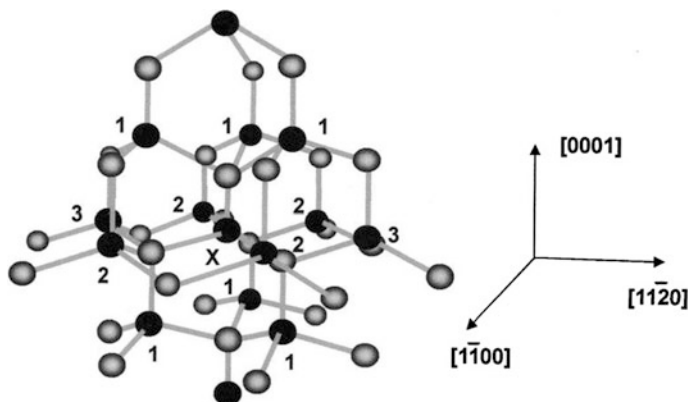


Fig. 6.6 Atomic structure and crystal axes for the model of the *m*-plane $\text{In}_{0.06}\text{Ga}_{0.94}\text{N}$. Black circles represent Ga or In atoms and gray ones N atoms. The atom labelled “X” is the X-ray absorbing In atom and the numbers of “1”, “2”, and “3” show three types of atomic positions

freestanding *m*-plane GaN substrates. The films were confirmed to be coherently grown by an X-ray reciprocal space mapping (XRSM) method [21], but the local structures around In atoms are still unknown. In this section, we present the results of In *K*-edge EXAFS analyses for one of the *m*-plane InGaN films, an $\text{In}_{0.06}\text{Ga}_{0.94}\text{N}$ film. The following three configurations were used: the electric field vector of the incident X-ray is parallel to the *c*-axis [0001], the *m*-axis $[1\bar{1}00]$, and the *a*-axis $[11\bar{2}0]$. A step-by-step procedure is introduced here to analyze the EXAFS data for the three different directions individually [22].

The $\text{In}_{0.06}\text{Ga}_{0.94}\text{N}$ film was deposited by MOVPE on a 1.5- μm -thick GaN buffer layer on a 325- μm -thick GaN substrate prepared by a halide vapor phase epitaxy method [23]. X-ray absorption measurements were made at a beam line of NW10A of Photon Factory Advanced Ring, KEK. XAFS data were collected with a double-crystal monochromator of Si (311) crystals. Indium K_{α} -fluorescence emission was measured using a 19-element Ge solid-state detector.

Figure 6.6 shows the atomic structure and the crystal axes of the wurtzite structure. The numbers of “1”, “2”, and “3” show three types of atomic positions.

For the analyses of the local structure for the second nearest neighbor (2NN) In–Ga and In–In, the following step-by-step procedure was used. Firstly, the $E \parallel c$ data were analyzed (E is the electric vector of the X-ray). In this configuration, only the type “1” atoms are detected by XAFS. Their effective coordination number N_1^* was set to be 12. The type “2” and “3” atoms are not detected because the plane including these types of atoms and the X-ray absorbing In atom (labelled “X” in Fig. 6.6) is perpendicular to E . Thus, $N_2^* = N_3^* = 0$. Strictly speaking, these six atoms labelled “1” are not equivalent for the strained *m*-plane wurtzite structure. However, it is quite difficult to analyze them separately because of too many fitting parameters. Thus, XAFS spectra were analyzed by assuming that the type “1” six Ga/In are equivalent.

Secondly, the $E \parallel m$ data, to which only the type “1” and “2” atoms contribute, were analyzed. Although σ may depend on the polarization direction of the incident X-ray, it was assumed to be isotropic in this study. For r , σ , and the ratio of $N_{\text{In-Ga}}^*/N_{\text{In-In}}^*$ for type “1” atoms, the values determined from the above $E \parallel c$ data fitting were used. N_1^* and N_2^* were fixed to be 3 and 9, respectively. At this step, r , σ , and $N_{\text{In-Ga}}^*/N_{\text{In-In}}^*$ for type “2” atoms were determined.

Finally, the $E \parallel a$ data were analyzed. In this configuration, all types of atoms contribute to the XAFS signal. The values of r , σ , and $N_{\text{In-Ga}}^*/N_{\text{In-In}}^*$ for type “1” and “2” atoms, which were obtained from the above two steps, were used. N_1^* and N_2^* , and N_3^* were set to be 3, 3, and 6, respectively.

We describe only preliminary results here. From the coordination numbers, In atoms were shown to be localized in all directions. The In–Ga distance in the c -direction was found to be longer than that in strain-free virtual crystals in spite of the compressive strain along the a - and c -axes, which can be explained by the partial strain-cancellation mechanism due to the anisotropic atomic structure on the m -plane. It was also found that the local strain in coherently grown m -plane InGaN is more relaxed than that in coherently grown c -plane InGaN.

6.3.4 *m*-Plane AlGaN Thin Films

Group III-nitride semiconductors have attracted much attention as materials for blue and ultraviolet light emitters [24]. The most studied member of this group, wurtzite GaN, has a direct band gap of 3.42 eV, while pseudo-binary alloys, $\text{Al}_x\text{Ga}_{1-x}\text{N}$, provide a direct gap up to 6.0 eV. It is attractive that it covers the wide range of ultraviolet rays. There are several XAFS studies of Ga K -edge for $\text{Al}_x\text{Ga}_{1-x}\text{N}$ ($x = 0\text{--}0.4$) [25] and ($x = 0\text{--}0.6$) [26] films, where the interatomic distances and the coordination numbers of Ga–Al and Ga–Ga pairs are discussed. On the other hand, quite a few Al K -edge XAFS studies have been reported for AlGaN films. Recently, it was pointed out that the problem is the reduction in oscillator strength of electron-hole pairs in quantum wells due to the QCSEs caused by the spontaneous and piezoelectric polarization discontinuity at c -plane heterointerfaces. To avoid these effects in AlGaN heterostructures, epitaxial growth in nonpolar orientation such as m - or a -plane has been attracting attention as shown in the previous section. However, the local structure and the basic physical properties of nonpolar AlGaN films are not fully understood.

In this section, the polarized XAFS studies for Al K -edge of m -plane $\text{Al}_x\text{Ga}_{1-x}\text{N}$ thin films are presented. In particular, the interatomic distance of 2NN Al–Al and Al–Ga atomic pairs and the localization of Al atoms in m -plane $\text{Al}_x\text{Ga}_{1-x}\text{N}$ are discussed [27].

140-nm-thick m -plane $\text{Al}_x\text{Ga}_{1-x}\text{N}$ films ($x = 0.32$ and 0.58) were grown by the NH_3 source molecular beam epitaxy method on m -plane GaN substrates. It is reported that the $\text{Al}_{0.32}\text{Ga}_{0.68}\text{N}$ film is pseudomorphic but the $\text{Al}_{0.58}\text{Ga}_{0.42}\text{N}$ film is

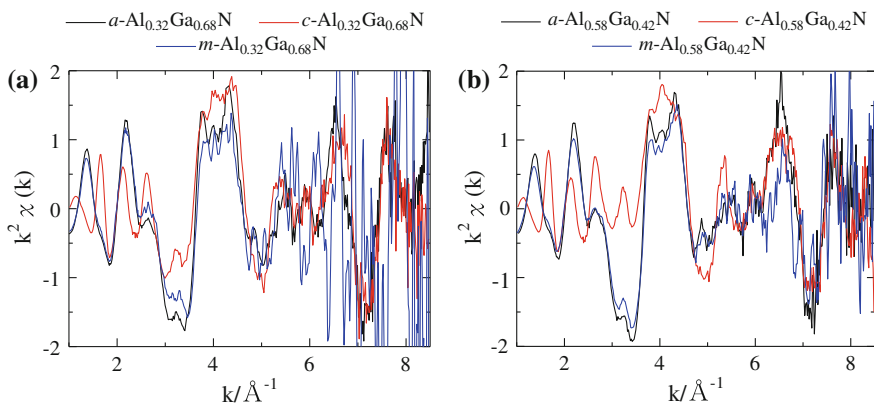


Fig. 6.7 EXAFS $k^2\chi(k)$ functions for Al K -edge of **a** the m -plane $\text{Al}_{0.32}\text{Ga}_{0.68}\text{N}$ film and **b** the $\text{Al}_{0.58}\text{Ga}_{0.42}\text{N}$ film [27]

partially or nearly fully relaxed [28]. Thus, it is important to study the local structure around Al atoms in the films by EXAFS.

Al K -edge (1560 eV) XAFS was measured on BL11A at Photon Factory, KEK, in fluorescence mode using a silicon drift detector. To measure the polarization dependent XAFS, we set the sample in three configurations with the electric field vector of X-ray parallel to the (1) a -axis $[11\bar{2}0]$, (2) c -axis $[0001]$, and (3) m -axis $[1\bar{1}00]$.

Figure 6.7 shows EXAFS $k^2\chi(k)$ spectra of Al K -edge for the m -plane (a) $\text{Al}_{0.32}\text{Ga}_{0.68}\text{N}$ and (b) $\text{Al}_{0.58}\text{Ga}_{0.42}\text{N}$ films, where “ a –”, “ c –”, and “ m –” denote that the polarization direction of the incident X-ray is parallel to the a -, c -, and m -axes, respectively. The data quality is acceptable.

Figures 6.8a, b show Fourier transforms for Al K -edge of the m -plane $\text{Al}_{0.32}\text{Ga}_{0.68}\text{N}$ and $\text{Al}_{0.58}\text{Ga}_{0.42}\text{N}$ films, respectively. The k -range for the Fourier transforms is $2.1\text{--}7.75\text{ \AA}^{-1}$. The authors analyzed polarization dependent XAFS data individually.

Although in general it is not so easy to analyze Al K -edge EXAFS because of the low energy of the K -edge and the narrow energy range of EXAFS region, the EXAFS curve fitting analysis was performed. The step-by-step procedure has been proposed for more precise polarization analysis in the previous section, however only the simple two-shell fitting are applied for Al–Al and Al–Ga pairs in each direction for this systems. The structural parameters obtained from the curve-fitting analysis are listed in Table 6.1.

For both the films, the interatomic distances of Al–Al and Al–Ga for the a -axis-polarized spectra are about $3.17 \pm 0.02\text{ \AA}$, which is close to Ga–Ga distance in GaN. On the other hand, those for the m -axis spectra are about $3.10 \pm 0.02\text{ \AA}$, which is close to Al–Al distance in AlN. It is suggested that local structures along the a -axis of the films are strongly affected by GaN substrates. This result is similar to that for m -plane $\text{In}_{0.06}\text{Ga}_{0.94}\text{N}$ XAFS, but in that case the 2NN interatomic

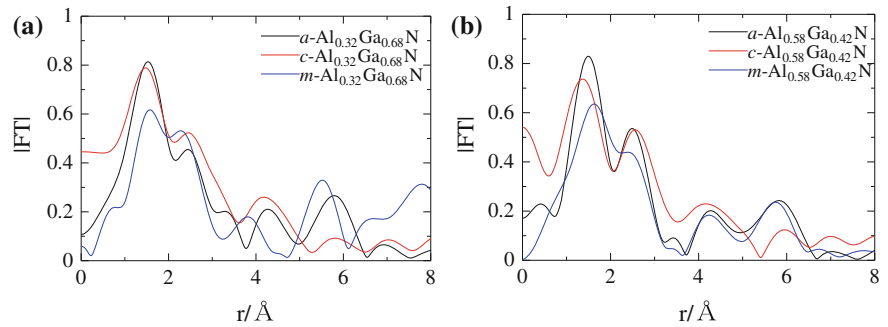


Fig. 6.8 Fourier transforms for Al *K*-edge of **a** the *m*-plane Al_{0.32}Ga_{0.68}N film and **b** the Al_{0.58}Ga_{0.42}N film [27]

Table 6.1 Structural parameters for the Al_xGa_{1-x}N films obtained from EXAFS curve fitting [27]

	$r_{\text{Al-Al}}/\text{\AA}$	$r_{\text{Al-Ga}}/\text{\AA}$	$N_{\text{Al}}/(N_{\text{Al}} + N_{\text{Ga}})$
<i>a</i> -Al _{0.32} Ga _{0.68} N	3.17	3.17	0.63
<i>c</i> -Al _{0.32} Ga _{0.68} N	3.11	3.18	0.67
<i>m</i> -Al _{0.32} Ga _{0.68} N	3.09	3.10	0.67
<i>a</i> -Al _{0.58} Ga _{0.42} N	3.17	3.18	0.58
<i>c</i> -Al _{0.58} Ga _{0.42} N	3.10	3.11	0.67
<i>m</i> -Al _{0.58} Ga _{0.42} N	3.11	3.12	0.59

distance (In–Ga) is shortened by the effect of GaN substrates. The interatomic distances of Al–Ga for the *c*-axis spectra are 3.18 ± 0.02 Å for Al_{0.32}Ga_{0.68}N and 3.11 ± 0.02 Å for Al_{0.58}Ga_{0.42}N, respectively. This difference may be caused by the fact that the Al_{0.32}Ga_{0.68}N film is pseudomorphic but that the Al_{0.58}Ga_{0.42}N film is partially or nearly fully relaxed. Yu et al. [26] studied Ga *K*-edge XAFS for Al_xGa_{1-x}N ($x = 0\text{--}0.6$) films and reported that the interatomic distance of Ga–Al is almost constant to be 3.10 Å and that of Ga–Ga changes linearly from 3.18 Å for GaN to 3.14 Å for Al_{0.6}Ga_{0.4}N. The present polarization dependent analysis shows that the interatomic distances of Al–Al and Al–Ga for the *m*-axis spectra are close to their Ga–Al distance (3.10 Å) and those for the *a*-axis spectra are close to their Ga–Ga distance for GaN (3.18 Å). The reason of these discrepancies may be that the present samples were grown on *m*-plane GaN substrates, while their ones on *c*-plane sapphire substrates.

The proportion of the 2NN Al atoms around Al atoms in the films is evaluated as $N_{\text{Al}}/(N_{\text{Al}} + N_{\text{Ga}})$. Those for the Al_{0.32}Ga_{0.68}N and Al_{0.58}Ga_{0.42}N films are around 0.63 ± 0.05 for each direction, suggesting that Al atoms are almost randomly distributed in the Al_{0.58}Ga_{0.42}N film but are localized in the Al_{0.32}Ga_{0.68}N film within the simple model. Polarization dependence of the localization of Al atoms was not observed. Yu et al. [26] reported that Ga and Al atoms are randomly distributed in the concentration range of $x = 0\text{--}0.6$. It is suggested that Al atoms are

localized in the $\text{Al}_{0.32}\text{Ga}_{0.68}\text{N}$ film. Hazu et al. have studied spatio-time-resolved cathode luminescence for an *m*-plane $\text{Al}_{0.25}\text{Ga}_{0.75}\text{N}$ film and suggested the localization of Al atoms [29].

6.3.5 *c*-Plane MgZnO Thin Film

$\text{Mg}_x\text{Zn}_{1-x}\text{O}$ has attracted a lot of attention as a useful semiconducting material, because the band gap can be controlled by Mg composition [30, 31]. This is expected to be a new LED and/or LD material in place of InGaN for marketing reasons. It is known that, after annealing, $\text{Mg}_x\text{Zn}_{1-x}\text{O}$ films with $x > 0.18$ decompose into a rock salt and a wurtzite phase, while the sample with $x = 0.12$ remains as a single phase of the wurtzite, in which the *c*-parameter in the wurtzite phase approximately follows Vegard's law [32]. In such a mixed crystal semiconductor as $\text{Mg}_x\text{Zn}_{1-x}\text{O}$, the local structure around each atom is important for the study of the light emitting mechanism. The local structure around the low concentration Mg atom is actually more attractive, since the interatomic distance does not simply correspond to Vegard's law. In this section, the Zn *K*-edge and Mg *K*-edge XAFS results for *c*-plane $\text{Mg}_{0.06}\text{Zn}_{0.94}\text{O}$ film are discussed [33].

The sample of $\text{Mg}_{0.06}\text{Zn}_{0.94}\text{O}$ film fabricated on (11 $\bar{2}$ 0) Al_2O_3 substrate by Helicon Wave excited Plasma Sputtering Epitaxy (HWPSE) method. For this technique, the plasma chamber is separated from the sputtering chamber in order to avoid the re-sputtering and surface damage caused by the high energy sputtering particles. The thickness of the $\text{Mg}_{0.06}\text{Zn}_{0.94}\text{O}$ layer is 780 nm on the buffer layer of AlN and Al_2O_3 substrate. The detailed sample preparation is described elsewhere [30]. XAFS measurements were performed at BL7C for Zn *K*-edge and BL11A for Mg *K*-edge in Photon Factory (KEK, Tsukuba) [33]. The fluorescence XAFS spectra were measured using a Lytle detector for Zn *K*-edge and a silicon drift detector for Mg *K*-edge. For the wurtzite anisotropy of this sample, the measurements were performed by polarized XAFS: (1) the electric field of the incident X-ray is horizontal and (2) the electric field is vertical to the sample plane of the $\text{Mg}_{0.06}\text{Zn}_{0.94}\text{O}$ film.

A. Results of Zn *K*-edge analyses

First, we describe the analytical procedure of Zn *K*-edge EXAFS for the $\text{Mg}_{0.06}\text{Zn}_{0.94}\text{O}$ film, which is not too thin for the condition that the fluorescence limit is applicable. Figure 6.9 shows the schematic measurement set up for the sample direction and X-ray beam. To give accurate XAFS spectra by the fluorescent method, two conditions are generally required; (1) thick sample and dilute target element or (2) thin sample and dense target element as shown in the previous section. In this sample, the condition for the measurement of Zn *K*-edge corresponds to the latter (2). However, unfortunately, the effective thickness of the vertical measurement is about 15 μm and out of the condition. Hence the vertical

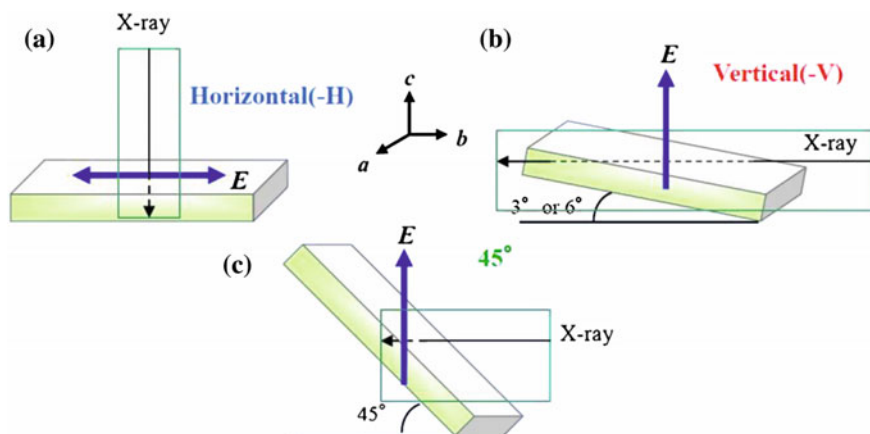


Fig. 6.9 The schematic measurement set up for the sample direction and X-ray beam. **a** The horizontal direction between the electric field of X-ray and the sample plane, **b** the vertical direction between them, and **c** 45° between them [33]

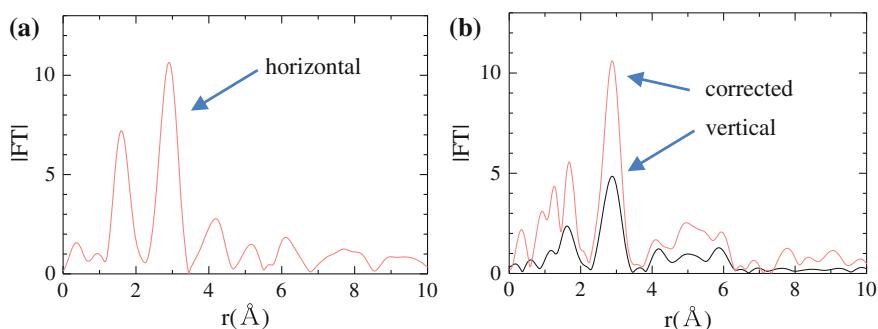


Fig. 6.10 Fourier transform of Zn *K*-edge for **a** the horizontal and **b** the vertical direction of the sample plane of $\text{Mg}_{0.06}\text{Zn}_{0.94}\text{O}$. In **b**, the corrected vertical spectrum is also shown [33]

measurement does not give a correct spectrum and the amplitude of the signal is reduced. The spectrum obtained with 45° between the sample plane and the X-ray beam (Fig. 6.9c) contains the equal amount of information of the horizontal and vertical. So the horizontal signal was subtracted from 45° signal to obtain the corrected vertical spectrum. Figure 6.10 shows the example of the Fourier transforms for the comparison with the horizontal and 45° spectra.

Next, we discuss the results of Zn *K*-edge analyses for the $\text{Mg}_{0.06}\text{Zn}_{0.94}\text{O}$ film. In Fig. 6.10b, the intensity of the Fourier transform for the corrected vertical data seems to be improved compared with the horizontal data. The first nearest peak corresponds to a Zn–O atomic pair and the second one to Zn–Zn and Zn–Mg atomic pairs. The contributions from Zn–Zn and Zn–Mg cannot be separated by the Fourier transform

method. The non-linear-least-square-fitting method was applied to these two nearest peaks. In the analyses for the second nearest peak, the contribution from Mg atoms cannot be recognized, due to the small amount of Mg atoms in $\text{Mg}_{0.06}\text{Zn}_{0.94}\text{O}$. The coordination numbers for the corrected vertical data are well improved to be close to the wurtzite structure. From these structural analyses for Zn K -edge EXAFS, the 1st nearest Zn–O distance is 1.963 Å, that of 2nd nearest Zn–Zn is 3.249 Å, and 4th nearest Zn–O is 3.861 Å. The local structure of the present $\text{Mg}_{0.06}\text{Zn}_{0.94}\text{O}$ layer is quite close to that of the bulk ZnO crystal [34].

B. Results of Mg K -edge analyses

In this subsection we discuss the result of the Mg K -edge XAFS analyses for the $\text{Mg}_{0.06}\text{Zn}_{0.94}\text{O}$ layer. Figure 6.11a shows the X-ray absorption spectra of Mg K -edge for $\text{Mg}_{0.06}\text{Zn}_{0.94}\text{O}$ for the horizontal direction (in plane) and the vertical direction (the sample is inclined about 6° to the c -axis). High quality data are obtained up to 1600 eV. Figures 6.11b, c show the EXAFS $k\chi(k)$ oscillation functions and their Fourier transforms for the horizontal (in plane) and the vertical (c -axis) direction, respectively. A clear polarization dependence is observed also in the Mg K -edge XAFS.

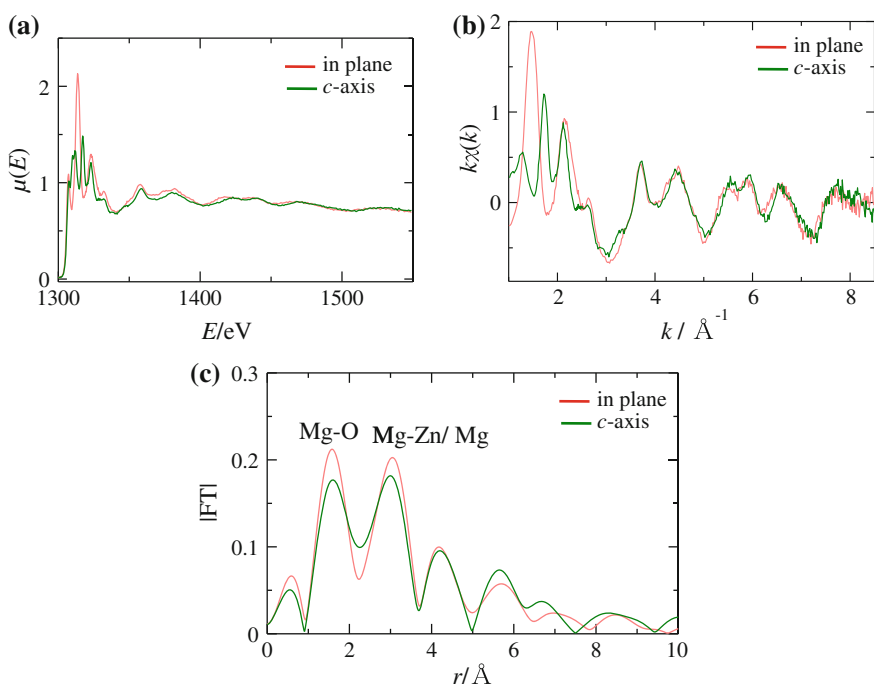


Fig. 6.11 **a** X-ray absorption spectra, **b** $k\chi(k)$ and **c** Fourier transforms of Mg K -edge for $\text{Mg}_{0.06}\text{Zn}_{0.94}\text{O}$ for the horizontal direction (in plane, red) and the vertical direction (the sample is inclined at about 6° to the c -axis, green) [33]

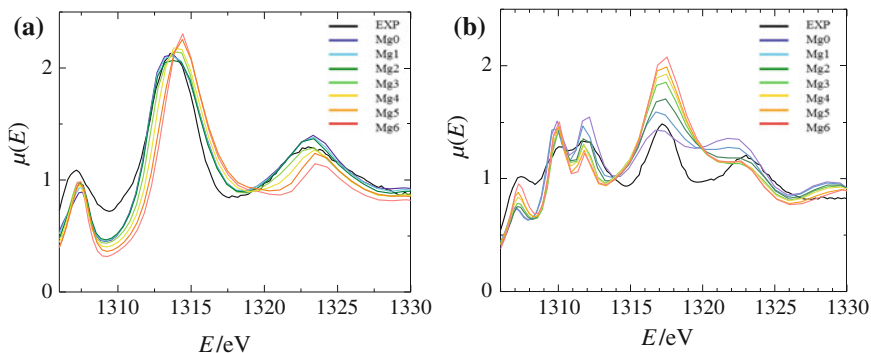


Fig. 6.12 Calculated XANES spectra with measured one (EXP) for **a** horizontal and **b** vertical directions. Mg0, Mg1, ... show the number of Mg atoms around an X-ray absorption Mg atom [33]

The data range of $k\chi(k)$ is limited to $8.2 \text{ \AA}^{-1}$ (in Fig. 6.11b), but the first and second peaks are well distinguished in the Fourier transform (in (c)). The first nearest peak corresponds to the Mg–O atomic pair and the second one does to the Mg–Zn and Mg–Mg pairs. Each contribution of Mg–Zn and Mg–Mg cannot be separated in FT as is the case of previous Zn *K*-edge EXAFS.

Although rather good quality XAFS data were obtained, the parameter fitting by non-linear-least-square method is not so easy, so we just compare the simulation spectra calculated by FEFF code with the experimental data. Figures 6.12 shows measured and calculated XANES for (a) horizontal and (b) vertical direction, respectively. For the model calculation, the atomic positions are fixed at the coordinates considered from the EXAFS result previously discussed and only the number of Mg atoms is changed as a parameter (the Zn atom in the 2NN position is replaced by Mg atom.) In Fig. 6.12a, the peaks at 1295, 1313, and 1323 eV are well reproduced by FEFF calculation. The peaks at 1313 and 1323 eV are sensitive to the number of surrounding Mg atoms; these intensities change and the positions are shifted to higher energy with increasing the number of Mg atoms. The energy shift is pronounced in the case that the number of Mg atoms is larger than 4. Similar features are observed in Fig. 6.12b: the peak at 1317 eV is sensitive to increasing Mg atoms and slightly shifts to higher energy, whereas peaks at 1312 and 1323 eV are decreasing. Judging from these simulation results, the number of surrounding Mg atoms in the second nearest shell from the X-ray absorbing Mg atom may be 0–3. Actually, for the limited non-linear-least-square-fitting in which the coordination number of Mg is fixed at the values of 0, 1, 2, ... 6, better fitting results were obtained for the model where less Mg atoms located in the 2NN positions.

It was concluded that Mg atoms are distributed randomly in the present $\text{Mg}_{0.06}\text{Zn}_{0.94}\text{O}$ layer. This result contrasts with the case of 3-nm-thick $\text{In}_x\text{Ga}_{1-x}\text{N}$ ($x = 0.145, 0.20, 0.275$) SQWs as previously discussed [4]. In the case of InGaN , the polarization dependence is remarkable and In atoms are aggregated in the vertical direction of $\text{In}_{0.2}\text{Ga}_{0.8}\text{N}$ but randomly distributed in the horizontal

direction. Recently, Shibata et al. studied the photoluminescence characterization of $\text{Mg}_x\text{Zn}_{1-x}\text{O}$ ($x = 0.05, 0.11, 0.15$) epitaxial films [31] and presented that the degree of the localization or the compositional fluctuation of Mg atoms increases with increasing the concentration of Mg.

According to the structural parameters obtained from the non-linear-least-square-fitting under the limited condition that no Mg atom exists in the 2NN from X-ray absorbing Mg atoms (not shown here), for the 1NN distance of Mg–O, Mg–O (horizontal) is 1.957 Å and Mg–O (vertical) is 1.983 Å. Thus, the local structure in the $\text{Mg}_{0.06}\text{Zn}_{0.94}\text{O}$ layer is almost the same as the bulk ZnO structure. For the 2NN, on the other hand, the atomic distances of Mg–Zn (3.283 Å for horizontal and 3.262 Å for vertical) are slightly longer than those of Zn–Zn (3.249 Å for horizontal and 3.220 Å for vertical from the present EXAFS study [33]) in the $\text{Mg}_{0.06}\text{Zn}_{0.94}\text{O}$ layer and those in bulk ZnO (3.239 Å for horizontal and 3.208 Å for vertical from [34]).

6.4 Summary

In this chapter we have provided several examples of XAFS application to the local structure analysis of semiconducting thin films: InGaN, AlGaIn, and MgZnO. For such thin films, it is not so easy to study the structure of active layers by conventional methods. XAFS can provide unique information on atomic distribution and electronic structure of the small area in the active layer of thin films. In general, XAFS is considered as a method to give isotropic structural information, but it is capable of analyzing anisotropic structure in epitaxial films by use of linearly-polarized X-ray from synchrotron radiation sources. While the XAFS analysis in low and medium energy region (soft and tender X-ray) is not so accurate in comparison with high energy region (hard X-ray), recent development of the high performance 3rd and higher generation synchrotron radiation source provides an advantage in lower energy XAFS studies.

Acknowledgments The authors gratefully acknowledge Prof. S.F. Chichibu and Dr. K. Nitta for significant contributions and fruitful collaborations. The authors thank the Photon Factory in KEK, and JASRI in SPring-8 for financial and experimental support.

References

1. G. Bunker, *Introduction to XAFS* (Cambridge University Press, Cambridge, 2010)
2. C.S. Schnorr, M.C. Ridgway, *X-ray Absorption Spectroscopy of Semiconductors* (Springer, Heidelberg, 2015)
3. P. Fornasini, *Synchrotron Radiation: Basics, Methods and Applications*, eds. by S. Mobilio, F. Boscherini, C. Meneghini, (Springer, Heidelberg, 2015) Chap. 6
4. D. Sayers, E.A. Stern, F. Lytle, *Phys. Rev. Lett.* **27**, 1204 (1971)

5. <http://www.exafscs.com/techpapers/index.html>
6. J. Als-Nielsen, D. McMorrow, *Elements of Modern X-ray Physics*, (Wiley, Chichester, 2011)
7. H. Sakane, T. Miyanaga, I. Watanabe, N. Matsubayashi, S. Ikeda, Y. Yokoyama, Jpn. J. Appl. Phys. **32**, 4641 (1993)
8. <http://cars.uchicago.edu/ifeffit/Downloads>
9. A.L. Ankudinov, J.J. Rehr, Phys. Rev. B **62**, 2437 (2000)
10. T. Miyanaga, T. Fujikawa, J. Spectrosc. Dyn. **1**, 4 (2011)
11. S. Nakamura, G. Fasol, *The Blue Laser Diode* (Springer, Heidelberg, 1997)
12. S.F. Chichibu, Y. Kawakami, T. Sota, in *Introduction to Nitride Semiconductor Blue Lasers and Light Emitting Diodes*, eds. by S. Nakamura, S.F. Chichibu (Taylor & Francis, London, 2000), Chap. 5
13. S. Chichibu, T. Azuhata, T. Sota, S. Nakamura, Appl. Phys. Lett. **69**, 4188 (1996)
14. T. Miyanaga, T. Azuhata, S. Matsuda, Y. Ishikawa, S. Sasaki, T. Uruga, H. Tanida, S.F. Chichibu, T. Sota, Phys. Rev. B **76**, 035314 (2007)
15. S.F. Chichibu, T. Sota, K. Wada, O. Brandt, K.H. Ploog, S.P. DenBaars, S. Nakamura, Phys. Status Solidi A **183**, 91 (2001)
16. P.R.C. Kent, A. Zunger, Appl. Phys. Lett. **79**, 1977 (2001)
17. T. Mukai, M. Yamada, S. Nakamura, Jpn. J. Appl. Phys., Part 1 **38**, 3976 (1999)
18. S. Sasaki, T. Miyanaga, T. Azuhata, T. Uruga, H. Tanida, S.F. Chichibu, T. Sota, A.I.P. Conf. Proc. **882**, 499 (2008)
19. S.F. Chichibu, A.C. Abare, M.P. Mack, M.S. Minsky, T. Deguchi, D. Cohen, P. Kozoboy, S.B. Fleischer, S. Keller, J.S. Speck, J.E. Bowers, E. Hu, U.K. Mishra, L.A. Coldren, S.P. DenBaars, K. Wada, T. Sota, S. Nakamura, Mater. Sci. Eng., B **59**, 298 (1999)
20. T. Takeuchi, S. Sota, M. Katsuragawa, M. Komori, H. Takeuchi, H. Amano, I. Akasaki, Jpn. J. Appl. Phys. **36**, L382 (1997)
21. S.F. Chichibu, H. Yamaguchi, L. Zhao, M. Kubota, T. Onuma, K. Okamoto, H. Ohta, Appl. Phys. Lett. **93**, 151908 (2008)
22. T. Miyanaga, T. Azuhata, K. Nitta, S.F. Chichibu, unpublished
23. T. Koyama, T. Onuma, H. Masui, A. Chakraborty, B.A. Haskell, S. Keller, U.K. Mishra, J.S. Speck, S. Nakamura, S.P. DenBaars, T. Sota, S.F. Chichibu, Appl. Phys. Lett. **89**, 091906 (2006)
24. F.H. Morkoc, S.N. Mohammad, Science **267**, 51 (1995)
25. K.E. Miyano, J.C. Woicik, L.H. Robins, C.E. Bouldin, D.K. Wickenden, Appl. Phys. Lett. **70**, 2108 (1997)
26. K.M. Yu, C.J. Glover, M.C. Ridgway, W.S. Wong, W. Yang, Appl. Phys. Lett. **75**, 4097 (1999)
27. T. Miyanaga, T. Azuhata, K. Nakajima, H. Nagoya, K. Hazu, S.F. Chichibu, J. Phys: Conf. Ser. **502**, 012031 (2014)
28. K. Hazu, T. Hoshi, M. Kagaya, T. Onuma, S.F. Chichibu, J. Appl. Phys. **107**, 033701 (2010)
29. K. Hazu, M. Kagaya, T. Hoshi, T. Onuma, S.F. Chichibu, J. Vac. Sci. Technol., B **29**, 21208 (2011)
30. T. Koyama, T. Ohmori, N. Shibata, T. Onuma, S.F. Chichibu, J. Vac. Sci. Technol., B **22**, 2220 (2002)
31. H. Shibata, H. Tampo, K. Matsubara, A. Yamada, K. Sakurai, S. Ishizuka, S. Niki, Appl. Phys. Lett. **90**, 124104 (2007)
32. H. Ryoken, N. Ohashi, I. Sakaguchi, Y. Adachi, S. Hishita, H. Haneda, J. Cryst. Growth **287**, 134 (2006)
33. T. Yamada, T. Miyanaga, T. Azuhata, T. Koyama, S.F. Chichibu, Y. Kitajima, e-J. Surf. Sci. Nanotech. **7**, 596 (2009)
34. H. Karzel, W. Potzel, M. Köfferlein, W. Schiessl, M. Steiner, U. Hiller, G.M. Kalvius, D.W. Mitchell, T.P. Das, P. Blaha, K. Schwarz, M.P. Pasternak, Phys. Rev. B **53**, 1425 (1996)

Chapter 7

Pulsed-Laser Generation of Nanostructures

Selcuk Akturk

Abstract We discussed two main approaches for generation of nanometer-scale structures using pulsed lasers. In the first, a bulk material is ablated in vacuum, gas or liquid and consequent re-solidification of the evaporated material in clusters generate the desired nanoparticles. While one can play with various experimental parameters (laser powers, type of liquids, background gas etc.) to control particle sizes, there is always a relatively broad distribution of particle sizes. Nevertheless, particles with sizes down to few nanometers are achievable. In the second approach, fs lasers are used to pattern material surfaces. This process is more controlled in the sense that the desired geometries are generated by scanning of the sample. On the other hand, best achievable resolutions are worse as compared to the former approach. Resolutions better than 100 nm over large surfaces is very difficult. Finally, special beam shaping methods forms an interesting frontier in pushing forward the fs laser nanomachining methods towards broader practical uses.

7.1 Introduction

Ablation of materials by pulsed lasers constitute a versatile approach to generation of nanometer-scale structures. In this chapter, we will discuss two distinct approaches for pulsed-laser generation of nanostructures. In the first case, laser pulses are sent to the surface of the target material in bulk form. The evaporated material then solidifies in the form of clusters of various sizes, usually in sub-micron range. The cluster formation can take place in liquid or gas environment. In the second approach that we will discuss, the process is more “controlled”. Here, ultrashort laser pulses (with femtosecond pulse durations) are focused to small spots and by working near the ablation-threshold, volumes with sub-micron dimensions are removed. Nanostructures of desired forms and geometries are

S. Akturk (✉)

Department of Physics, Istanbul Technical University, 34469 Maslak, Istanbul, Turkey
e-mail: akturks@itu.edu.tr

generated by scanning the sample under laser illumination. We will describe the first approach (which is a relatively more mature field, and many good reviews already exist) in a brief overview manner, while we will elaborate more on the second, more recently emerging case.

7.2 Self-formation of Nanostructures Through Pulsed-Laser Ablation

Laser-assisted formation of nanostructures has a history of nearly three decades. In one of the earlier works, continuous-wave laser are used to manipulate the size distribution of metal clusters [1]. The physical mechanism of size change is basically absorption of laser energy by the particles, consequent heating and division. The method is thus suitable for processing of nanoparticles, already prepared by other means. In 1997, Morales and Lieber demonstrated the use of pulsed laser ablation in a gas environment to generate semiconductor nanowires [2]. Cluster formation and vapor-liquid-solid (VLS) growth yields the wire-type structures. Following years has seen an expanding use of laser ablation to generate nanoparticles of various materials and geometries.

Pulsed-laser self-formation of nanoparticles is mainly performed in gas/vacuum or liquid environments. Experimental principles of both cases is outlined in Fig. 7.1. In the first case, the target material is ablated with laser pulses and the evaporated material forms clusters on a stopping target. Physical mechanisms of particles formation depends strongly on the nature and the conditions of the environment [3].

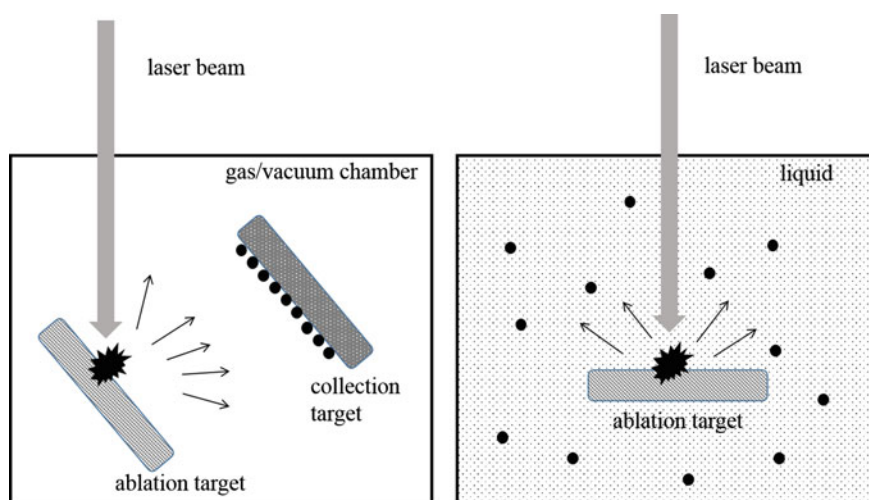


Fig. 7.1 Schematic representation of laser-assisted nanoparticle generation in gas/vacuum (*left*) and liquid (*right*) environments

Nanoparticles may form in the gas phase or from nucleation on the target substrate. An extensive review of methods and applications for the gas-phase case is collected by Kruis et al. [4]. Particles sizes are typically below 100 nm. The necessity of plasma formation on the ablation target material usually calls for nanosecond-pulsed lasers. It is also possible to use picosecond and femtosecond pulsed lasers for the same purpose, even though the particle formation mechanisms are quite different [5]. Processes performed in vacuum allow a more detailed investigation of events taking place from the laser absorption up to the final particle solidification [6].

In the second and more commonly used scheme of self-formation of nanoparticles, the ablation process takes place in a liquid environment (Fig. 7.1). Desired material of the bulk form is immersed in a background liquid, and then ablated with a pulsed laser, forming an emulsion of nanoparticles [7]. Particles sizes, again typically below 100 nm are obtained and controlled via laser fluence (pulse energy per unit area) and background liquid. Most commonly used materials are metals (Au, Ag, Ti etc.) [8] and semiconductors [9], and most common background liquids are water and ethanol. Different background liquids are also investigated for size stabilization [10]. As with the gas/vacuum case, ultrashort pulse durations are studied for the same purpose, both from a fundamental point of view (to investigate nucleation dynamics), as well as to optimize particle size distributions [11].

Lastly, nanoparticles formed by the laser processes briefly outlined here also find range of applications, especially in biology, biomedicine and chemistry. We refer the reader to excellent reviews already published in the literature [12, 13].

7.3 Formation of Nanoscale Patterns by Femtosecond Laser Ablation

7.3.1 General Principles

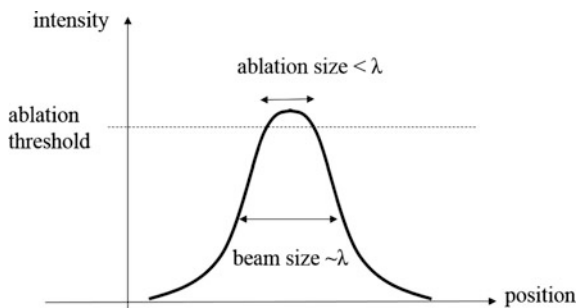
As mentioned above, self-formation of nanostructures through laser ablation in vacuum, gas or liquid environments has a relatively long history, and there are many good reviews in the literature. Hence, we presented a brief flavor from a quite broad range of methods, materials and applications. In this section we will describe, in a deeper detail, a relatively younger field of using femtosecond (fs) laser pulses for generation of nanometer-scale structures. The two processes described in the previous and present sections are fundamentally different. In the former, nanoparticles are self-formed into clusters from a gas and/or plasma phase medium (formed from bulk, using a laser). In the latter case, the process is more “controlled” in the sense that the laser beam has a well-defined ablation volume, and the process is performed by scanning the sample over the desired pattern. One can also make a crude categorization of the two cases as bottom-up and top-down, even though the nomenclature is not as distinct as the case of chemical processes, for example.

Before delving into the experimental details, it would be useful to briefly discuss the spirit of material processing with fs laser pulses. In general, laser ablation processes are based on absorption of light energy (through whatever physical process) by the electrons of the medium and consequent heating, melting and vaporization [14]. Fs laser pulses has several aspects that separate them from longer-pulse lasers. First of all, the pulse duration is typically orders of magnitude shorter than heat diffusion times, as a result, there is essentially no thermalization during the pulse period [15–17]. For transparent materials, long-pulse ablation is initiated by noise electrons, through avalanche ionization, while fs laser generate plasma via tunnel and multiphoton ionizations [18, 19]. As the latter processes are deterministic, while the former being more dependent on the presence of impurities, fs laser ablation yield much more precise, repeatable and well-defined structures (free of heat affected zones, burrs etc.) [20].

While the field of laser material processing dates back to essentially a couple years after the invention of the laser, nanometer-scale resolutions (here we define resolution as smallest ablation dimensions) have only been possible after the emergence of fs laser material processing. In a typical laser material processing experiment, the resolution is defined mainly by the overall heat affected zone and also by the focused laser beam spot size. Furthermore, the minimum focusable spot sizes are determined by wave nature of light and limited to about the wavelength of the laser (called diffraction limit). As a result of combination of both effects, even when UV lasers (e.g. excimers) are used, best achievable resolutions for CW and long-pulsed lasers are limited to well above micrometers.

The situation changes dramatically for fs laser pulses, due to the deterministic nature and minimal heat diffusion effects mentioned above. First of all, transverse spatial modes of stable laser cavities are Gaussian beams and in the most common lowest order (the so-called TEM_{00} mode) have a single-peak bell-like shapes [21]. The Gaussian shape is also preserved through focusing, with transformation to a different spot size. When a fs laser beam is focused, to a near-diffraction-limited spot, of diameter say $\sim 1 \mu\text{m}$ (for a visible or near-infrared laser), the intensity distribution would look schematically like the cross section illustrated in Fig. 7.2. As a result, by adjusting the laser intensity to proper levels, such that only the central region of the focused beam exceeds the ablation threshold, resolutions well

Fig. 7.2 Generation of sub-diffraction-limit ablation sizes by working near critical intensity



below the diffraction limit, and in the nanometer range can be achieved [22, 23]. In one of the earliest proofs of this principle of “optics at critical intensity”, resolutions of ~ 30 nm are obtained with a laser wavelength at 1053 nm [22]. The ultimate resolution of this approach is limited by the pulse-to-pulse stability of the laser system and mechanical stability of the overall setup. The method is suitable for a broad range of materials including metals, dielectrics and polymers [24].

7.3.2 Femtosecond Laser Nanomachining on Thin Films with Bessel Beams

Nanometer-scale ablation by fs laser pulses explained above by itself is not sufficient to generate functional nanostructures. The process simply yields a sensitive “tool” to process larger-scale materials. One method for generating functional nanostructures with this tool is to start with “thin films” (films with sub-micron thicknesses) of desired materials, deposited on a suitable substrate, and generate a two dimensional ablation pattern, via scanning of the fs laser beam [25, 26].

This nanofabrication method, although limited to materials of which thin films can be formed, yields a direct, versatile and relatively simple approach. A typical schematic of such experiments is shown in Fig. 7.3. With a fs laser of high repetition rate, millimeter-scale regions can be processed within minutes. As a result, the approach is particularly suitable for prototype investigations (not so much for serial production). Most processes are performed in ambient air.

In the experimental realization of fs laser nanomachining, particularly on large (tens of microns to millimeter-scale) surfaces, optical alignments and sample positioning are quite critical and requires extreme care. High resolutions require high-numerical aperture focusing, and the resulting tightly-focused beams diverge very rapidly. Roughly speaking, beams focused to transverse sizes of $\sim$ wavelength stay intense only within a longitudinal depth of $\sim$ one wavelength. Since the essence of fs nanomachining is working sensitively close to the ablation threshold, minute changes in the intensity due to beam divergence, sample positioning and sample surface roughness can yield deleterious effects.

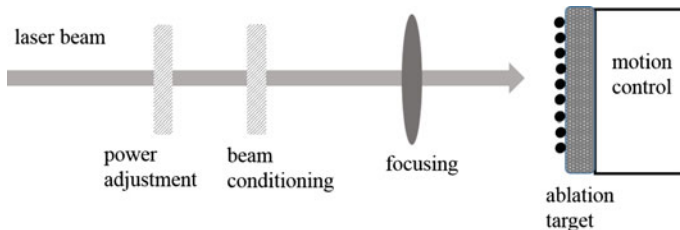


Fig. 7.3 Typical layout of a fs laser material processing experiment

These critical constraints resulting from rapid diffraction of focused Gaussian beams can be remedied by exploiting the so-called “non-diffracting beams” [27]. These beams have transverse field envelope profiles in the form of Bessel function of the first kind and have the property of retaining the same intensity and beam profile regardless of free-space propagation [28, 29]. Gaussian beam outputs of typical laser cavities can be turned into quasi-Bessel beams by using conical lenses, or axicons [30, 31]. Some example profiles of theoretical and experimental Bessel beams is shown in Fig. 7.4. Defining the width of a Bessel beam as the breadth of the central peak, the focal depths (length of Bessel zones) of Bessel beams can be one or two orders of magnitude longer as compared to Gaussian beams of the same width.

The potential advantages of Bessel beams in fs nanomachining should be clear. Since the beam size does not change and the intensity changes very slowly within the Bessel zone, the constraints related to positioning and alignments mentioned above are significantly relaxed. Furthermore, it is much easier to use Bessel beams in the ultraviolet wavelengths (by using a suitable axicon), as compared to Gaussian

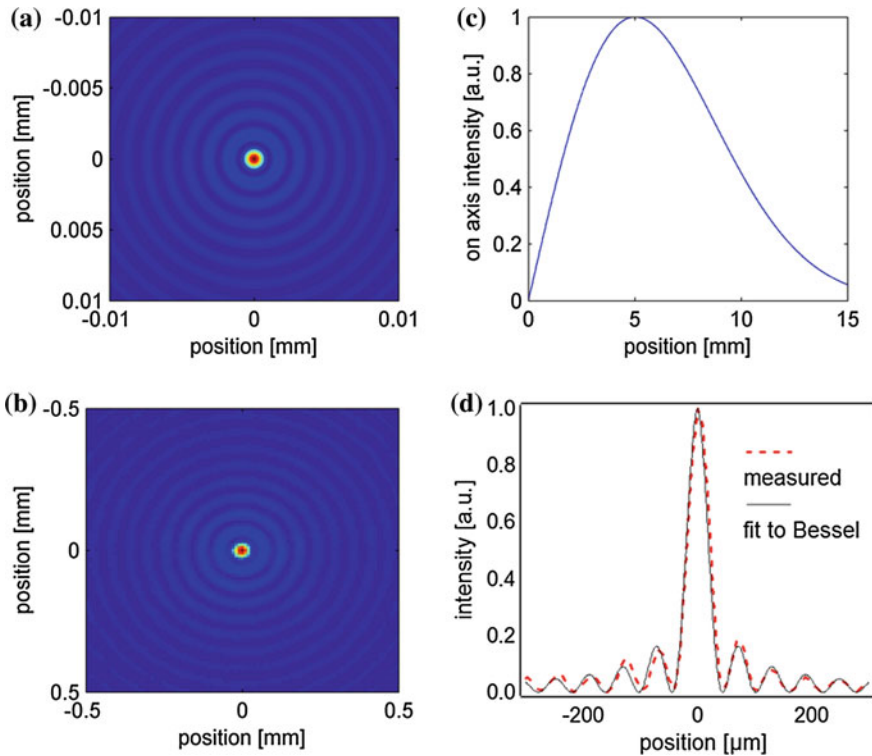


Fig. 7.4 Theoretical transverse profile (a) and longitudinal intensity variation (b) of a Bessel beam. Experimental Bessel beams generated by an axicon has a very similar transverse profile (c) closely following the Bessel function (d)

beams, which require highly costly aberration-corrected objectives [32]. Note that, when working close to ablation threshold, multiple-rings surrounding the main peak do not contribute to the ablation pattern.

7.3.3 Experimental Results

The virtues of Bessel beams are exploited in many recent fs laser ablation experiments. In a transparent sample, Bessel beams can form high-aspect-ratio channels without mechanical scanning [33–35]. Similarly, it is also possible to ablate surfaces (in this case, one does not exploit long-focal-line nature of Bessel beams, but more the non-diffracting nature) [36]. In this section, we will present some experimental results on using Bessel beam fs lasers to generate nanostructures on thin film surfaces.

Figure 7.5 shows nanometer scale channels and nanoparticle arrays formed on Au thin films (~ 20 nm thickness) deposited on glass [37]. The laser used in this process is a Yb:Glass based chirped-pulse amplification (CPA) system, delivering 550 fs pulses at 1 kHz. Second-harmonic of the output (515 nm) is generated to enhance resolution. A glass axicon with 40° base angle is used to generate Bessel beams. The diameter of Bessel beams decreases with increasing axicon base angle [38], and this axicon is very close to the maximum, allowed for avoiding total internal reflection inside the glass.

Metallic nanostructures of types shown in Fig. 7.5 can be readily used for applications. Metal nanostructure arrays are particularly suitable for plasmonics applications [39]. Plasmonic behaviors are observed in structures generated with this method [40].

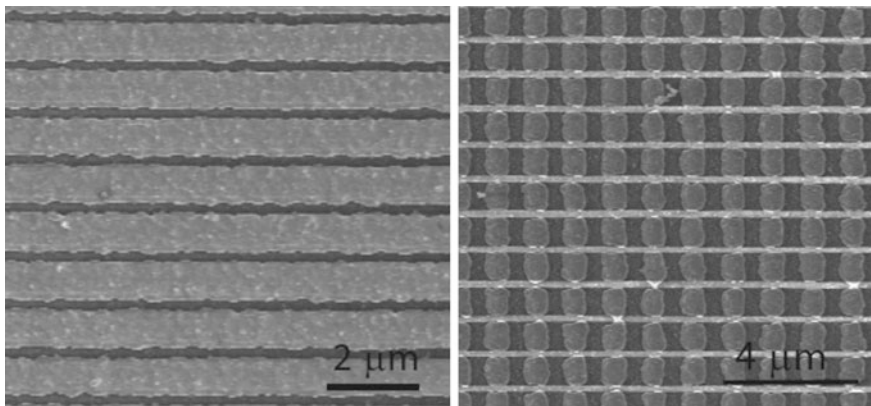


Fig. 7.5 Nanometer scale channels (*left*) and nanoparticle arrays (*right*) formed on Au thin films

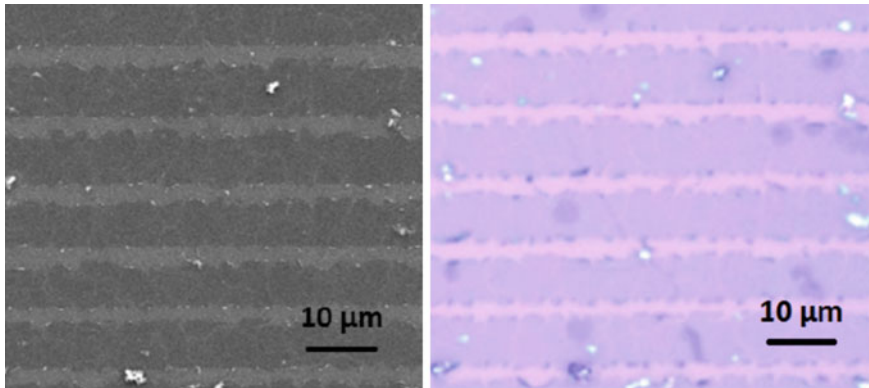


Fig. 7.6 SEM (*left*) and optical (*right*) images of graphene-on-SiO₂ samples ablated with fs laser Bessel beams

The second example that we present is on the emerging field of two-dimensional (single-atomic layer) materials such as graphene [41]. These materials have exotic physical behaviors and potentials for application in the future of electronics. Yet, high resolution patterning of graphene and similar structures is still a challenge. In this regard, the methods summarized in this section provides important opportunities. In recent proof-of-principle experiments, nanometer-scale patterning of graphene surface is demonstrated [42]. An example of graphene surface, patterned with fs laser Bessel beams (with experimental parameters similar to above) is shown in Fig. 7.6.

7.4 Conclusions

In this chapter, we discussed two main approaches for generation of nanometer-scale structures using pulsed lasers. In the first, a bulk material is ablated in vacuum, gas or liquid and consequent re-solidification of the evaporated material in clusters generate the desired nanoparticles. While one can play with various experimental parameters (laser powers, type of liquids, background gas etc.) to control particle sizes, there is always a relatively broad distribution of particle sizes. Nevertheless, particles with sizes down to few nanometers are achievable. In the second approach, fs lasers are used to pattern material surfaces. This process is more controlled in the sense that the desired geometries are generated by scanning of the sample. On the other hand, best achievable resolutions are worse as compared to the former approach. Resolutions better than 100 nm over large surfaces is very difficult. Finally, special beam shaping methods forms an interesting frontier in pushing forward the fs laser nanomachining methods towards broader practical uses.

References

1. M. Vollmer, R. Weidenauer, W. Hoheisel, U. Schulte, F. Träger, Size manipulation of metal particles with laser light. *Phys. Rev. B.* **40**(18), 12509–12512 (1989)
2. A.M. Morales, C.M. Lieber, A laser ablation method for the synthesis of crystalline semiconductor nanowires. *Science* **279**(5348), 208–211 (1998)
3. D.B. Geoghegan, A.A. Puretzky, G. Duscher, S.J. Pennycook, Time-resolved imaging of gas phase nanoparticle synthesis by laser ablation. *Appl. Phys. Lett.* **72**(23), 2987–2989 (1998)
4. F.E. Kruis, H. Fissan, A. Peled, Synthesis of nanoparticles in the gas phase for electronic, optical and magnetic applications—a review. *J. Aerosol Sci.* **29**(5–6), 511–535 (1998)
5. S. Eliezer, N. Eliaz, E. Grossman, D. Fisher, I. Gouzman, Z. Henis et al., Synthesis of nanoparticles with femtosecond laser pulses. *Phys. Rev. B.* **69**(14), 144119 (2004)
6. S. Amoruso, G. Ausanio, R. Bruzzese, M. Vitiello, X. Wang, Femtosecond laser pulse irradiation of solid targets as a general route to nanoparticle formation in a vacuum. *Phys. Rev. B.* **71**(3), 033406 (2005)
7. S.I. Dolgaev, A.V. Simakin, V.V. Voronov, G.A. Shafeev, F. Bozon-Verduraz, Nanoparticles produced by laser ablation of solids in liquid environment. *Appl. Surf. Sci.* **186**(1–4), 546–551 (2002)
8. A.V. Simakin, V.V. Voronov, N.A. Kirichenko, G.A. Shafeev, Nanoparticles produced by laser ablation of solids in liquid environment. *Appl. Phys. A* **79**(4–6), 1127–1132 (2004)
9. K.V. Anikin, N.N. Melnik, A.V. Simakin, G.A. Shafeev, V.V. Voronov, A.G. Vitukhnovsky, Formation of ZnSe and CdS quantum dots via laser ablation in liquids. *Chem. Phys. Lett.* **366**(3–4), 357–360 (2002)
10. J.-P. Sylvestre, A.V. Kabashin, E. Sacher, M. Meunier, J.H.T. Luong, Stabilization and size control of gold nanoparticles during laser ablation in aqueous cyclodextrins. *J. Am. Chem. Soc.* **126**(23), 7176–7177 (2004)
11. J.-P. Sylvestre, A.V. Kabashin, E. Sacher, M. Meunier, Femtosecond laser ablation of gold in water: influence of the laser-produced plasma on the nanoparticle size distribution. *Appl. Phys. A* **80**(4), 753–758 (2004)
12. I.H. El-Sayed, X. Huang, M.A. El-Sayed, Surface plasmon resonance scattering and absorption of anti-EGFR antibody conjugated gold nanoparticles in cancer diagnostics: applications in oral cancer. *Nano Lett.* **5**(5), 829–834 (2005)
13. M.-C. Daniel, D. Astruc, Gold nanoparticles: assembly, supramolecular chemistry, quantum-size-related properties, and applications toward biology, catalysis, and nanotechnology. *Chem. Rev.* **104**(1), 293–346 (2004)
14. W.M. Steen, *Laser Material Processing*, 3rd edn. (Springer, London, 2008)
15. B.C. Stuart, M.D. Feit, S. Herman, A.M. Rubenchik, B.W. Shore, M.D. Perry, Nanosecond-to-femtosecond laser-induced breakdown in dielectrics. *Phys. Rev. B.* **53**(4), 1749 (1996)
16. B.C. Stuart, M.D. Feit, A.M. Rubenchik, B.W. Shore, M.D. Perry, Laser-induced damage in dielectrics with nanosecond to subpicosecond pulses. *Phys. Rev. Lett.* **74**(12), 2248 (1995)
17. S. Nolte, C. Momma, H. Jacobs, A. Tünnermann, B.N. Chichkov, B. Wellegehausen et al., Ablation of metals by ultrashort laser pulses. *J. Opt. Soc. Am. B.* **14**(10), 2716–2722 (1997)
18. X. Liu, D. Du, G. Mourou, Laser ablation and micromachining with ultrashort laser pulses. *IEEE J. Quantum Electron.* **33**(10), 1706–1716 (1997)
19. D. Du, X. Liu, G. Korn, J. Squier, G. Mourou, Laser-induced breakdown by impact ionization in SiO₂ with pulse widths from 7 ns to 150 fs. *Appl. Phys. Lett.* **64**(23), 3071–3073 (1994)
20. R.R. Gattass, E. Mazur, Femtosecond laser micromachining in transparent materials. *Nat. Photon.* **2**(4), 219–225 (2008)
21. A.E. Siegman, *Lasers* (Univ Science Books, 1986)
22. A.P. Joglekar, H. Liu, E. Meyhöfer, G. Mourou, A.J. Hunt, Optics at critical intensity: applications to nanomorphing. *Proc. Natl. Acad. Sci. USA* **101**(16), 5856–5861 (2004)

23. P.P. Pronko, S.K. Dutta, J. Squier, J.V. Rudd, D. Du, G. Mourou, Machining of sub-micron holes using a femtosecond laser at 800 nm. *Opt. Commun.* **114**(1–2), 106–110 (1995)
24. F. Korte, J. Serbin, J. Koch, A. Egbert, C. Fallnich, A. Ostendorf et al., Towards nanostructuring with femtosecond laser pulses. *Appl. Phys. A* **77**(2), 229–235 (2003)
25. J. Kim, S. Na, Metal thin film ablation with femtosecond pulsed laser. *Opt. Laser Technol.* **39**(7), 1443–1448 (2007)
26. V. Mizeikis, S. Juodkazis, K. Sun, H. Misawa, Fabrication of micro- and nanostructures in thin metallic films by femtosecond laser ablation. [cited 2015 Jun 11], pp. 75910 J–75910 J–11 (2010). <http://dx.doi.org/10.1117/12.843074>
27. K. Dholakia, Optics: Against the spread of the light. *Nature* **451**(7177), 413 (2008)
28. J. Durnin, J.J. Miceli, J.H. Eberly, Diffraction-free beams. *Phys. Rev. Lett.* **58**(15), 1499–1501 (1987)
29. J. Durnin, Exact solutions for nondiffracting beams. I. The scalar theory. *J. Opt. Soc. Am. A* **4**(4), 651–654 (1987)
30. J.H. McLeod, Axicons and their uses. *J Opt Soc Am.* **50**(2), 166–169 (1960)
31. J. Arlt, K. Dholakia, Generation of high-order Bessel beams by use of an axicon. *Opt. Commun.* **177**(1–6), 297–301 (2000)
32. R. Sahin, T. Ersoy, S. Akturk, Ablation of metal thin films using femtosecond laser Bessel vortex beams. *Appl. Phys. A* **9**, 1–5 (2014)
33. F. Courvoisier, J. Zhang, M.K. Bhuyan, M. Jacquot, J.M. Dudley, Applications of femtosecond Bessel beams to laser ablation. *Appl. Phys. A*. 1–6
34. M.K. Bhuyan, F. Courvoisier, P.A. Lacourt, M. Jacquot, R. Salut, L. Furfaro et al., High aspect ratio nanochannel machining using single shot femtosecond Bessel beams. *Appl. Phys. Lett.* **97**(8), 081102 (2010)
35. M.K. Bhuyan, F. Courvoisier, P.-A. Lacourt, M. Jacquot, L. Furfaro, M.J. Withford et al., High aspect ratio taper-free microchannel fabrication using femtosecond Bessel beams. *Opt. Express* **18**(2), 566–574 (2010)
36. F. Courvoisier, P.-A. Lacourt, M. Jacquot, M.K. Bhuyan, L. Furfaro, J.M. Dudley, Surface nanoprocessing with nondiffracting femtosecond Bessel beams. *Opt. Lett.* **34**(20), 3163–3165 (2009)
37. B. Yalizay, T. Ersoy, B. Soyly, S. Akturk, Fabrication of nanometer-size structures in metal thin films using femtosecond laser Bessel beams. *Appl. Phys. Lett.* **100**(3), 031104–031104–3 (2012)
38. S. Akturk, B. Zhou, B. Pasquiou, M. Franco, A. Mysyrowicz, Intensity distribution around the focal regions of real axicons. *Opt. Commun.* **281**(17), 4240–4244 (2008)
39. S.A. Maier, *Plasmonics: Fundamentals and Applications* (Springer, 2007), 234 pp
40. R. Sahin, Y. Morova, E. Simsek, S. Akturk, Bessel-beam-written nanoslit arrays and characterization of their optical response. *Appl. Phys. Lett.* **102**(19), 193106 (2013)
41. A.K. Geim, K.S. Novoselov, The rise of graphene. *Nat. Mater.* **6**(3), 183–191 (2007)
42. R. Sahin, E. Simsek, S. Akturk, Nanoscale patterning of graphene through femtosecond laser ablation. *Appl. Phys. Lett.* **104**(5), 053118 (2014)

Chapter 8

Graphene for Silicon Microelectronics: Ab Initio Modeling of Graphene Nucleation and Growth

Jarek Dabrowski, Gunther Lippert and Grzegorz Lupina

Abstract Graphene electronics is expected to complement the conventional Si technologies. Graphene processing should thus be compatible with the mainstream Si technology: CMOS. Ideally, it should be possible to grow graphene directly on a Si wafer, but this does not work. Large area graphene can be grown on Cu or on Ni, its transfer to silicon must then follow, which is problematic. Researchers try therefore to grow graphene on CMOS compatible substrates, such as on Ge/Si(001) wafers. Ab initio modeling, particularly when used in combination with experimental data, can elucidate the mechanisms that govern the process of nucleation and growth of graphene, and thus provide assistance in the design of experiments and production processes. We overview our results in this context, starting from atomic C deposited on (chemically inert) graphene, through the similar cases of Si deposited on graphene and C deposited on hexagonal boron nitride, and the case of carbon on a non-inert insulator (SiO₂-like surface of mica), up to C atoms and hydrocarbon molecules building graphene on Ge(001) surfaces.

8.1 Introduction

Graphene, a two-dimensional sheet of sp²-bonded carbon atoms, is a semi-metal with linear energy-momentum relation around the Fermi energy. This photon-like law has its origin in the symmetry of the atomic lattice and makes the charge carriers massless, just like photons are. One of the consequences is that the sheet conductivity of even a relatively highly defected graphene sheet may be comparable

J. Dabrowski (✉) · G. Lippert · G. Lupina
IHP, Im Technologiepark 25, 15236 Frankfurt(Oder), Germany
e-mail: dabrowski@ihpmicroelectronics.com

G. Lippert
e-mail: lippert@ihpmicroelectronics.com

G. Lupina
e-mail: lupina@ihpmicroelectronics.com

to what would be needed for a heterojunction bipolar transistor (HBT) designed to operate at terahertz frequencies. And perfect enough graphene used as the channel material may enable a field effect transistor (FET) to work in the THz range. Graphene is thus supposed to give the traditional microelectronics based on Si FETs or HBTs the chance to cross the terahertz boundary [1–3].

Graphene can be formed from SiC on Si [4, 5] or grown on a metal substrate (usually, on Cu) and then transferred to Si [6]. However, practical constraints make both these approaches difficult to integrate with the chip production. For example, transfer from copper leads to contamination with Cu atoms [7]. Graphene can also be grown directly on Si or on SiO₂ by plasma-enhanced chemical vapor deposition (PE-CVD) [8]. Yet such films have too many graphene layers for FET-like applications. Moreover, chemical interaction of C with Si results in the formation of carbide, excluding the usage of this material as the base electrode in HBT-like devices. On the other hand, graphene can be grown on Ge by CVD [9] or even by molecular beam epitaxy (MBE), i.e., from atomic source [10]. This is good news, because Ge is a material that is fully compatible with the standard silicon technology.

In this chapter we work out some theoretical aspects of the nucleation and growth of graphene on semi-metallic and non-metallic substrates. We put most weight on the chemically simplest case, which is deposition of carbon atoms in a vacuum MBE chamber. We begin with the single-species system: carbon on graphene. Comparison to the growth of Si films on graphene is then done [11]. We use these two cases together with that of carbon on hexagonal boron nitride (hBN) to illustrate the behavior of atoms on a substrate with which they interact only weakly. We then turn attention to the substrate that is insulating and in most cases chemically inert, but yet highly reactive to carbon: to hexagonal SiO₂ surfaces and to similar materials, exemplified by mica. Finally, we analyze the process of nucleation and growth of graphene on the Ge(001) surface, for which we make also a brief excursion into the CVD process. The latter section is preceded by remarks on the graphene base transistor (GBT) [12–14], a device in which graphene grown directly on Ge may be a good combination of materials.

In the latter context, we performed calculations for electron tunneling through graphene sheet [13]. It turns out that tunneling probability may have to be accounted for in the simulations of GBT electrical performance, particularly when the lateral momentum is conserved in the tunneling process, as it may be the case if hexagonal BN (hBN) is used to separate the bottom electrode from the base.

Hexagonal BN is generally interesting as a substrate for graphene because its interaction with carbon is weak enough for high carrier mobility to be retained in graphene but seems to be strong enough to allow for some epitaxial relationship between the substrate and the film. It may be viewed as a potential substrate also for CMOS-GFET integration. We learned that graphene nucleates on h-BN at pre-existing defect sites, including C and B₂O₃ impurities, and estimated the sticking coefficient of C as a function of average distance between the adsorbing sites [15].

One of the first successful experimental growth studies of graphene on an insulating surface have however been performed by MBE deposition of C on mica, i.e., on a surface closely resembling SiO_2 but having few surface defects. From theoretical analysis we learned that due to strong interaction of C with O and Si, only deposition on defect-free, insulating oxide surfaces may lead to graphene films of acceptable quality [16–18]. While this can be achieved on mica, the density of nucleation sites on SiO_2 is too high: the mica process cannot be transferred to SiO_2 .

The final substrate considered in this Chapter is germanium. In contrast to the case of Si, deposition of C onto Ge does not deteriorate the substrate by carbide formation. However, graphene grown by MBE on Ge(001) is separated from the substrate by a poorly conducting film that is probably built of nanocrystalline graphene interconnected by Ge atoms [10] and which may be better to be eliminated if an application in a THz device is targeted. An ab initio study of the interaction of carbon and hydrocarbons with Ge(001) provided insight into the nucleation of this film.

8.2 Approach

Electronic structures, total energies, forces, and atomic positions were calculated in the framework of the Density Functional theory (DFT) as implemented in the *Quantum Espresso* package designed for massively parallel computations [19]. The exchange and correlation energy of electrons was approximated by the gradient corrected functional of Perdew, Burke and Ernzerhof [20]. Spin polarization was accounted for. The atoms were described by the pseudopotential method using custom ultrasoft potentials and a plane wave basis set with the energy cut-off of 30 Ry for the wave functions and 240 eV for the electron density. The calculations for Ge(001) substrates (rectangular surface cell) were done with two special k -points from the Brillouin zone (or with a sampling derived from this k -point set), while the calculations for substrates with hexagonal surface cells (hexagonal boron nitride, model mica, graphene) were done with the Γ sampling (or with a Γ -derived sampling).¹ Structures were assumed to be converged when the maximum force dropped below 0.01 eV/Å and the expected change of the total energy by further relaxation was insignificant for the final conclusion. Energy barriers on migration and reaction paths and the corresponding atomic configurations have been obtained using the Nudged Elastic Band/Climbing Image (NEB/CI) algorithm [21, 22],

¹Derived means here “using the same set but taken from Brillouin zone of a larger surface cell”. This is done with caution, as it costs computing time. For example, calculation using the Γ point from the 24×24 cell of graphene (this is a fully converged k -point set) requires 10 times more wall time (or 10 times more cores if the same wall time is requested) than the standard calculation done with the Γ point from the 6×6 cell (this is a poor set for adsorption energies, but a reasonable one for diffusion barriers). Therefore, the structures are computed using the fundamental k -point set and only in the most interesting cases refined using the fully converged set.

which searches for saddle points on the energy surface. Where applicable, the energies and barriers are compatible with those obtained in previous DFT studies [23–25].

8.3 Carbon on Graphene

When carbon atoms are deposited on any material and graphene then nucleates and grows, sooner or later more and more of the incoming carbon atoms do not hit the original substrate but they land on the newly grown graphene film. Carbon on graphene, while being the chemically simplest system among those treated in this Chapter, is therefore also the one of the fundamental importance.

A carbon atom approaches graphene from vacuum without any barrier. When adsorbed, its most stable position is a bridge site: it stands above two neighboring C atoms of graphene, making two bonds to their protruding orbitals. These were originally p_z orbitals, but the interaction with the additional atom adds an s component, transforming them towards a hybrid with sp^3 character. This creates tension in the substrate, because in response to the changed hybridization, the preferred angles to the neighboring atoms change as well, from sp^2 towards tetrahedral sp^3 angles. A tiny elevation is thus produced, with the adsorbed C on its top (Fig. 8.1). The adsorption energy, equal to the desorption barrier, is small: 1.3 eV, only by 0.7 eV higher than the barrier for diffusion. The consequence is that when the temperature of the substrate is as high as during typical deposition (above 800 °C), most of the adsorbed C desorbs after diffusing for only short time. The probability that a deposited C atom will combine with another one into a C_2 dimer can be roughly estimated by the product of the deposition rate R and the lifetime t_{life} of the ad-atom on the surface (Fig. 8.2). When $p = t_{\text{life}}R$ approaches unity, the adsorbed C may be hit by another atom arriving from vacuum.² But at typical deposition temperatures of around 850 °C (the dashed line in Fig. 8.2) and realistic deposition rates R the life time is way too short for spontaneous formation of carbon clusters that stick to graphene.

The probability of C_2 formation in fact higher. This is also because a C atom sitting on the surface can kick its C neighbor out from the sheet and take its place in the crystal. The result of equivalent to the C ad-atom punching through the graphene, from its top to its bottom surface. Once on the other side, the atom may bind itself to the underlying material, so that the life time of atomic C in the adsorbed state increases. Nevertheless, it turns out that *trimers* (C_3) are unbound on graphene: they desorb immediately after being formed. This means that preexisting nucleation sites are needed for another graphene layer to form on top of the complete one (Fig. 8.3).

²This ignores the probability that two ad-atoms find one another; for full treatment, see Fig. 8.5a.



Fig. 8.1 Adsorption, diffusion and desorption of C atoms on graphene. The gray trails connect the initial and final atomic configurations on the NEB/CI reaction path

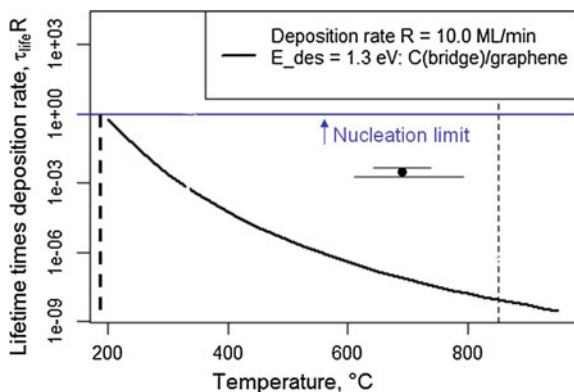


Fig. 8.2 The C_2 dimer formation probability p on graphene, estimated as $t_{life}R$, where t_{life} is the lifetime of C ad-atom on the surface and R is the deposition rate. A typical deposition rate $R = 10$ monolayers per minute is assumed. The lifetime is obtained from $t_{life} = f_{des} \exp(-E_{des}/kT)$, where E_{des} is the desorption barrier of 1.3 eV (Fig. 8.1) and the attempt frequency f_{des} is taken to be $f_{des} = 10^{13} \text{ s}^{-1}$

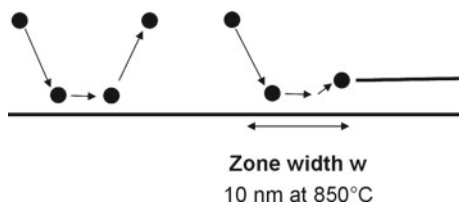


Fig. 8.3 Collection zone around an adsorbing defect on graphene (e.g., step edge). To be adsorbed, the C atom must fall within the distance w from the defect. The atoms that land further away, desorb to vacuum without contributing to the graphene growth

Because the time span t_{life} between the adsorption and desorption events is limited, so is the distance walked on the surface by the adsorbed atom. To the growth of graphene can therefore contribute only those of the impinging atoms, which hit the surface close enough to the nucleation center. Using the random walk theory one can easily show [15] that the collection zone width w (defined as such

distance from a linear adsorbing boundary within which ad-atoms can be treated as contributing to growth and beyond which they can be treated as non-contributing) equals to

$$w = s \sqrt{\frac{f_{\text{mig}}}{f_{\text{des}}}} \exp\left(\frac{E_{\text{des}} - E_{\text{mig}}}{2kT}\right), \quad (8.1)$$

where s is a distance of the order of the jump length, f_{mig} and f_{des} are attempt frequencies for diffusion and desorption, E_{mig} and E_{des} are the corresponding energy barriers, and kT is the thermal energy. Please note that $1/w$ is thermally activated with *only half* of the difference between the desorption and diffusion barriers. The factor $\frac{1}{2}$ in the exponential term stems from the random character of the walk. Indeed, it is the well-known result of the random walk theory that the distance by which a random walker departs from the initial point after N jumps scales as $\sqrt{N}$.

At 850 °C, the estimated zone width is $w = 10$ nm, or several lattice constants. The MBE growth temperature cannot be much smaller than that; the graphene quality would otherwise deteriorate. The sticking coefficient of C on good graphene is thus tiny: very many monolayers must be deposited to produce a single ad-layer.

8.4 Silicon on Graphene

In order to build a useful graphene transistor, one must deposit some material on the graphene sheet. In the case of a GBT, this can be a silicon film, for example.

Figure 8.4 illustrates a somewhat intriguing effect observed [11] after MBE deposition of Si atoms onto graphene (transferred from Cu onto oxidized Si). Though the film is in general rough and grainy, one notes that in the case of sample 1 there are regions where a relative smooth Si layer was produced. A closer examination reveals that where the film is smooth, the graphene was somewhat

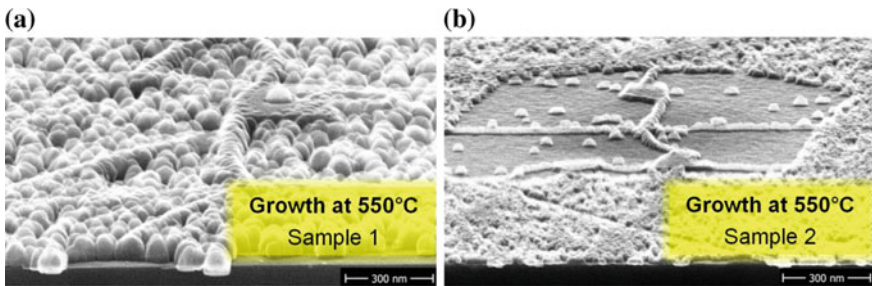


Fig. 8.4 Scanning electron microscope (SEM) images of two graphene samples after similar MBE deposition process of Si at 550 °C [11]. The Si overlayer in sample 1 is closed but inhomogeneous, apart from quite a flat region that grew on an island of a thicker (probably bilayer) graphene. A similar though larger island is visible on sample 2, but in this case it is not overgrown

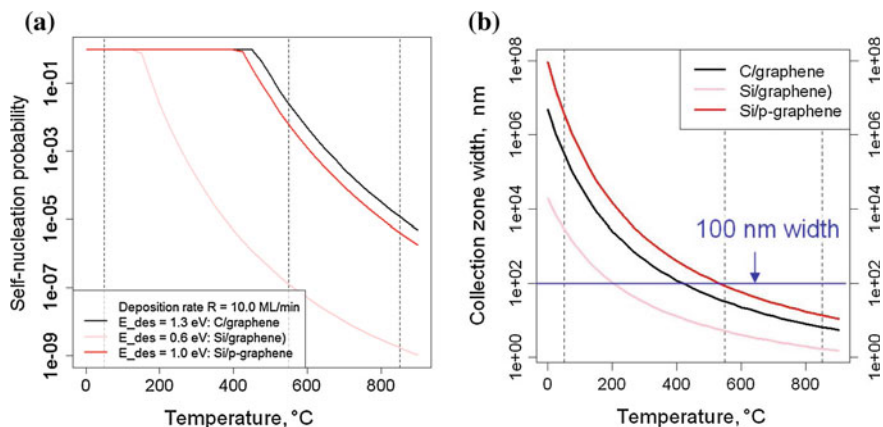


Fig. 8.5 **a** Spontaneous nucleation probabilities and **b** collection zone widths w , obtained for MBE deposition of Si atoms [11] and C atoms [15] on graphene from ab initio results by applying the random walk theory [15]

thicker, probably consisting of two carbon layers (such islands appear often on single-layer graphene sheets). But in contrast to that, the same kind of islands on sample 2 were not overgrown. Graphene has been transferred to both samples in the same process; the difference is that graphene on sample 1 stems from different provider than on sample 2.

The solution to this puzzle is suggested by results of ab initio calculations [11]. First, one finds that Si atoms stick to graphene even worse than C atoms do (Fig. 8.5). The adsorption energy (and the desorption barrier) of Si on undoped graphene is only 0.58 eV, less than half of that for C. On p-type graphene it may increase noticeably due to electron transfer from Si (Fig. 8.5). The self-nucleation probability³ is low at 550 °C and above, but many dimers form below about 200 °C for undoped or about 450 °C for p-doped graphene because two ad-atoms spent enough time on the surface and visit enough sites there to find one another (Fig. 8.5a); yet this contributes only to Si loss, because in contrast to C dimers, Si dimers do not stick to graphene at all, apart from being attracted to it by weak van der Waals forces. But when nucleation sites are present on graphene, Si atoms may stick to them, and the collection zone increases strongly when the graphene is p-doped (Fig. 8.5b). At 550 °C (the dashed line in Fig. 8.5b), the collection zone width on p-type graphene is estimated to be of the order of 100 nm, which compares well to the distance between the grains in Fig. 8.4. Overgrowth with Si is therefore possible even at 550 °C if graphene is p-doped and if there is enough nucleation sites available. Indeed, the nucleation was found to correlate with the surface concentration of PMMA rests.⁴

³Figure 8.5a includes the probability that two ad-atoms find one another on the surface.

⁴PMMA is the polymer used during graphene transfer from Cu to the target wafer.

8.5 Carbon on h-BN

Hexagonal boron nitride (hBN) is a layered compound with a graphene-like atomic structure. The graphene lattice sites are now occupied by alternating B and N atoms, and the lattice constant differs by only 4 % from the lattice constant of graphene. Consequently, some epitaxial relationship may be achieved in the hBN/graphene system, whereby the mismatch of the lattice constant leads to the appearance of Moiré patterns in AFM and STM images and to electronic super-structure effects.

When C atoms are deposited on hBN, graphene can be formed, but the morphology of the resulting film varies. We have therefore analyzed by ab initio calculations the influence of typical impurities on the nucleation of graphene on hBN and compared the predictions to the experimental data obtained on hBN flakes with non-negligible concentration of carbon, oxygen, and boron oxide (Fig. 8.6).

Figure 8.7 compares the adsorption data for C ad-atoms and ad-dimers on perfect hBN and perfect graphene. It is clear that sticking of carbon to both substrates is comparable and weak. It is thus no surprise that also on hBN the self-nucleation path of carbon is broken by desorption of carbon trimers, C_3 .

It may be interesting to verify whether the formation of C tetramers, C_4 , from two dimers may provide a nucleation path bypassing the trimer desorption. The answer is: yes, it can, but this mechanism has low efficiency. During the whole deposition process, only about 200 tetramers were created in this way in the area imaged in Fig. 8.6. This makes a negligible fraction of the number of carbon atoms that are incorporated into graphene during the same time.

But C atoms, dimers, and also bigger C clusters including C_3 are trapped by C–C donor-acceptor pairs: two substitutional C atoms occupying two neighboring sites in the hBN lattice (Fig. 8.8). This explains the appearance of isolated graphene islands.

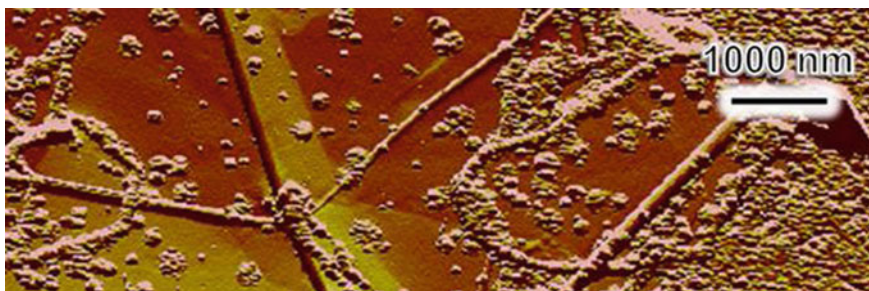


Fig. 8.6 AFM image of hBN after deposition of C atoms by MBE at 850 °C [15]. Flat areas, isolated islands, lines of islands, island-free steps, and areas covered by coalesced islands are recognizable

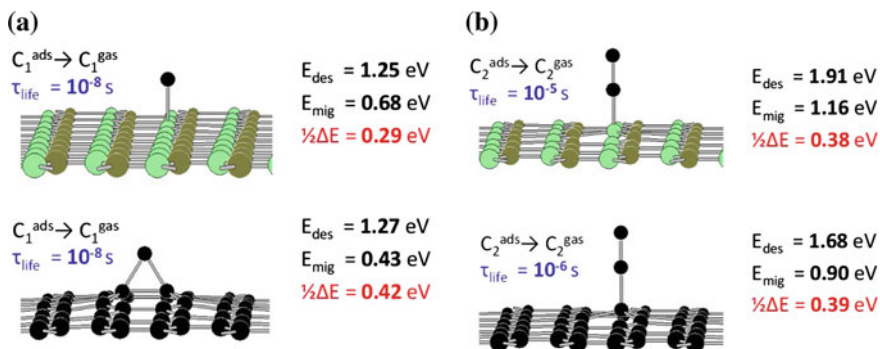


Fig. 8.7 Comparison of the atomic structures and energies on perfect hBN and on perfect graphene for **a** C ad-atoms and **b** C₂ ad-dimers [15]. Carbon is *black*, boron is *dark green*, nitrogen is *light green*, $\Delta E = E_{\text{des}} - E_{\text{mig}}$. Life times between adsorption and desorption events are given for $T_{\text{substrate}} = 850^\circ\text{C}$, assuming the attempt frequencies f_{des} and f_{mig} to be equal to 10^{13} s^{-1}

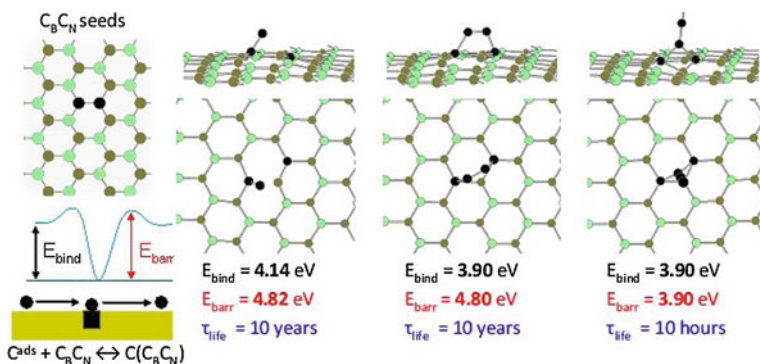


Fig. 8.8 Trapping of C ad-atoms, ad-dimers and ad-trimers by C–C donor-acceptor pairs

From the ab initio data and from random walk theory it follows that [15]:

1. MBE carbon desorbs from perfect hBN and from perfect graphene, at least at substrate temperatures as high as needed to obtain reasonably crystalline films.
2. The observed island heights are compatible with the total exposure to carbon and with theoretical sticking coefficients obtained for the density of adsorbing boundaries suggested by the analysis of the Raman spectra.
3. The sticking coefficient computed for average the monatomic step density observed on the fresh substrate is high enough to justify the assumption that the flat areas visible in AFM images are covered by monolayer graphene.
4. The islands in the densely covered areas may have nucleated on B₂O₃ inclusions and on oxygen-carbon donor-acceptor pairs. Contribution from carbon-carbon donor-acceptor pairs and from larger C [?] precipitates (graphene-like inclusions)

is possible as well, provided that the size distribution of carbon clusters in the virgin hBN substrate is far from thermal equilibrium.

5. The lines of islands are likely to be associated with nucleation on boron-terminated substrate steps that have been oxidized by ambient oxygen.

8.6 Carbon on Mica

The case of atomic carbon deposited on the surface of mica [16–18] is interesting for several reasons. First, the surface of mica is built mostly of SiO_2 that, due to the hexagonal arrangement of surface atoms, is free of dangling bonds. SiO_2 is the natural insulator used in Si microelectronics, although the technological SiO_2 is amorphous and its surface has a certain density of defects with Si bonds that are not saturated by oxygen nor by another Si atom. Second, the defect density on mica surfaces is usually low; this includes low concentration of steps, as mica is a layered material. Third, atomic C combines easily with oxygen (burning to CO) and with Si (forming SiC). Given this third property, it is intriguing that high quality graphene flakes can be grown on mica, and that large areas of mica may remain unaffected even after prolonged exposure to high flux of carbon atoms (Fig. 8.9).

Three different hypotheses can be formulated [26] to explain this effect (Fig. 8.10):

- A. The surface of mica reflects carbon atoms and graphene grows only at defects.
- B. Carbon ad-atoms diffuse on the surface of mica to nucleation centers (defects) faster than they bind the surface oxygen atoms into volatile CO molecules;
- C. Carbon atoms do reduce the surface (CO is emitted) and then intervene into oxygen vacancies to make bonds with Si, but the substitutional CO produced in this way diffuses to nucleation centers faster than it can escape as CO after reacting with a surface oxygen atom, while the oxygen removed from the surface is efficiently restored by O diffusion from subsurface and/or bulk regions of mica.

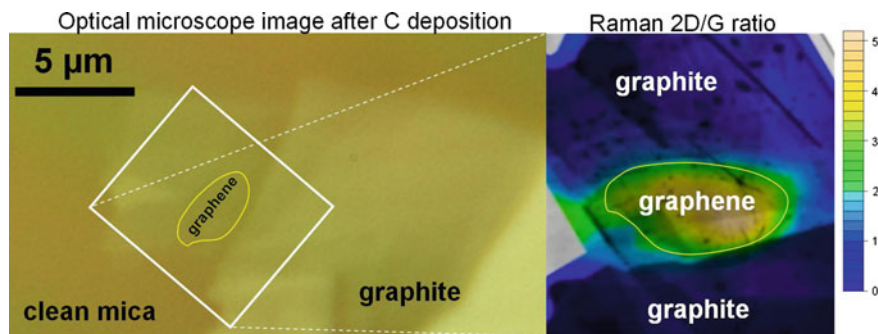


Fig. 8.9 MBE graphene on mica. *Left* Optical image. *Right* AFM overlaid with Raman data

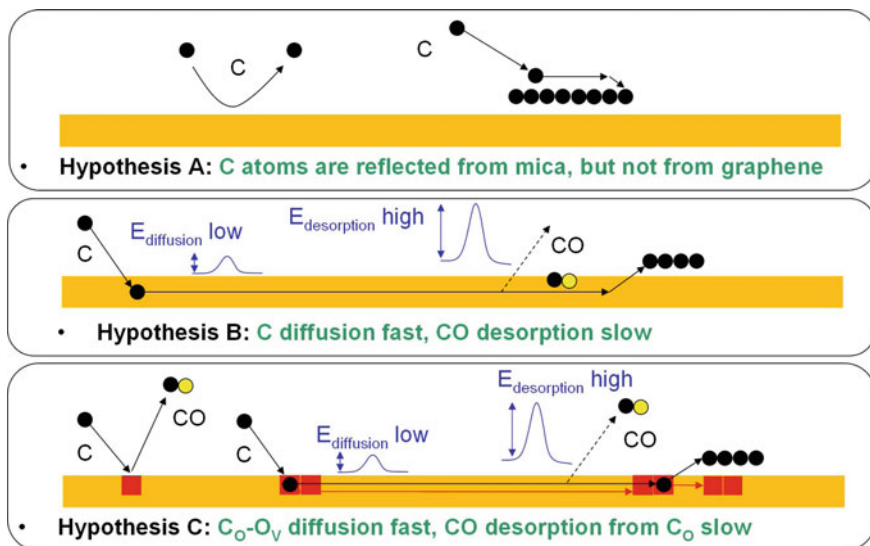
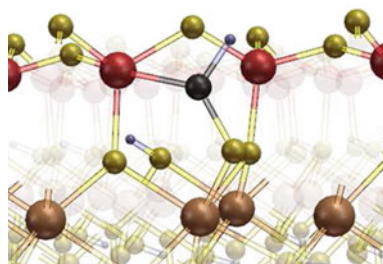


Fig. 8.10 Schematic illustration of hypotheses explaining the formation of graphene on mica



Ground state (C_I)
CO emission: 10 ms

Fig. 8.11 Atomic structure of C on mica, as acquired after ten of ten molecular dynamic runs. C is black, O is yellow, Si is red, Al is brown, H is blue. At the substrate temperatures used in the experiment, 10 ms suffice for this atom to escape as CO

Hypothesis A can be tested by ab initio molecular dynamics: a carbon atom with velocity similar to that of atoms in the MBE beam is placed above the surface of mica or of hexagonal SiO_2 . Its interaction with the surface is then monitored. After a few runs with various initial parameters, one can obtain a reasonable picture of what can and what cannot happen. And it appears that hypothesis A fails (Fig. 8.11).

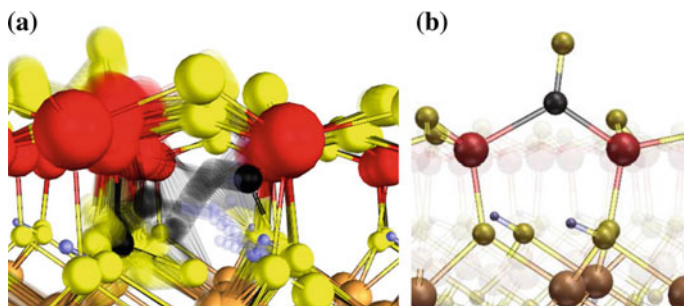


Fig. 8.12 **a** C atom (black) diffusing under the surface of model muscovite mica. **b** Second ground state of adsorbed C, with the energy the same as that in Fig. 8.11 but with only 10 ns of conversion time to CO. O is yellow, Si is red, Al is brown, H is blue

To test hypothesis B, one should therefore search for such diffusion paths, which avoid a close contact between C and twofold-coordinated oxygen. Such a path may indeed exist under the surface (Fig. 8.12a). But the barrier for diffusion turns out to be clearly higher than that for a detour that the diffusing atom can take to the surface and then towards its escape as CO. Moreover, the C atom can reach the alternative ground state (Fig. 8.12b), from which it escapes as CO even faster than from the ground state shown in Fig. 8.11. Thus, hypothesis B fails as well.

Somewhat disappointingly, the remaining hypothesis C also fails. On the one hand, oxygen vacancy diffusion is possible. The vacancy must be positively charged, or the barrier is high. The same is known in SiO₂, and the barriers are similar. But although carbon can diffuse on the surface by a vacancy-mediated mechanism, the diffusion barrier is higher than the barrier for CO emission.

One should therefore re-examine the hypothesis A, keeping in mind that the reflection of the incoming C atoms does not have to be complete, because oxygen vacancy diffusion may help to restore the surface stoichiometry to one that contains enough oxygen to suppress the formation of carbide that would survive the contact with ambient oxygen when the sample is removed from the reaction chamber.

The reason why C is reflected from the surface may be associated with the internal excitation of the C electrons caused by the interaction between the carbon and oxygen orbitals before the bond is formed. Figure 8.13 illustrates this effect. It does not lead to internal excitation during unconstrained DFT molecular dynamics, because such calculations force the system to the ground state. But the energy transfer from the C 2p electrons to mica does require certain time. If this time is shorter than the time after which the internally excited atom would be reflected to vacuum (i.e., shorter than about 60 fs), the atom eventually is indeed reflected back. This is confirmed by DFT calculations with constrained orbital occupation.

Energy transfer from the atom to the substrate is expected to be more efficient when there are partially occupied states that are sufficiently close in energy and in the distance to the excited C 2p orbitals. This means that surface regions which are defected are more sticky to C; graphene and then graphite can grow there. All in

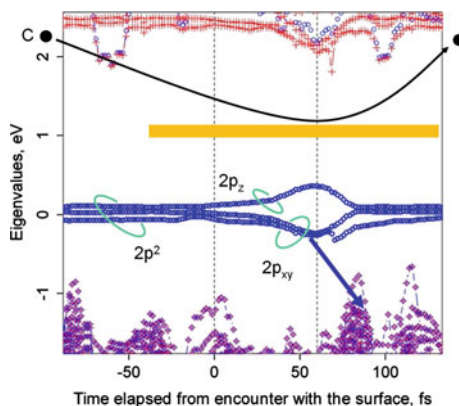


Fig. 8.13 Pauli repulsion effect. Blue C $2p^2$ orbital, initially degenerate and with spin 1. Violet and red valence and conduction states of mica (O orbitals). C p_x and p_y try to make bonds to O, $2p_z$ and $2s^2$ are repelled from VB top. When p_z remains occupied, the C atom turns back. This is confirmed by DFT calculations with constrained orbital occupation: all 10 molecular dynamics runs resulted with reflection back to vacuum

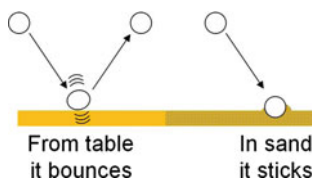


Fig. 8.14 Pingpong ball analogy of the Pauli repulsion effect. The ball bounces from the table because it becomes internally excited. If the excitation energy is efficiently transferred to the target, as to the beach sand, the ball stops

one, the behavior of carbon on the mica surface seems to resemble the behavior of a pingpong ball (Fig. 8.14), which bounces from a table (perfect areas of the surface) but sticks to sand (defected areas of the surface).

Figure 8.15 summarizes the proposed growth mechanism. It takes into account the diffusivity of O vacancies. The seed may be either a dust particle, or damaged surface, or the area where the vacancies cannot become positively charged and for that reason the interaction with C gradually converts the oxide there into carbide.

For a number of reasons, mica cannot be used in CMOS technology. The technological consequence of the above mechanism is thus that the growth of graphene on oxides with twofold-coordinated O, like on SiO_2 , is hard to achieve: the surface density of defects is too high there. Such a high defect density combined with low surface diffusivity of C results in a nanographene-like film that is then difficult to overgrow with graphene crystals of at least μm -size, as needed for the transistors.

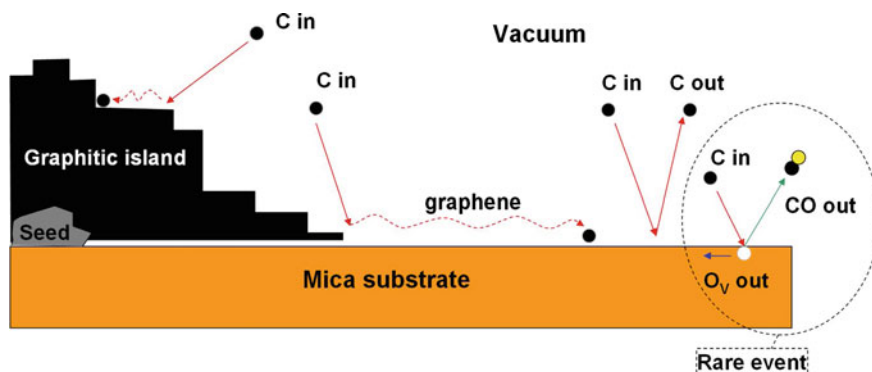


Fig. 8.15 The proposed mechanism of graphene and graphite growth from C atomic beam on mica

8.7 Graphene Base Transistor

In a Field Effect Transistor (FET), the current flowing through the channel formed below the gate oxide is controlled by the potential of the gate electrode on the opposite side of the gate oxide: the channel, the gate oxide, and the gate electrode form a capacitor. In a graphene FET (GFET), the channel is made of graphene, so that the current that is being controlled flows in the plane of the graphene sheet.

The Graphene Base Transistor (GBT) is based on a completely different device concept, akin to that of a bipolar transistor or simply to that of a vacuum tube: the charge carrier flow is controlled by a base electrode placed between the cathode (emitter) and the anode (collector). The carriers flow in the direction normal to the plane of the base electrode, which is graphene in the GBT (Fig. 8.16a).

One of the practical problems of a GFET is that when a semi-metal is used as the channel material, then the transistor cannot be switched off. The gate voltage modulation can only modulate the channel resistance, but the channel remains always conductive. In other words, a GFET is a poor switch, and as such it is not suited for logics: in a microprocessor, transistors in the OFF state should conduct no current, otherwise unacceptable power losses occur and the circuit is not only too expensive in use, but it simply heats up so much that it cannot work at all. The situation in a GBT is totally different: when a suitable emitter-collector bias is applied, the emitter-base bias modulation can switch the state of the transistor between non-conductive OFF and conductive ON (Fig. 8.16a).

Figure 8.16b demonstrates the influence of the graphene sheet on the transmission probability. Transmission probabilities were computed from self-consistent band structures of two unbiased Co electrodes placed at various distances from one another; for the separation of 1.9 nm, one calculation was done with a graphene sheet placed midway between the electrodes. It is clear that insertion of graphene increases the transmission, but the effect is not dramatic: the graphene sheet

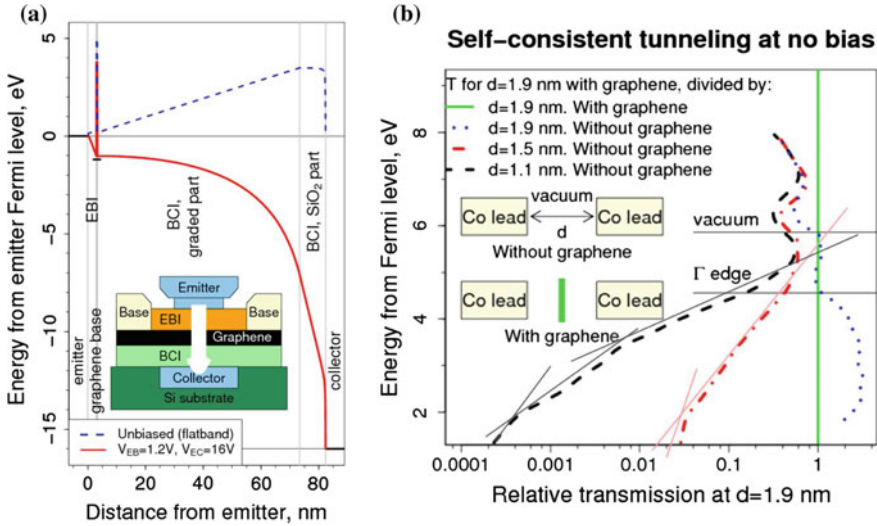


Fig. 8.16 **a** The GBT concept with graded collector insulation [13]; blue line is the unbiased device, red line is the device in the ON state, ECI and BCI are the base-emitter and base-collector insulators. **b** Energy distribution of the relative transmission probability for a graphene sheet placed centrally between two Co electrodes separated by 1.9 nm. Various lines correspond to various reference configurations. Γ edge is the computed position of graphene conduction band bottom at Γ

behaves approximately as a tunneling barrier. The reason is that the electrons traveling in the direction normal to the sheet see the forbidden gap that graphene has at Γ .

This result relies however on the conservation of the lateral momentum of electrons and is therefore strictly valid only if graphene is not a strong perturbation to the symmetry of the system. In most of the practically useful cases, graphene would destroy the lateral symmetry completely, so that efficient channels away from Γ may appear even for electrons traveling along the normal. The only exception may be graphene placed on hBN. Therefore, for most of the layers that separate graphene from the emitter and collector electrodes (Fig. 8.16a) it may be still legitimate to treat graphene as fully (or nearly fully) permeable to emitter-collector electrons.

The calculations summarized in Fig. 8.17 have been done for a device in which the graphene base electrode is placed on a Ge(001) film, whereby graphene was assumed to be a moderately strong barrier for electrons traveling across the sheet. The results indicate that such a device may be capable of operating in the THz range. GBTs built with graphene placed between a Ge(001) wafer and overgrown with a germanium film have been analysed by other researches [27], with similar conclusions: that such devices may be able to operate with terahertz frequencies.

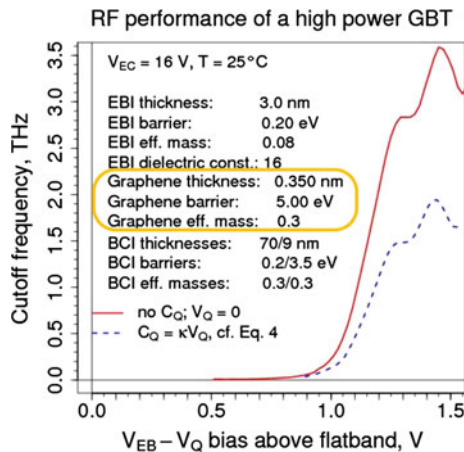


Fig. 8.17 High-frequency performance (cutoff frequency) of the GBT, simulated for a GBT with Ge(001) “insulation” between an erbium germanide emitter and graphene, assuming that graphene is a moderate tunneling barrier for electrons. The *red solid* and the *blue dashed* line correspond to the simulation done with and without taking the quantum capacitance of graphene into account. It follows that the Ge/graphene device should be capable of operating in the THz range. In addition, the device is unique in that it can produce high output power: the design of the base-collector insulator allows for high voltage to be applied between the emitter and the collector

8.8 Carbon on Germanium

As noted in Sect. 8.7, from the design point of view it would be good to have Ge (001) as the substrate on which graphene is grown in the active device region. For this purpose, relatively small graphene flakes suffice: the width of the device may be smaller than a micron, and the whole area may be of the order of a μm^2 . On the other hand, if a good graphene could be grown on germanium in the Si wafer-size scale (that is, with the diameter of 300 mm or more), then some of the problems associated with the graphene transfer to target Si wafer would be solved: in contrast to atoms from the metal substrates used nowadays for graphene growth, Ge atoms does not contaminate the production tools. This would open other technological options for production of CMOS-compatible graphene-based devices.

Graphene can indeed be grown on Ge(001) substrates, including Ge(001) films deposited by a CMOS-compatible technique on Si(001) wafers. Figure 8.18 illustrates this by some experimental data obtained for MBE graphene on samples cut from patterned Ge(001)/Si(001) wafers. Germanium pillars have been embedded into a SiO_2 matrix and then exposed to atomic carbon in an MBE chamber. Raman spectroscopy analysis shows that there is a dramatic difference in the quality of the film obtained on the Ge pillars and on the surrounding SiO_2 . This is exemplified in the upper panel of Fig. 8.18 by the 2D/G Raman mode ratio. The higher the ratio is, the more does the film resemble a single-layer graphene; in principle, the ratio

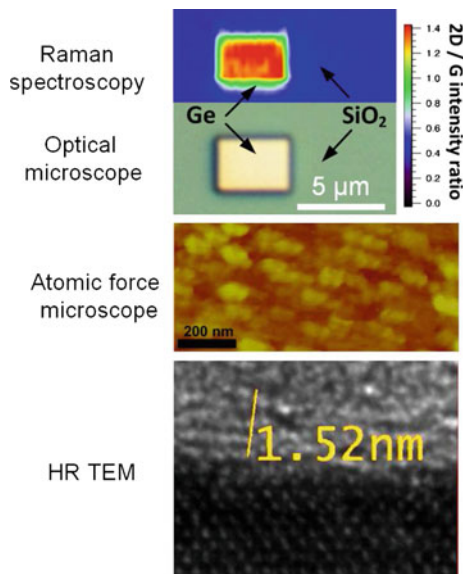


Fig. 8.18 MBE graphene on germanium. From *top* Map of 2D/G Raman mode intensity ratio, and the optical microscope image of this area (Ge pillar surrounded by SiO₂); high ratio (*red*) is seen on Ge and indicates the presence of graphene, while low ratio (*blue*) is seen on SiO₂ and indicates a predominantly sp²-bonded carbon film. Atomic force microscope (AFM) image taken from the region covered by graphene; the growth process resulted in roughening of the surface. High resolution transmission electron microscopy (HR TEM) cross-section of a MBE graphene/Ge(001) sample with a Ge overlayer deposited on top of the graphene after the growth of graphene was complete; several graphene layers on top of crystalline Ge(001) can be recognized

should exceed 1 if the Raman signal originates in such a graphene.⁵ The film obtained on germanium may be described as a single-layer graphene or a few-layer graphene consisting of layers decoupled from one another [10], and the domain size estimated from the Raman spectrum is some tens of nanometers. Scanning electron microscopy images indicate the presence of much larger flakes of micrometer size (Fig. 8.19); in addition, the straight edges of the graphene flakes seen in Fig. 8.19 suggest that the flakes consist of at most few domains of monocrystalline graphene. The difference in the estimated domain size may be associated with the few-layer structure of the graphene, as discussed further on: the flakes have grown on top of the interface layer built of smaller graphene crystallites that is still not fully covered by the flakes of better graphene and remains to be seen as the gray areas in Fig. 8.19.

Here, we are most interested in the reason for the significant difference between the growth of graphene on Ge(001) pillars and on the surrounding SiO₂ planes.

⁵The 2D/G ratio is not the only criterion here, and it is not the absolute one. For more discussion, see [10] (the case of MBE graphene on Ge) and references therein.



Fig. 8.19 Scanning electron microscope (SEM) image of graphene on a Ge pillar such as that in Fig. 8.18. *Black areas* are more conducting than the gray areas and are interpreted as graphene flakes. The *gray areas* may be associated with a more polycrystalline film, possibly with Ge intervening at grain boundaries

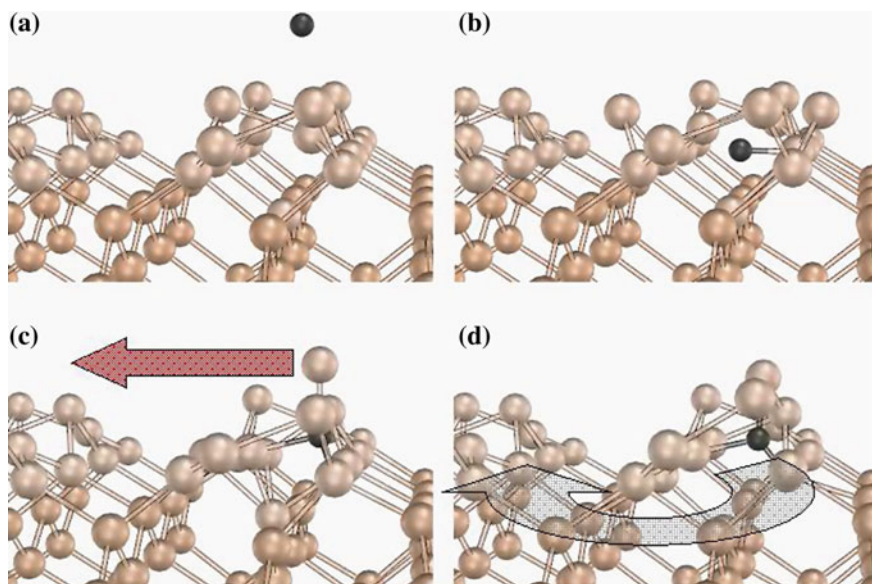


Fig. 8.20 Molecular dynamics simulation of the reaction of C with Ge(001). **a** The incoming atom. **b** A Ge dimer breaks, the C atom lands under the surface. **c** The C atom may substitute a Ge atom from the broken dimer; the other Ge can now diffuse on *top* of the surface. **d** Alternatively, the Ge dimer survives and the C atom diffuses *under* the surface

Figure 8.20 uses the atomic structures obtained during *ab initio* molecular dynamics simulation of the interaction of an incoming C atom with the Ge(001) surface to illustrate two fundamental processes invoked by this interaction. When the C atom hits the surface, about 4.4 eV is released. Part of this energy is received by two Ge dimer atoms, and part is retained by the C atom. The Ge dimer initially breaks, and the C atom finds itself under the surface (Fig. 8.20b). The C atom may now substitute one of the atoms of the broken dimer, transferring the other one into

a Ge ad-atom that diffuses on the surface (Fig. 8.20c). Alternatively, the Ge dimer may survive; in this case, the C atom becomes an Interstitialcy that diffuses under the surface. The probability that it as next substitutes a fourfold-coordinated Ge atom is low (because the barrier for this process is high). The Interstitialcy avoids also going deeper into the bulk of Ge; this is because lot of stress would be generated. Since the GeC phase (in contrast to the SiC phase) is unstable and the C Interstitialcy is mobile, longer deposition times results in the formation of graphene flakes.

Besides the instability of Ge carbide, the key reason for the radically different behavior of C on SiO₂ (and also on Si) than on Ge is therefore good (sub-)surface diffusivity of the C Interstitialcy under (and of small C clusters—on) the Ge(001) surface. This is probably similar to the behavior of the C Interstitialcy under the Si (001) surface; however, in the later case the SiC phase is stable and its formation (presumably happening by the surface Si ejection process similar to the Ge ejection from the Ge–Ge surface dimer) efficiently competes with the formation of graphene.

The ejection of surface Ge by C atoms leads to problems: the ejected Ge atoms attach themselves preferably to the edges of the growing graphene. They may thus remain in the grain boundaries, preventing the formation of larger graphene crystals.

On the other hand, when a small graphene molecule is built by coalescence of interstitialcies arriving from under the surface, it is likely to make as many possible bonds to the substrate (Fig. 8.21a). This parallel configuration is the ground state of the molecule. Tiny C clusters (dimers, trimers, and even hexamer strings) diffuse easily as well—but on top of, not under the surface. Coalescence of C strings may end up in a metastable molecule that is raised (Fig. 8.21b). Such cases are unfavorable for a good graphene, but here the ejected Ge have a positive contribution: they help to “glue” the molecule back to the substrate (Fig. 8.22).

Figure 8.23 contains a summary of the basic processes leading to nucleation and growth of graphene on Ge(001) from atomic carbon. The kick-in process is the reverse process to Ge ejection (or kick-out) by carbon and it produces a complete Ge dimer and a subsurface C Interstitialcy. Note that kick-out may be initiated not only by the incoming C atom, but also by a diffusing C Interstitialcy. For simplicity, the contribution of carbon strings to the growing graphene is not shown in Fig. 8.23.

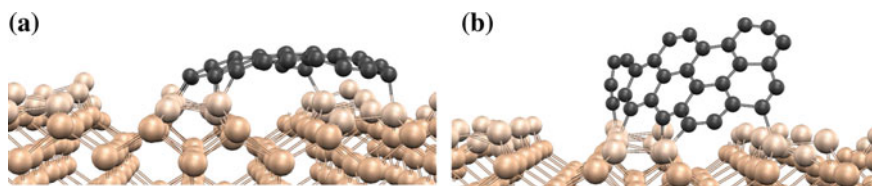


Fig. 8.21 A small graphene molecule **a** lying and **b** standing on Ge(001)

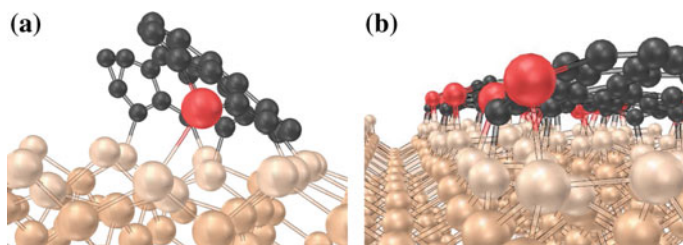


Fig. 8.22 **a** A small graphene molecule, half-standing on and partially glued to Ge(001) by an ejected Ge atom (red). **b** A piece of graphene with Ge atoms attached to its edge

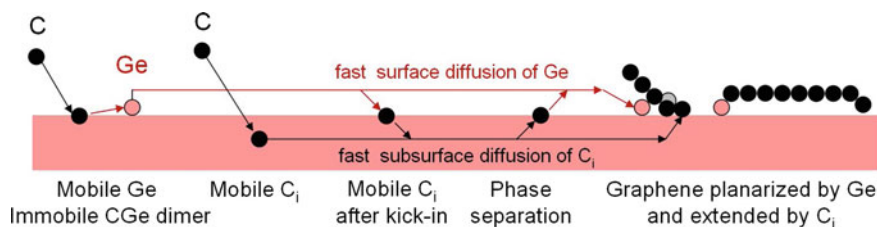


Fig. 8.23 Schematic presentation of the graphene formation on Ge(001) from atomic carbon

The first layer of graphene on Ge(001) is therefore quite poor, presumably due to its nanocrystalline nature stabilized by Ge atoms preventing good coalescence of graphene islands. Polycrystallinity and Ge ejection lead also to the roughening of the surface. Nevertheless, the density of such defects as standing graphene molecules or highly disordered regions of sp^2 carbon can be low enough to allow for this interfacial layer to be overgrown by a much better graphene sheet [10].

Ejection of surface Ge appears to be largely an effect associated with the presence of atomic carbon. When hydrocarbon molecules are supplied instead, the major diffusing species are—according to our *ab initio* calculations—such their fragments as CH_2 or C_2H_3 (depending on the precursor used). Consequently, there seems to be the possibility to suppress the formation of the interface layer in a CVD process; yet up to now, the interface layer might be there (Fig. 8.24, cf. the gray areas).

The problem that is common to the MBE and the CVD growth of graphene on germanium is the high substrate temperature needed to produce good graphene. Significantly better graphene is grown at temperatures close to the melting point of Ge, but at such conditions there is additional, strong surface roughening effect. The major challenge is therefore to reduce the growth temperature down into the range at which the Ge/Si(001) is stable against roughening induced by surface melting.

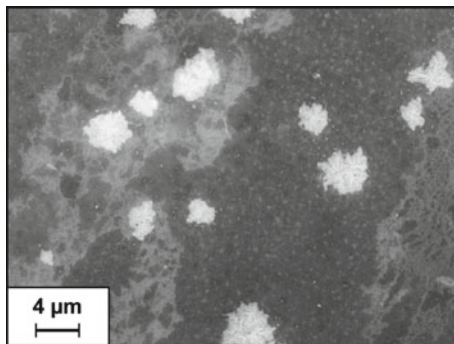


Fig. 8.24 Scanning electron microscope (SEM) image of graphene grown in a cold wall CVD chamber from C_2H_4 . *Black areas* are graphene flakes. This is confirmed by the observation of the Dirac cone in angle-resolved ultraviolet photoelectron spectroscopy (ARUPS) [28]

8.9 Summary and Conclusions

Physical mechanisms involved in the process of graphene growth on several substrates are analyzed on the basis of theoretical (ab initio DFT) and experimental data. The substrates considered here are: graphene, hexagonal boron nitride (hBN), hexagonal SiO_2 and mica, and Ge(001). The focus is on the deposition of atomic carbon, apart from some remarks on CVD from hydrocarbons on Ge(001). A link to the device design and fabrication is provided by sections on the deposition of Si on graphene and on the performance analysis of a graphene base transistor (GBT).

C ad-atoms diffuse fast on graphene and on hBN. But this does not lead to efficient self-nucleation, because the binding of trimers (C_3) to such substrates is practically null. In order to overgrow these materials with graphene, one must provide high enough density of nucleation centers. Ab initio data can be used in a simple formula derived by the random walk theory to estimate the collection zone width w along a linear defect to which carbon sticks; from that one can obtain the sticking coefficient σ of C atoms as $\sigma = w/L$, where L is the average distance between such absorbing defects. Sticking coefficient on high-quality substrates turns out to be low, because w is small, due to desorption of ad-atoms at temperatures of interest.

Chemical interaction of C atoms with perfect graphene or perfect hBN is weak. The opposite is true for SiO_2 -like substrates, including mica, where it leads to CO and SiC formation. In spite of that, atomic C leaves large areas of mica unreacted, and graphene is grown. This is attributable to (a) reflection of many C atoms back to vacuum by a mechanism associated with internal excitation of C 2p electrons before a chemical bond to the *perfect* substrate is formed, and (b) to surface re-oxidation by O vacancy diffusion. This works on mica and should work on a perfect hexagonal SiO_2 . But the growth on SiO_2 and on similar substrates fails, probably

due to too high density of energy-dissipative centers, to which the carbon atoms stick.

Atomic carbon interacts strongly with Ge(001) as well. However, there are two key differences between Ge and SiO₂. First, C atoms and small C clusters diffuse efficiently on Ge(001), so that the crystallites that are already formed can collect the carbon that fell down on the surface far away from them. Second, no stable carbide phase is built by Ge and C; this is in contrast to stability of SiC. The end effect is that Ge can be overgrown with good MBE graphene. The process is not unproblematic, though. The interaction of C atoms with Ge(001) produces mobile Ge atoms, which decorate edges of the growing graphene and may hinder the formation of a wide-area graphene of sufficiently high sheet conductivity. As hydrocarbons do not eject Ge ad-atoms, this may possibly be improved when CVD is used instead of MBE.

Acknowledgments The authors wish to dedicate this work to the memory of Prof. Wolfgang Mehr, the inventor and the pioneer of research on graphene base transistors. We thank our IHP colleagues: Yuji Yamamoto for CVD growth of Ge(001) films on Si wafers, Oksana Fursenko for taking AFM images, Markus Schubert for TEM and EDX characterization, Julia Kitzmann for graphene transfer, Thomas Schroeder for discussions, ideas, and encouragement, and Andre Wolff for technology support. Financial support from the European Commission through the GRADE project (No. 317839) and computing time support from the Jülich Supercomputing Center of the John von Neumann Institute for Computing (project hfo06) are gratefully acknowledged.

References

1. A.K. Geim, Graphene: Status and Prospects. *Science* **324**, 1530–1534 (2009)
2. Y.M. Lin, C. Dimitrakopoulos, K.A. Jenkins, D.B. Farmer, H.Y. Chiu, A. Grill, P. Avouris, 100-GHz transistors from wafer-scale epitaxial graphene. *Science* **327**, 662–662 (2010)
3. F. Schwierz, Graphene transistors. *Nat. Nanotechnol.* **5**, 487–496 (2010)
4. H.C. Kang, H. Karasawa, Y. Miyamoto, H. Handa, T. Suemitsu, M. Suemitsu, T. Otsuji, Epitaxial graphene field-effect transistors on silicon substrates. *Solid State Electron.* **54**, 1010–1014 (2010)
5. J. Kedzierski, P.L. Hsu, P. Healey, P.W. Wyatt, C.L. Keast, M. Sprinkle, C. Berger, W.A. de Heer, Epitaxial graphene transistors on SiC substrates. *IEEE Trans. Electron. Dev.* **55**, 2078–2085 (2008)
6. K.S. Kim, Y. Zhao, H. Jang, S.Y. Lee, J.M. Kim, K.S. Kim, J.H. Ahn, P. Kim, J.Y. Choi, B.H. Hong, Large-scale pattern growth of graphene films for stretchable transparent electrodes. *Nature* **457**, 706–710 (2009)
7. G. Lupina, J. Kitzmann, I. Costina, M. Lukosius, C. Wenger, A. Wolff, S. Vaziri, M. Östling, I. Pasternak, A. Krajewska, W. Strupinski, S. Kataria, A. Gahoi, M.C. Lemme, G. Ruhl, G. Zoth, O. Luxenhofer, W. Mehr, Residual metallic contamination of transferred chemical vapor deposited graphene. *ACS Nano* **9**, 4776–4785 (2015)
8. A. Malesevic, R. Vitchev, K. Schouteden, A. Volodin, L. Zhang, G. Van Tendeloo, A. Vanhulsel, C. Van Haesendonck, Synthesis of few-layer graphene via microwave PECVD. *Nanotechnology* **19**, 305604 (2008)
9. G. Wang, M. Zhang, Y. Zhu, G. Ding, D. Jiang, Q. Guo, S. Liu, X. Xie, P.K. Chu, Z. Di, X. Wang, Direct growth of graphene film on germanium substrate. *Sci. Rep.* **3**, 2465 (2013)

10. G. Lippert, J. Dabrowski, T. Schroeder, M.A. Schubert, Y. Yamamoto, F. Herziger, J. Maultzsch, J. Baringhaus, C. Tegenkamp, M.C. Asensio, J. Avila, G. Lupina, Graphene grown on Ge(001) from atomic source. *Carbon* **75**, 104–112 (2014)
11. G. Lupina, J. Kitzmann, M. Lukosius, J. Dabrowski, A. Wolff, W. Mehr, Deposition of thin silicon layers on transferred large area graphene. *Appl. Phys. Lett.* **103**, 263101 (2013)
12. F. Driussi, P. Palestri, L. Selmi, Modeling, simulation and design of the vertical graphene base transistor. *Microelectron. Eng.* 109:338–341; S. Venica, F. Driussi, P. Palestri, D. Esseni, S. Vaziri, L. Selmi (2014) Simulation of DC and RF performance of the graphene base transistor. *IEEE Trans. Electron Devices* **61**, 2570–2576 (2013)
13. W. Mehr, J. Dabrowski, J.C. Scheytt, G. Lippert, Y.H. Xie, M.C. Lemme, M. Östling, G. Lupina, Vertical graphene base transistor. *IEEE Electron Device Lett.* **33**, 691–693 (2012)
14. S. Vaziri, G. Lupina, C. Henkel, A.D. Smith, M. Östling, J. Dabrowski, G. Lippert, W. Mehr, M.C. Lemme, A Graphene-Based Hot Electron Transistor. *ACS Nano Lett.* 13:1435–1439; Vaziri S, Lupina G, Paussa A, Smith AD, Henkel C, Lippert G, Dabrowski J, Mehr W, Östling M, and Lemme MC (2013) A manufacturable process integration approach for graphene devices. *Solid-State Electronics* **84**, 185–190 (2013)
15. J. Dabrowski, G. Lippert, T. Schroeder, G. Lupina, Role of defects in the process of graphene growth on hexagonal boron nitride from atomic carbon. *Appl. Phys. Lett.* **105**, 191610 (2014)
16. G. Lippert, J. Dabrowski, M.C. Lemme, C. Marcus, O. Seifarth, G. Lupina, Direct graphene growth on insulator. *Phys. Status Solidi B* **248**, 2619–2622 (2011)
17. G. Lippert, J. Dabrowski, Y. Yamamoto, F. Herziger, J. Maultzsch, J. Baringhaus, C. Tegenkamp, M.C. Lemme, W. Mehr, G. Lupina, Molecular beam epitaxy of graphene on mica. *Phys. Status Solidi B* **249**, 2305–2312 (2012)
18. G. Lippert, J. Dabrowski, Y. Yamamoto, F. Herziger, J. Maultzsch, M.C. Lemme, W. Mehr, G. Lupina, Molecular beam growth of micrometer-size graphene on mica. *Carbon* **52**, 40–48 (2013)
19. P. Giannozzi et al., QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Cond. Mat.* **21**, 395502 (2009)
20. J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996)
21. G. Henkelman, B.P. Uberuaga, Hannes Jónsson, A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **113**, 9901–9904 (2000)
22. H. Jónsson, G. Mills, K.W. Jacobsen, Nudged Elastic Band Method for Finding Minimum Energy Paths of Transitions. in *Classical and Quantum Dynamics in Condensed Phase Simulations* eds. by B.J. Berne, G. Ciccotti, D.F. Coker (World Scientific, Singapore, 1998), p. 385
23. C. Ataca, S. Ciraci, Functionalization of BN honeycomb structure by adsorption and substitution of foreign atoms. *Phys. Rev. B* **82**, 165402 (2010)
24. N. Berseneva, A. Gulans, A.V. Krasheninnikov, R.M. Nieminen, Electronic structure of boron nitride sheets doped with carbon from first-principles calculations. *Phys. Rev. B* **87**, 035404 (2013)
25. A. van de Walle, M. Asta, P.W. Voorhees, First-principles calculation of the effect of strain on the diffusion of Ge adatoms on Si and Ge(001) surfaces. *Phys. Rev. B* **67**, 041308 (2013)
26. J. Dabrowski, A. Fleszar, G. Lippert, G. Lupina, Ab initio Modeling of Growth of Grapene for Silicon-Compatible Microelectronics, in *Proceedings of NIC Symposium, 12–13 February 2014*, ed. by K. Binder, G. Mnster, M. Kremer (Jülich, Germany, Forschungszentrum Jülich, Jülich, 2014), p. 207
27. G. Lippert, J. Dabrowski, W. Mehr, G. Baccarani, A. Gnudi, V. Di Lecce (2015) Low band gap graphene transistor. European patent application 14194150.0
28. M.C. Asensio, J. Avila, J. Baringhaus, I. Colambo, J. Dabrowski, G. Lippert, G. Lupina (2015) unpublished

Chapter 9

Recent Progress on Nonlocal Graphene/Surface Plasmons

Norman J.M. Horing, A. Iurov, G. Gumbs, A. Politano
and G. Chiarello

Abstract We review recent experimental and theoretical studies of the non-local plasmon dispersion relations of both single and double layers of graphene which are Coulomb-coupled to a thick conducting medium. High-resolution electron energy loss spectroscopy (HREELS) was employed in the investigations. A mean-field theory (R.P.A.) formulation was used to simulate and explain the experimental results, with the undamped plasmon excitation spectrum calculated for arbitrary wave number. Our numerical calculations show that when the separation a between a graphene layer and the surface is less than a critical value $a_c = 0.4k_F^{-1}$, the lower acoustic plasmon is overdamped. This result seems to explain the experimentally observed behavior for the plasmon mode intensity as a function of wave vector. The damping, as well as the critical distance, changes in the presence of an energy bandgap for graphene. We also report similar damping features of the plasmon modes for a pair of graphene layers. However, the main difference arising in the case when there are two layers is that if the separation between the layer nearest the surface and the surface is less than a_c , then both the symmetric and antisymmetric modes become damped, in different ranges of wave vector.

N.J.M. Horing (✉)

Department of Physics, Stevens Institute of Technology, Hoboken, NJ 07030, USA
e-mail: nhoring@stevens.edu

A. Iurov

Center for High Technology Materials, University of New Mexico,
Albuquerque, NM 87106, USA

G. Gumbs

Department of Physics, Hunter College, Cuny, NY 01065, USA

G. Gumbs

Donostia International Physics Center (DIPC), P de Manuel Lardizabal, 4,
20018 San Sebastian, Basque Country, Spain

A. Politano · G. Chiarello

Dipartimento di Fisica, Università degli Studi della Calabria, 87036 Rende (Cs), Italy

9.1 Introduction

This review article is concerned with plasma phenomenology in graphene [1–8]; both nonlocal as well as local features are discussed, including wavevector shifting of plasmon modes and Landau damping. It is hard to overstate the importance of graphene (and similar materials, such as silicene [9]). As an ultrathin layer of carbon/graphite just a single atom thick in a planar hexagonal honeycombed array, graphene has an energy spectrum proportional to linear momentum that is characteristic of relativistic Dirac electrons with no gap between the conduction and valence bands and, correspondingly, no mass in its native state. It has been the subject of truly extraordinary research efforts because of its great potential to bring about a massive change in the electronic device and computer industries based on its exceptional “device-friendly” electrical conductivity and sensing properties. The electrical mobility in graphene reaches up to $200,000 \text{ cm}^2 \text{ Vs}^{-1}$, over two orders of magnitude greater than that of present silicon-based materials, over 20 times that of gallium arsenide and over twice that of indium antimonide. It has a long mean-free-path ($l \sim 400 \text{ nm}$ at room temperature) and it is stable to 3000 K, as well as having great strength. Its planar form is consonant with well-developed present device fabrication techniques. Graphene-based chemical sensors can detect one part per billion of various active gases with high sensitivity based on adsorbed molecules acting as donors/acceptors, jointly with graphene’s high conductivity and low noise. These, and other features of graphene, have commanded the attention of the scientific/engineering world, generating a massive flood of research activity.

Graphene’s electromagnetic/optical response features are also important and its dielectric and plasmonic properties (local and nonlocal) play a major role in this. As indicated above, this subject will be discussed in detail here in two important experimental configurations: (1) free-standing graphene, and (2) graphene Coulomb-coupled with a slab of substrate plasma. The pertinent equation for the fundamental inverse dielectric “screening” function of the latter involves the corresponding function for the slab of substrate plasma alone: we start the discussion with the theoretical background for that since it may be expected to play an important role in further future studies. The slab plasma discussion is followed by a review of the present graphene plasma literature and recent theoretical results. Finally, and most importantly, there is a thorough discussion of the subject from an experimental point of view.

9.2 Nonlocal Dielectric Response of a Slab Plasma

Plasma nonlocality has been a subject of great interest in both classical and quantum physics. It is important in dynamic plasma screening phenomenology, including the plasmon spectrum, intensity and natural damping, as well as in the static shielding of charge interactions. In low dimensional systems, the boundaries involved can significantly affect such features of plasma dynamics, including all aspects of nonlocal

behavior. Moreover, the associated geometric constraints have become ever more important as semiconductor device sizes have diminished, now at the nanometer level and even entering the regime of a “single atom thick” layer, such as graphene and silicene. Of course, “thick” and semi-infinite plasma layers also involve nonlocal plasmas, and their dielectric response properties will be discussed here on both a stand-alone basis, as well as in the role of a substrate interacting with a thin two-dimensional sheet of graphene. This review will address both the theoretical and experimental aspects of such a nonlocal system starting with the analysis of the dynamic, nonlocal inverse dielectric response function $K(1, 2)$ of a finite slab of plasma in the random phase approximation (RPA) ($1 = \mathbf{r}_1, t_1$; $2 = \mathbf{r}_2, t_2$). This function is the space-time matrix inverse of the direct dielectric function $\epsilon(1, 2)$, such that the effective potential $V(1)$ generated by an applied potential $U(2)$ is given by [10–14].

$$V(1) = \int d2 K(1, 2) U(2) \quad \text{or} \quad K(1, 2) = \delta V(1) / \delta U(2), \quad (9.1)$$

for a linear medium. Since $V(1)$ is developed by the density perturbation $\rho(1)$ due to polarization of the plasma (with interparticle coulomb potential v) by $U(2)$, it is given by

$$V(1) = U(1) + \int d3 v(1 - 3) \rho(3), \quad (9.2)$$

leading to the integral equation ($\delta(1 - 2)$ is a space-time Dirac delta function):

$$\begin{aligned} K(1, 2) &= \frac{\delta V(1)}{\delta U(2)} = \delta(1 - 2) + \int d3 v(1 - 3) \frac{\delta \rho(3)}{\delta U(2)} \\ &= \delta(1 - 2) + \int d4 \int d3 v(1 - 3) \frac{\delta \rho(3)}{\delta V(4)} \frac{\delta V(4)}{\delta U(2)}, \end{aligned} \quad (9.3)$$

or

$$K(1, 2) = \delta(1 - 2) - \int d4 \alpha(1, 4) K(4, 2) \quad (9.4)$$

where the polarizability $\alpha(1, 4)$ is defined as

$$\alpha(1, 4) = - \int d3 v(1 - 3) \frac{\delta \rho(3)}{\delta V(4)} \quad (9.5)$$

and

$$R(3, 4) = \frac{\delta \rho(3)}{\delta V(4)} = i G_1(1, 2) G_1(2, 1^+) \quad (9.6)$$

is the “ring diagram” of the density perturbation response function R , expressed in terms of noninteracting one-electron thermodynamic Green’s functions G_1 within the random phase approximation (RPA) [15]. Such RPA results for the density perturbation response function and the effective potential, as well as the inverse dielectric function, for three-, two- and one-dimensional plasmas based on the use of noninteracting one-electron Green’s functions in (9.6) are very well known. Our focus here will be on a *thick* plasma slab (including the semi-infinite plasma limit) in which it is embedded in a substrate in the vicinity of a thin two-dimensional plasma such as graphene or silicene [9].

Of course, plasma boundaries have been the subject of intensive investigation in the past. Older reviews, articles and books by Ritchie [16], Raether [17, 18], Feibelman [19], Liebsch [20], Kushwaha [21] and others summarize a great deal of work on the role of boundaries and nonlocality in thick bounded plasmas. A particularly informative study was carried out by Newns [22], which was further elaborated by Horing [23] in the determination of antisymmetric (A) and symmetric (S) effective potentials $V^{(A,S)}$ generated by corresponding applied potentials $U^{(A,S)}$ associated with sources $\nabla_2^2 U(2) = 4\pi S(2)$ [note that the convention $\nabla_2^2 U(2) \rightarrow 4\pi S(2)$ employed here differs from the usual one in the sign of $S(2)$]. In this analysis for infinite-barrier-model boundaries, the effective potentials $V^{(A,S)}$ (as well as $S^{(A,S)}$) of a slab of plasma in the region $0 \leq z \leq d/2$ are continued antisymmetrically (A) or symmetrically (S) into the adjoining region $d/2 < z < d$ across the midpoint $z = d/2$; and it yields results in a mixed $(z, \bar{Q})$ and ω representation given as follows for the regions $z \leq 0$ and $0 \leq z \leq d$ [note that translational invariance in the lateral (x, y) plane permits representation in terms of a single conjugate lateral wavevector $\bar{R} = (x_1 - x_2, y_1 - y_2) \rightarrow \bar{Q} = (Q_x, Q_y)$]:

$$\begin{aligned} V_Q^{(A,S)}(z, \omega) = & \eta_+(-z) \left\{ e^{Qz} \left[V_Q^{(A,S)}(0, \omega) - U_Q^{I(A,S)}(0, \omega) \right] + U_Q^{I(A,S)}(z, \omega) \right\} \\ & + \eta_+(z) \eta_+(d-z) \left[\left(-\frac{4}{d} \right) \sum_{q^{(A,S)}} \sum_{q'^{(A,S)}} \cos(q^{(A,S)} z) E_{Qqq'}^{-1} \left[2\pi S_{Qq'}^{(A,S)}(\omega) \right. \right. \\ & \left. \left. + QV_Q^{(A,S)}(0, \omega) - 2QU_Q^{I(A,S)}(0, \omega) \right] \right], \end{aligned} \quad (9.7)$$

where $\eta_+(z)$ is the Heaviside unit step function and

$$U_Q^{I(A,S)}(z, z', \omega) = -\frac{2\pi}{Q} \int_{-\infty}^0 dz' e^{-Q|z-z'|} S_Q^{(A,S)}(z', \omega) \quad \text{for } z < 0, \quad (9.8)$$

represents the potential in the region where $z < 0$ due to that part of the source charge $S_Q^{(A,S)}(z'; \omega)$ lying within $z < 0$ and

$$V_Q^{(A,S)}(0, \omega) = \left(1 + \varepsilon_Q^{(A,S)}\right)^{-1} \left(2U_Q^{I(A,S)}(0, \omega) - \frac{8\pi}{d} \varepsilon_Q^{(A,S)} \sum_{q^{(A,S)}} \sum_{q'^{(A,S)}} E_{Qqq'}^{-1} S_{Qq'}^{(A,S)}(\omega)\right) \quad (9.9)$$

and

$$\varepsilon_Q^{(A,S)} \equiv \varepsilon_Q^{(A,S)}(\omega) = \left(\frac{4Q}{d} \sum_{q^{(A,S)}} \sum_{q'^{(A,S)}} E_{Qqq'}^{-1}\right)^{-1}. \quad (9.10)$$

In these equations above we have

$$E_{Qqq'} \equiv E_{Qqq'}(\omega) = (Q^2 + q^2) \delta_{qq'} / \eta_q + 4\pi R_{Qqq'} \quad (9.11)$$

$$= 4\pi (\Delta_{Qq} \delta_{qq'} / \eta_q - A_{Qqq'}), \quad (9.12)$$

with

$$R_{Qqq'} = D_{Qq} \delta_{qq'} / \eta_q - A_{Qqq'} \quad (9.13)$$

as the density-perturbation response function $R(1, 2) = \delta\rho(1)/\delta V(2)$ in (Q, q, q') -representation having a diagonal part D and a non-diagonal part $-A$. In this, $\Delta_{Qq} = (4\pi)^{-1}(Q^2 + q^2) + D_{Qq}$. The notations (A, S) refer to antisymmetric or symmetric potentials and sources, and for the antisymmetric case, $q^{(A)}, q'^{(A)} = (2n+1)\pi/d, n=0, 1, 2, \dots, \infty$ and $\eta_q \equiv 1$, whereas for the symmetric case, $q^{(S)}, q'^{(S)} = 2n\pi/d, n=0, 1, 2, 3, \dots, \infty$ and $\eta_q = 1$ for $q > 0$, but $\eta_q = \frac{1}{2}$ for $q = 0$. Both of the Fourier series sums over $q^{(A)} = (2n+1)\frac{\pi}{d}$ and over $q^{(S)} = \frac{2n\pi}{d}$ represent the same function in the half interval $[0 \leq z < \frac{d}{2}]$, but in the other half interval $[\frac{d}{2} < z \leq d]$, they represent antisymmetric or symmetric continuations, respectively, across $z = \frac{d}{2}$. The Fourier series transforms employed here are defined as $[\mathbf{r} = (\bar{R}, z)$ and $\mathbf{q} = (\bar{Q}, q)]:$

$$f^{(A,S)}(\bar{R}, z) = \frac{2}{d} \sum_{q^{(A,S)}} \eta_q \int \frac{d^2 \bar{Q}}{(2\pi)^2} e^{i\bar{Q} \cdot \bar{R}} \cos(q^{(A,S)} z) f_{Qq}^{(A,S)}; \quad (9.14)$$

and $f_{Qq}^{(A,S)} = \int_0^d dz \int d^2 \bar{R} e^{i\bar{Q} \cdot \bar{R}} \cos(q^{(A,S)} z) f^{(A,S)}(\bar{R}, z).$

The diagonal part $(D_{Qq} \delta_{qq'} / \eta_q)$ of the density perturbation response function (R) corresponds to the familiar noninteracting Fermi gas result subject to periodic boundary conditions, whereas the non-diagonal part (A_{Qq}) arises from boundary

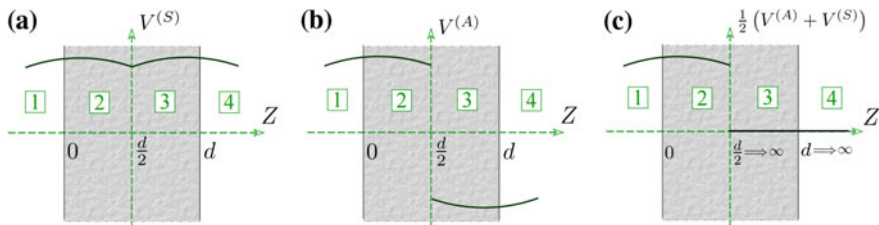


Fig. 9.1 Slab plasma in region II where $z = [0, d/2)$ with potential V continued across $z = d/2$ into region $z = [d/2, d]$: **a** $V^{(S)}$ symmetrically continued potential; **b** $V^{(A)}$ antisymmetrically continued potential; **c** V average of the symmetric potential $V^{(S)}$ and antisymmetric potential $V^{(A)}$ with equal weight as $d \rightarrow \infty$ ($d/2 \rightarrow \infty$)

conditions forcing the electron Green's functions G_1 to vanish at the boundaries (having effects localized within small spatial regions limited to an inverse Fermi wavenumber near the boundaries), for the infinite barrier model considered here.

While each of the two Fourier series for $V_Q^{(A)}(z, \omega)$ and $V_Q^{(S)}(z, \omega)$ individually properly represent the effective potential in the slab region $0 \leq z \leq d/2$, as generated by corresponding applied potentials $U^{(A)}$ and $U^{(S)}$ and associated sources, it is necessary to average them to eliminate the appearance of spurious sources in the continuation region $d/2 < z < d$, which is distant for large d (Fig. 9.1). Thus, we have

$$V_Q(z, \omega) = \frac{1}{2} \left[V_Q^{(A)}(z, \omega) + V_Q^{(S)}(z, \omega) \right] = \frac{1}{2} \sum_{A+S} V_Q^{(A,S)}(z, \omega) \text{ for large } d \quad (9.15)$$

and

$$U_Q(z, \omega) = \frac{1}{2} \sum_{A+S} U_Q^{(A,S)}(z, \omega). \quad (9.16)$$

It is at once apparent that (9.1) yields $K(1, 1')$ directly upon choosing the Dirac delta function $U(2) = \delta(2 - 1')$ as a model potential,

$$V_{1'}(1) = \int d(2)K(1, 2)U(2) = \int d(2)K(1, 2)\delta(2 - 1') = K(1, 1'). \quad (9.17)$$

Using this, one can describe $K(1, 1')$ for a thick slab of plasma with planar boundaries by using (9.7) for $V^{(A,S)}(1)$ jointly with (9.15), generated by the mathematical model potential $U(2) = \delta(2 - 1')$. These considerations yield $V_Q(z, \omega)$ for a thick slab as (define $\tilde{E}_{Qqq'} = \eta_q E_{Qqq'}$)

$$\begin{aligned}
V_Q(z, \omega) = & \frac{1}{2} \eta_+(-z) \left[e^{Qz} \left(\sum_{A+S} V_Q^{(A,S)}(0, \omega) - 2U_Q^{I(A,S)}(0, \omega) \right) + 2U_Q^{I(A,S)}(z, \omega) \right] \\
& - \frac{1}{2} \eta_+(z) \left[\sum_{A+S} \left(\frac{4}{d} \right) \sum_{q^{(A,S)}} \sum_{q'(A,S)} \cos(q^{(A,S)} z) \tilde{E}_{Qqq'}^{-1} \left[2\pi S_{Qq'}^{(A,S)}(\omega) \right. \right. \\
& \left. \left. + QV_Q^{(A,S)}(0, \omega) - 2QU_Q^{I(A,S)}(0, \omega) \right] \right],
\end{aligned} \tag{9.18}$$

where terms of order $\frac{1}{d}$ are neglected (but could be important in a calculation of surface energy), and

$$V_Q^{(A,S)}(0, \omega) = \left(1 + \varepsilon_Q^{(A,S)} \right)^{-1} \left(2U_Q^{I(A,S)}(0, \omega) - \frac{8\pi}{d} \varepsilon_Q^{(A,S)} \sum_{q^{(A,S)}} \sum_{q'(A,S)} \tilde{E}_{Qqq'}^{-1} S_{Qq'}^{(A,S)}(\omega) \right) \tag{9.19}$$

and

$$\left(\varepsilon_Q^{(A,S)} \right)^{-1} = \frac{4Q}{d} \sum_{q^{(A,S)}} \sum_{q'(A,S)} \tilde{E}_{Qqq'}^{-1}. \tag{9.20}$$

The resulting solution for the inverse dielectric function $K(1, 1') = V_{1'}(1)$ of a thick slab, neglecting terms of order $O(1/d)$, may be written in a mixed representation as

$$\begin{aligned}
K(Q, z, z'; \omega) = & \frac{1}{2} \eta_+(-z) \left[e^{Qz} \left(\sum_{A+S} V_Q^{(A,S)}(0, \omega) - Q^{-1} \delta'(z') - \delta(z') \right) \right. \\
& \left. + Q^{-1} \delta'(z') e^{-Q|z|} + \delta(z') \Theta(z) e^{-Q|z|} + 2\eta_+(-z') \delta(z - z') \right] \\
& - \frac{1}{2} \eta_+(z) \sum_{A+S} \left(\frac{4}{d} \right) \sum_{q^{(A,S)}} \sum_{q'(A,S)} \cos(q^{(A,S)} z) \tilde{E}_{Qqq'}^{-1} \left[QV_Q^{(A,S)}(0, \omega) \right. \\
& \left. - (q'^2 + Q^2) \cos(q'(A, S)z') \eta_+(z') - Q\delta(z') \right]
\end{aligned} \tag{9.21}$$

where $\Theta(z) \equiv \eta_+(z) - \eta + (-z)$ and $\delta'(z) \equiv d(\delta(z))/dz$, with

$$V_Q^{(A,S)}(0, \omega) = (1 + \varepsilon_Q^{(A,S)})^{-1} \left\{ Q^{-1} \delta'(z') + \delta(z') - \frac{8\pi}{d} \varepsilon_Q^{(A,S)} \sum_{q^{(A,S)}} \sum_{q'^{(A,S)}} \tilde{E}_{Qqq'}^{-1} \right. \\ \left. \times \left[\frac{\delta'(z')}{2\pi} - \left(\frac{q'^2 + Q^2}{2\pi} \right) \cos(q'^{(A,S)} z') \eta_+(z') \right] \right\}. \quad (9.22)$$

Equations (9.21, 9.22, 9.17) above include the role of non-diagonal density response matrix elements, $-A_{Qqq'}$ as well as the diagonal ones, $D_{Qqq'} \delta_{qq'}/\eta_q$.

Considering the semi-infinite limit $d \rightarrow \infty$, it is reasonable to neglect the nondiagonal elements $-A_{Qqq'}$ as well as terms of order $O(1/d)$, and passing to the continuum limit $\sum_{q^{(A,S)}} \rightarrow \frac{d}{2\pi} \int_0^\infty dq$ (also $\Sigma_{A+S} \rightarrow 2$), we have

$$K(\overline{Q}, z, z'; \omega) = \eta_+(-z) \left(\delta(z - z') - \frac{\varepsilon_Q}{1 + \varepsilon_Q} e^{Qz} \delta(z') \right. \\ \left. + \frac{2\varepsilon_Q}{1 + \varepsilon_Q} K_\infty^{3D}(\overline{Q}, z'; \omega) e^{Qz} \eta_+(z') \right) \\ + \eta_+(z) \left[v_\infty^{3D}(\overline{Q}, z; \omega) \left(\frac{Q\varepsilon_Q}{1 + \varepsilon_Q} \delta(z') - \frac{2Q\varepsilon_Q}{1 + \varepsilon_Q} K_\infty^{3D}(\overline{Q}, z'; \omega) \eta_+(z') \right) \right. \\ \left. + [K_\infty^{3D}(\overline{Q}, z + z'; \omega) + K_\infty^{3D}(\overline{Q}, z - z'; \omega)] \eta_+(z') \right], \quad (9.23)$$

where

$$(\varepsilon_Q^{(A,S)})^{-1} = \varepsilon_Q^{-1} = \frac{2Q}{\pi} \int_0^\infty dq [(Q^2 + q^2) \varepsilon_\infty^{3D}(\mathbf{q}, \omega)]^{-1}, \quad (9.24)$$

and $\varepsilon_\infty^{3D}(\mathbf{q}, \omega)$ (3D denotes three dimensional) is the dynamic, nonlocal bulk dielectric function of the thick-slab material. Furthermore, (9.23) involves the definitions

$$K_\infty^{3D}(\overline{Q}, z'; \omega) = \pi^{-1} \int_0^\infty dq \cos(qz') / \varepsilon_\infty^{3D}(\mathbf{q}, \omega) \quad (9.25)$$

and

$$v_\infty^{3D}(\overline{Q}, z'; \omega) = 2\pi^{-1} \int_0^\infty dq \cos(qz') / [(q^2 + Q^2) \varepsilon_\infty^{3D}(\mathbf{q}, \omega)]. \quad (9.26)$$

A fuller account of the derivation of these results is provided in [21].

If the bulk plasma within the semi-infinite slab is fully local in the sense that $\varepsilon_{\infty}^{3D}(\mathbf{q}, \omega) \rightarrow \varepsilon_{\infty}^{3D}(\omega) \equiv \varepsilon(\omega)$, then the corresponding local inverse dielectric function is given by

$$K(\overline{Q}, z, z'; \omega) = \eta_+(-z) \left[\delta(z - z') + \delta(z') e^{Qz} \left(\frac{1 - \varepsilon(\omega)}{1 + \varepsilon(\omega)} \right) \right] \\ + \eta_+(z) \left[\frac{\delta(z - z')}{\varepsilon(\omega)} + \delta(z') e^{-Qz} \frac{1}{\varepsilon(\omega)} \left(\frac{\varepsilon(\omega) - 1}{\varepsilon(\omega) + 1} \right) \right]. \quad (9.27)$$

While (9.27) is useful because of its relative simplicity, it is important to understand that its validity is restricted because the q -integrations in (9.24–9.26) extend over an infinite integration range. This admixes the effects of the boundary/image length scale with that of the q -nonlocality dependence, eliminating the possibility of an unqualified limit $q \rightarrow 0$ and modifying the plasmon dependence on Q : Furthermore, the imaginary part of $\varepsilon_{\infty}^{3D}(\mathbf{q}, \omega)$ participates in these q -integrations, imparting a damping contribution to the surface plasmon roots even in the low- Q limit (which is not represented in (9.27)). In fact, the “nonlocal” linear Q -correction to the surface plasmon, and its imaginary part involving damping, depend sensitively on the nature of the bounding surface. Nevertheless, there is a regime in which (9.27) is meaningful. This can be seen by examining the parameter measuring the importance of nonlocality in the bulk dielectric function $\varepsilon_{\infty}^{3D}(\mathbf{q}, \omega)$, namely $p_c^M \sim \left(m\omega_p^2/E_{\text{thermal}} \right)^{1/2}$, where the characteristic thermal energy $E_{\text{thermal}} = \mu$ is the Fermi energy in the degenerate case at hand. For $q \ll p_c^M$ it is reasonable to neglect nonlocality, at least in the surface plasmon frequency $\omega_s = \omega_p/\sqrt{2}$ as well as in comparison with other more pertinent sources of nonlocal behavior (but it cannot be neglected in the damping of the surface plasmon). For Cu, $\omega_p \sim 1.66 \times 10^{16} \text{ s}^{-1}$, we have the critical nonlocality wavenumber parameter as $p_c^M \sim 10^{10} \text{ m}^{-1}$. By the way of comparison, the corresponding critical wavenumber for massless graphene, measured by $p_c^G \sim p_F E_F / \hbar \omega_p$ with $\omega_p = \sqrt{(3.22 \times 10^{19}) p_c^G}$, is given by $p_c^G \sim 4 \times 10^{-2} \text{ m}^{-1}$. Thus, while nonlocality is important in graphene for wavenumbers in the vicinity of (and above) p_c^G , it is relatively unimportant in surface plasmon phenomenology up to $p_c^M \sim 10^{10} \text{ m}^{-1}$. The high wavenumber range between p_c^G and p_c^M will be discussed in detail here taking account of graphene nonlocality, but not that of the metal. Initially, we review the fully local range for wavenumbers well below both p_c^G and p_c^M (which, of course, excludes intrinsically nonlocal damping).

Note that henceforth the symbols Q and $q_{\parallel}$ for lateral wavevector are used interchangeably, $Q \equiv q_{\parallel}$.

9.3 2D Plasma Coulomb-Coupled with a Slab Plasma

To address the problem of a two dimensional sheet of plasma interacting with a slab of substrate plasma, it is reasonable to expect that the perturbed densities of the two are mutually independent and additive, $\rho = \rho_{Sl} + \rho_{2D}$, to the extent that tunneling between the two plasmonic systems is negligible. Accordingly, the polarizabilities are additive and the inverse dielectric function K for the composite system is given by (subscripts “Sl” and “2D” refer to the slab plasma and 2D sheet, respectively)

$$K^{-1} = \varepsilon = 1 + \alpha_{Sl} + \alpha_{2D} \equiv \varepsilon_{Sl} + \alpha_{2D} = K_{Sl}^{-1} + \alpha_{2D} \quad (9.28)$$

or

$$K = K_{Sl} - K_{Sl} \cdot \alpha_{2D} \cdot K, \quad (9.29)$$

where K_{Sl} refers to the inverse dielectric function of the slab plasma alone, as discussed above. This integral equation is readily solved in position representation (bear in mind that the symbol for lateral wavenumber, formerly Q above, will now be $q_{\parallel}$ [$Q \equiv q_{\parallel}$], which we suppress along with frequency ω):

$$K(z_1, z_2) = K_{Sl}(z_1, z_2) - \int_{-\infty}^{\infty} dz' \int_{-\infty}^{\infty} dz'' K_{Sl}(z_1, z') \alpha_{2D}(z', z'') K(z'', z_2). \quad (9.30)$$

Simplification occurs because localization of the 2D sheet to the plane at $z = a$ means that

$$\alpha_{2D}(z', z'') = \int_{-\infty}^{\infty} dz''' v(z', z''') D(z''', z'') \quad (9.31)$$

where (v is the Coulomb potential)

$$D(z''', z'') = \Pi_{2D}^{(0)}(q_{\parallel}, \omega) \delta(z''' - a) \delta(z'' - a) \equiv -R_{2D}^{(0)}(q_{\parallel}, \omega) \delta(z''' - a) \delta(z'' - a), \quad (9.32)$$

with $\Pi_{2D}^{(0)}(q_{\parallel}, \omega) \equiv -R_{2D}^{(0)}(q_{\parallel}, \omega)$ as the 2D ring diagram of the RPA for the 2D plasma sheet. This yields

$$K(z_1, z_2) = K_{Sl}(z_1, z_2) - \Pi_{2D}^{(0)}(q_{\parallel}, \omega) \int_{-\infty}^{\infty} dz' K_{Sl}(z_1, z') v(z' - a) K(a, z_2) \quad (9.33)$$

and the solution for $K(a, z_2)$ may be obtained algebraically as

$$K(a, z_2) = \frac{K_{SI}(a, z_2)}{S_C(q_{||}, \omega)}, \quad (9.34)$$

where the “inverse dispersion factor” $S_C(q_{||}, \omega)$, whose zeros determine the coupled plasmon frequencies (and amplitudes) of the composite system is given by

$$S_C(q_{||}, \omega) \equiv 1 + \Pi_{2D}^{(0)}(q_{||}, \omega) \left\{ \int_{-\infty}^{\infty} dz' K_{SI}(a, z') v(z' - a) \right\}. \quad (9.35)$$

The resulting inverse dielectric function for the composite system is given in mixed z - $q_{||}$ - ω representation by [12, 14, 24]

$$K(z_1, z_2) = K_{SI}(z_1, z_2) - \Pi_{2D}^{(0)}(q_{||}, \omega) \frac{K_{SI}(a, z_2)}{S_C(q_{||}, \omega)} \left\{ \int_{-\infty}^{\infty} dz' K_{SI}(z_1, z') v(z' - a) \right\}. \quad (9.36)$$

The fully local composite system result for $K(z_1, z_2)$ of a semi-infinite substrate plasma coupled with a 2D sheet plasma is obtained by employing K_{SI} from (9.25) in (9.35) and (9.36). Correspondingly, the interacting plasmon roots arise from the vanishing of $S_C(q_{||}, \omega)$ of (9.34) using the local limit of $\Pi_{2D}^{(0)}(q_{||}, \omega)$ given as [25]

$$\Pi_{2D}^{(0)}(q_{||}, \omega) \approx C q_{||}^2 / \omega^2 \quad \text{with} \quad C = \frac{2\mu}{\pi \hbar^2} \left\{ 1 - \frac{\Delta^2}{\mu^2} \right\}, \quad (9.37)$$

which accomodates both gapless as well as gapped graphene, with Δ as the gap between the valence and conduction bands. The fully local plasmon modes follow as (ϵ_s is the background dielectric constant)

$$\omega^2 = \frac{\pi e^2 C}{\epsilon_s} q_{||} + \left(\frac{\omega_p}{2} \right)^2 \pm \sqrt{\frac{\pi e^2 C \omega_p^2}{\epsilon_s} e^{-2aq_{||}} q_{||} + \left[\left(\frac{\omega_p}{2} \right)^2 - \frac{C e^2 \pi}{\epsilon_s} q_{||} \right]^2}, \quad (9.38)$$

and for large separation $aq_{||} \rightarrow \infty$, the plasmons disengage (ω_s is the decoupled surface plasmon and ω_{2D}^2 is the decoupled 2D plasmon)

$$\omega_+^2 = \frac{\omega_p^2}{2} \equiv \omega_s^2; \quad \omega_-^2 = \frac{2\pi e^2 C}{\epsilon_s} q_{||} \equiv \omega_{2D}^2, \quad (9.39)$$

whereas small separation yields the fully coupled hybridized modes as

$$\omega_+^2 = \omega_s^2 + \omega_{2D}^2 - 2aq_{\parallel} \frac{\omega_{2D}^2 \omega_s^2}{\omega_s^2 + \omega_{2D}^2}; \quad \omega_-^2 = 2aq_{\parallel} \frac{\omega_{2D}^2 \omega_s^2}{\omega_s^2 + \omega_{2D}^2}, \quad (9.40)$$

with plasmon linearity exhibited at low wavenumber ($q_{\parallel}a \ll 1$) as

$$\omega_+ \approx \omega_s + \frac{\pi C e^2}{\varepsilon_s \omega_s} q_{\parallel}; \quad \omega_- \approx 2e \sqrt{\frac{\pi a C}{\varepsilon_s}} q_{\parallel}. \quad (9.41)$$

Both of these modes are linear in lateral wavenumber, differing from the $q_{\parallel}^{\frac{1}{2}}$ -dependence for free-standing graphene or a 2DEG [25].

Generalizing our considerations to include an arbitrary number N of 2D layers which are Coulomb-coupled to a semi-infinite substrate as well as to each other, we have

$$\sum_{j=1}^N \left\{ \delta_{jj'} + \Pi_{2D;j}^{(0)}(q_{\parallel}, \omega) \int_{-\infty}^{\infty} dz' K_{SI}(a_{j'}, z') v(z' - a_j) \right\} K(a_j, z_2) = K_{SI}(a_{j'}, z_2). \quad (9.42)$$

Employing the local approximation for K_{SI} , i.e.

$$K_{SI}(z, z'; q_{\parallel}, \omega) = \theta(z) \left\{ \delta(z - z') + \delta(z') e^{-q_{\parallel} z} \left[\frac{1 - \varepsilon_B(\omega)}{1 + \varepsilon_B(\omega)} \right] \right\} \\ + \theta(-z) \left\{ \frac{\delta(z - z')}{\varepsilon_B(\omega)} + \delta(z') e^{q_{\parallel} z} \frac{1}{\varepsilon_B(\omega)} \left[\frac{\varepsilon_B(\omega) - 1}{\varepsilon_B(\omega) + 1} \right] \right\},$$

leads to the matrix form

$$\begin{pmatrix} K(a_1, z_2) \\ K(a_2, z_2) \\ \vdots \\ K(a_N, z_2) \end{pmatrix} = \frac{1}{S_c^{(N)}(q_{\parallel}, \omega)} \mathcal{N}^{\leftrightarrow(N)}(q_{\parallel}, \omega) \begin{pmatrix} K_{SI}(a_1, z_2) \\ K_{SI}(a_2, z_2) \\ \vdots \\ K_{SI}(a_N, z_2) \end{pmatrix} \quad (9.43)$$

where $S_c^{(N)}(q_{\parallel}, \omega) = \text{Det} \mathcal{M}^{\leftrightarrow(N)}(q_{\parallel}, \omega)$ with matrix elements given by

$$\begin{aligned}
\mathcal{M}_{jj'}^{(N)}(q_{\parallel}, \omega) &= \delta_{jj'} + \Pi_{2D;j}^{(0)}(q_{\parallel}, \omega) \int_{-\infty}^{\infty} dz' K_{SI}(a_j, z') v(z' - a_j) \\
&= \delta_{jj'} + \frac{2\pi e^2}{\epsilon_s q_{\parallel}} \Pi_{2D;j}^{(0)}(q_{\parallel}, \omega) \\
&\quad \times \left[e^{-q_{\parallel} |a_j - a_{j'}|} + e^{-q_{\parallel} (a_j + a_{j'})} \frac{\omega_p^2}{2\omega^2 - \omega_p^2} \right].
\end{aligned} \tag{9.44}$$

and $\mathcal{N}^{\leftrightarrow(N)}(q_{\parallel}, \omega)$ is the transpose cofactor matrix of $\mathcal{M}^{\leftrightarrow(N)}(q_{\parallel}, \omega)$.

To address the role of graphene nonlocality in the coupled plasmon spectrum in the wavenumber region between p_c^G and p_c^M , the nonlocal graphene polarizability ring diagram may be employed in (9.36) and (9.35), with K_{SI} again taken to be local as K_{SI} of (9.25). The nonlocal gapless graphene polarizability was derived by Wünsch et al. [26] and by Hwang and Das Sarma [27] in the degenerate limit; for convenience it is exhibited in Appendix 1. The corresponding result for gapped graphene was derived by Pyatkovskij [25] and is exhibited in Appendix 2.

9.4 Numerical Results for Plasmon Dispersion: Graphene Layers Interacting with Semi-infinite Conductor

Calculated numerical results for the plasmon dispersion of a system consisting of a semi-infinite conducting medium which is Coulomb-coupled to a single layer of ungapped graphene are shown in Fig. 9.2, and for two layers they are exhibited in Fig. 9.3. We note that the plasmon spectra and Landau damping both depend crucially on the separation between the constituents. Landau damping is determined by the imaginary part of the polarization function. For a single layer, Fig. 9.2 illustrates that when the plasmon mode enters a region with $\Im m I^{(0)}(q, \omega) \neq 0$, the mode is damped. Our calculations also show that when the distance a is less than a critical value $a_c = 0.4k_F^{-1}$, the lower acoustic plasmon is overdamped. On the other hand, our calculation shows that when the conducting surface-MLG separation exceeds the critical distance a_c , the intensity of the low-frequency plasmon in the long wavelength regime is high, and may be detected up to some cut-off wave vector q_c . For $q_{\parallel} > q_c$, the intensity of ω_c is very weak until the plasmon wave vector exceeds some value q'_c . However, as indicated above, for gapless graphene, when the surface-MLG separation is less than a_c , the surface plasmon intensity in the long wavelength regime is weak and the mode only appears at shorter wavelength when $q_{\parallel} > q'_c$. This behavior seems consistent with the reported experimental data: It is obviously the case if the plasmon branch goes below the main diagonal $\omega = v_F q$. Such damping, as well as the critical distance, changes in the presence of an energy bandgap for graphene.

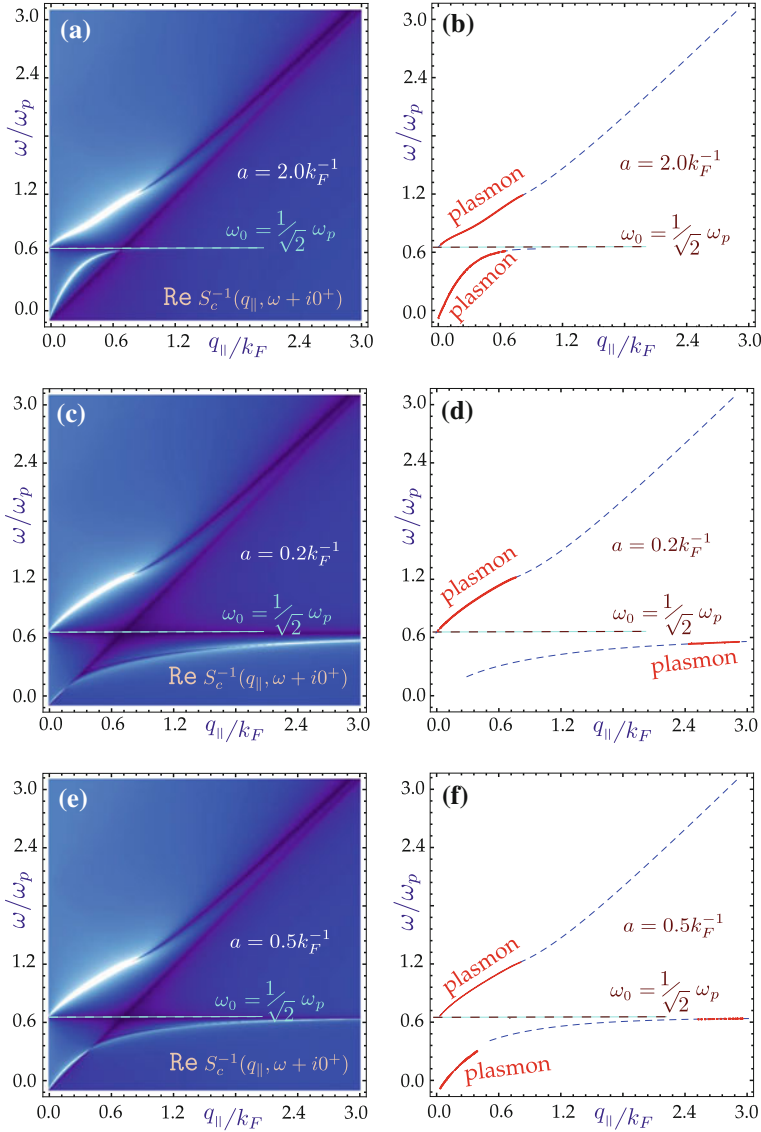


Fig. 9.2 Plasmon modes for a semi-infinite conductor interacting with a monolayer of ungapped graphene with various surface-to-layer separations. The *left panels a, c and e* present density plots of the inverse dispersion function $S_c^{-1}(q_{||}, \omega + i0^+)$ with peaks corresponding to the plasmon modes. The *right panels b, d and f* show exact numerical solutions for the plasmon branches, both Landau damped and undamped. The distances chosen are $ak_F = 0.2, 1.0$ and 0.5 , correspondingly. All the plots are provided for extrinsic graphene (doped) with zero energy bandgap

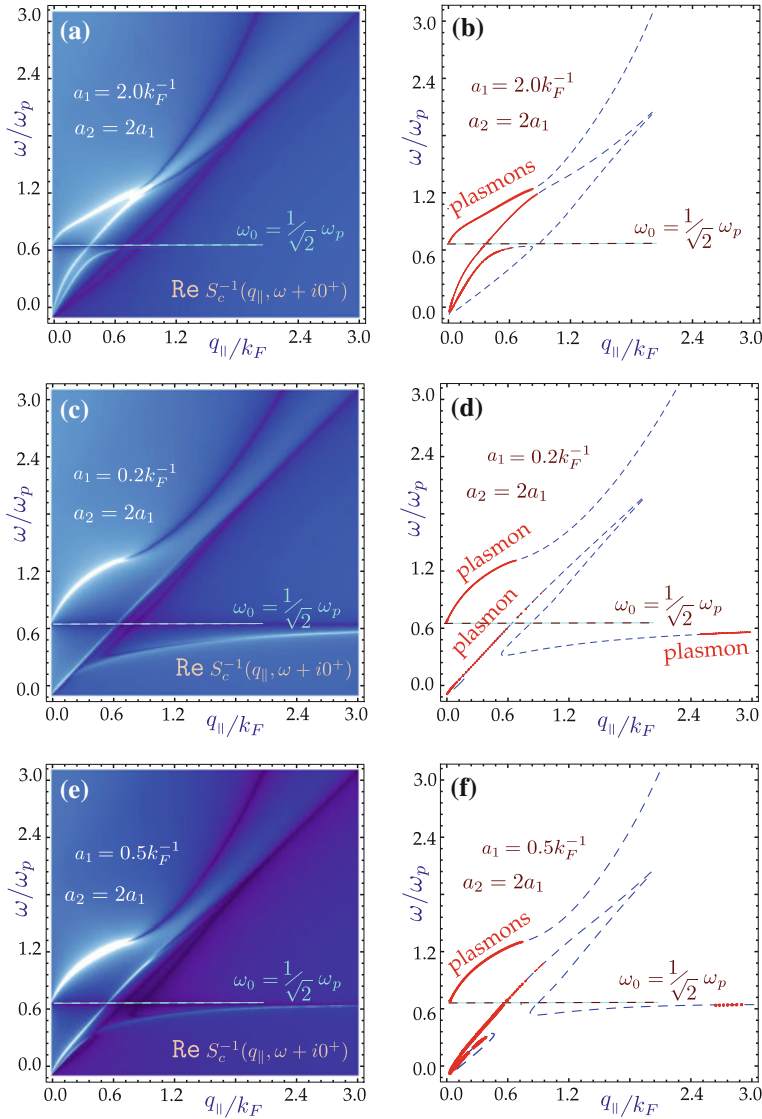


Fig. 9.3 Plasmon excitations for a semi-infinite conducting medium interacting with two ungapped graphene layers located at chosen distances from the surface. The *left panels a, c and e* give density plots of the inverse dispersion function $1/S_c^{N=2}(q_{||}, \omega + i0^+)$ with peaks corresponding to the plasmon modes. The *right panels b, d and f* show the exact numerical solutions for the plasmon branches, both Landau damped and undamped. The plots show results for various distances between the surface and the layers: $a_1 k_F = 2.0, 0.2$ and 0.5 , respectively. The second layer is placed at a distance a_2 , which is twice as large as a_1 . All the plots are provided for extrinsic graphene (doped) with zero energy bandgap

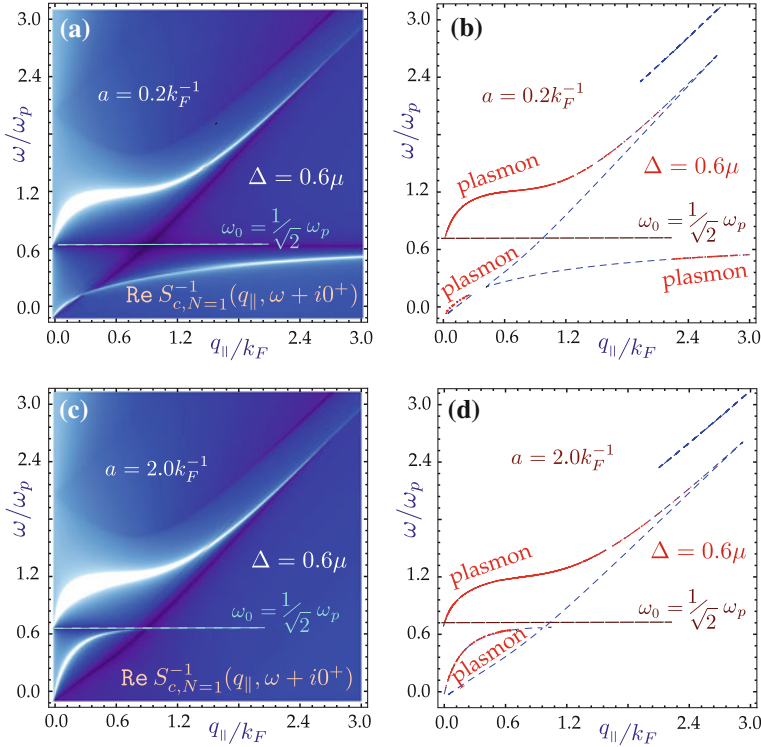


Fig. 9.4 Plasmon excitation spectra for a semi-infinite conductor Coulomb coupled with a single sheet of monolayer gapped graphene located at a chosen distance from the surface. The *left-hand panels* exhibit density plots of the inverse dispersion function $1/S_c^{N=1}(q_{||}, \omega + i0^+)$ with peaks corresponding to the plasmon modes. The *right panels* show the numerical solutions for the Landau damped (*dashed blue lines*) and undamped (*red curves*) plasmon branches

Similar calculated results for two ungapped graphene layers and a surface are presented in Fig. 9.3. The principal difference arising in the case when there are two Coulomb-coupled layers is that if the separation between the layer nearest the surface and the surface is less than a_c , then both the symmetric and antisymmetric modes become damped in different ranges of wave vector. It is important to mention that the upper plasmon branch (symmetric mode) remains almost unchanged for all cases, with either one or two graphene layers.

The role played by the energy band gap is an important part of our considerations. For monolayer graphene, the existence of an energy gap leads to an extended region of undamped plasmons [25]. We have focused on the regions outside of the single-particle excitation continuum with $\text{Im } \Pi_{2D}^{(0)}(q_{||}, \omega) = 0$, since the plasmons in these regions are not Landau-damped. In Fig. 9.4, we have plotted the plasmon dispersion relation for the single graphene layer with an energy bandgap for two chosen distances between the graphene layer and the conducting surface. These

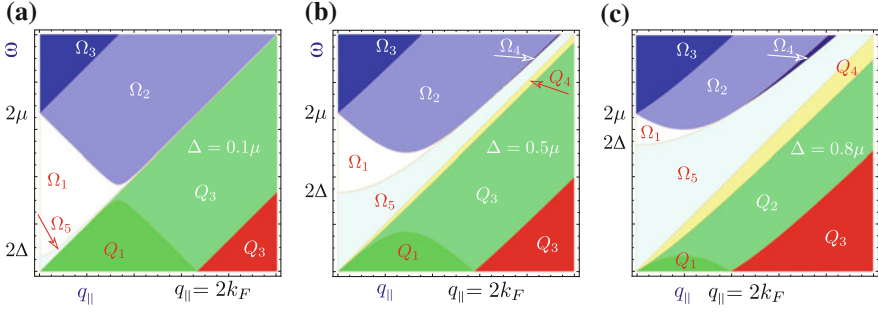


Fig. 9.5 Schematic illustrations of the $(\omega - q_{\parallel})$ -space regions having differing analytical expressions for the non-interacting polarization function given in Appendix 2 as taken from [25]. Both real and imaginary parts each have distinct analytical expressions, as described in [25]. The regions with $\omega > \hbar v_F q_{\parallel}$ above the main diagonal are denoted by Ω_1 – Ω_5 , while the regions below this diagonal are denoted by Q_1 – Q_4 . Regions with zero $\text{Im}I_{2D}^{(0)}(q, \omega)$, namely Q_4 , Ω_1 and Ω_5 correspond to regions where undamped plasmons exist. Each schematic plot illustrates the case with a different energy bandgap: **a** $\Delta = 0.1 \mu$; **b** $\Delta = 0.5 \mu$; and **c** $\Delta = 0.8 \mu$

results show that for a conducting substrate surface plasmon frequency denoted by $\omega_s = \omega_p / \sqrt{2}$, there is a coupled/hybrid plasmon frequency ω_c of the composite system which lies below ω_s . Furthermore, the intensity of this mode depends on the distance of the graphene layer from the surface of the conductor as well as on the energy band gap between the valence and conduction bands of the graphene monolayer (MLG).

In Fig. 9.5, we illustrate the regions in which both the real and imaginary parts of the polarization function have distinct analytical expressions for several bandgaps. As mentioned above, we have given particular attention to these regions with $\text{Im}I^{(0)}(q, \omega) = 0$ outside of the single-particle excitation continuum, since the plasmons in these regions are not Landau-damped. The detailed analytic expressions for the polarization function in each region are given in Appendix 2.

9.5 Experimental Studies on Epitaxial Graphene

Electron energy loss spectroscopy (EELS) is the main experimental technique employed in the investigation of collective electronic excitations (see [28–30] for a review of this technique). In particular, while photons can only access the long-wavelength region, EELS is able to scan an extended range in reciprocal space (up to 0.4 – $0.5 \text{ \AA}^{-1}$) [29], encompassing both the long- and the short-wavelength regions. Figure 9.6 shows a typical loss spectrum of plasmon modes for graphene.

While a square-root-like dispersion of the intraband plasmon has been reported for free-standing graphene and graphene/SiC, experiments performed on monolayer

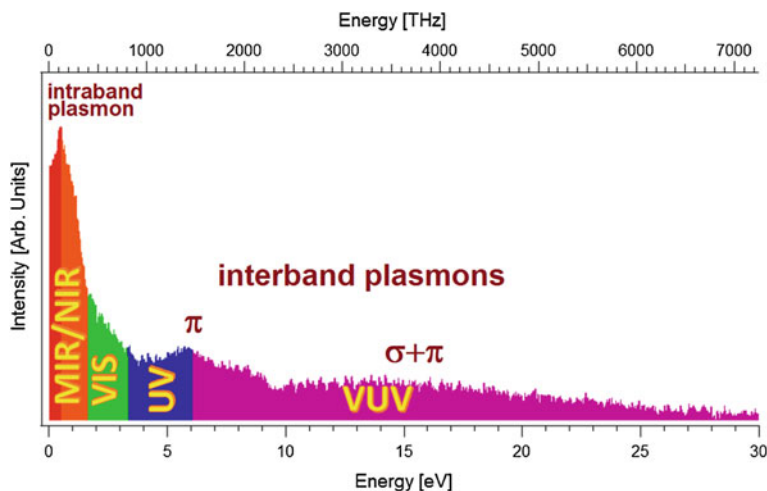
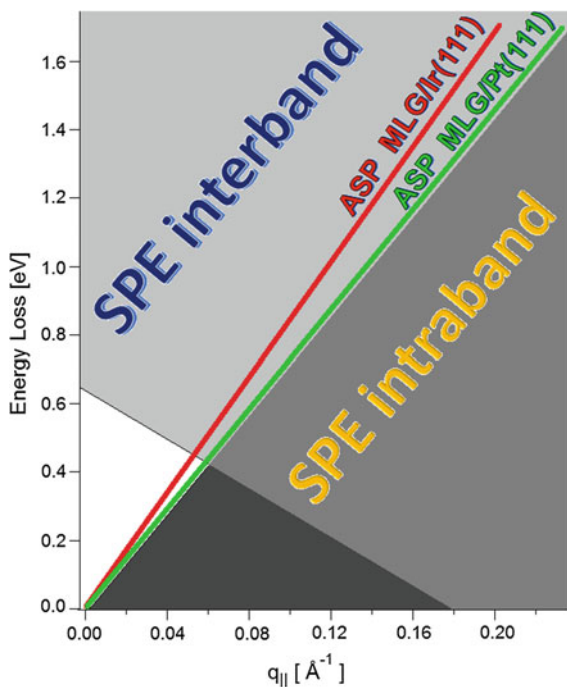


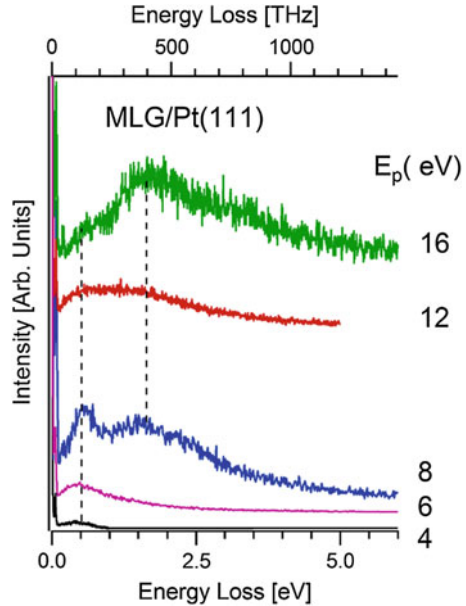
Fig. 9.6 Intraband and interband plasmons in epitaxial graphene on Pt(111)

Fig. 9.7 Dispersion relation of the ASP in MLG/Pt(111) (data taken from [32]) and MLG/Ir(111) (data taken from [34])



graphene grown on metals exhibit a linear dispersion for the intraband plasmon. In particular, experiments for graphene on Pt(111) [31–33] and on Ir(111) [34] have shown a nearly identical dispersion relation (Fig. 9.7).

Fig. 9.8 Dependence of the HREELS spectrum for monolayer graphene (MLG) grown on Pt(111) on the primary electron beam energy. Spectra have been recorded in the same scattering geometry (incidence angle $\theta_i = 55^\circ$ and scattering angle $\theta_s = 51^\circ$, respectively, measured with respect to the sample normal). In the top axis, the energy scale is reported in THz



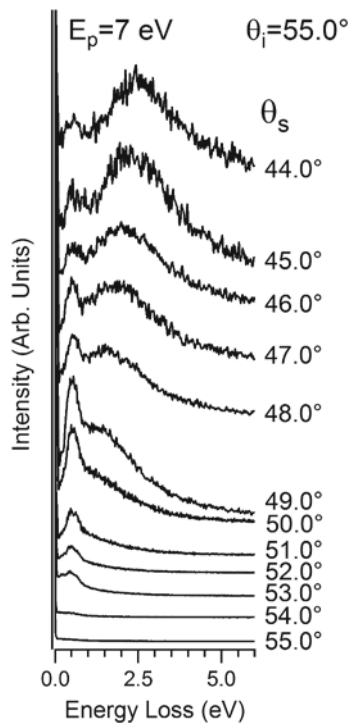
The experimental study of plasmon modes of graphene/metals merits special attention. Graphene-metal contacts [35–44] are omnipresent and inevitable components of graphene-based devices and, thus, the investigation of the nature and dispersion of plasmon modes at graphene/metal interfaces is a key step toward engineering plasmonic applications of graphene.

The linear dispersion of 2D plasmons at graphene/metal interfaces is a consequence of the coexistence of the bounded metal electron gas with the π -charge density of graphene in close proximity. It resembles the acoustic surface plasmon (ASP) at metal surfaces with a partially occupied surface state band within a wide bulk energy gap [45, 46]. The non-local character of the dielectric function [11, 14, 24] and of the screening processes in graphene [47] prevents the sheet plasmon from being screened out by the 3D bulk states of the metal substrate.

The group velocity of intraband plasmons at graphene/metal interfaces ($\approx 10^6$ m/s), extracted from the slope of the dispersion relation, is similar to those calculated for ASP, thus indicating a similar nature of the two collective excitations.

The group velocity of the sheet plasmons at graphene/metal interfaces is about two orders of magnitude lower than the speed of light, and thus its direct excitation by light is not possible. However, nanometer-size objects at such surfaces, including atomic steps or molecular structures, can facilitate coupling between the sheet plasmons and light. The linear behavior of its dispersion implies that both phase and group velocities of the collective excitation are the same, so signals can be transmitted undistorted along the surface with potential applications in graphene-based nano-optical devices.

Fig. 9.9 HREELS spectra for MLG/Pt(111), recorded as a function of the scattering angle. The primary energy is 7 eV, while the incidence angle is fixed at 55° with respect to the sample normal



In addition to the ordinary intraband plasmon, a few recent theoretical works show the existence of a nonlinear plasmon in free-standing doped graphene [48, 49]. Although a similar excitation has also been observed experimentally [33, 50] (see the nonlinear mode in Figs. 9.8, 9.9 and 9.10), its origin is still unclear and currently under investigation.

The electronic response of self-doped graphene supported on metal substrates shows a remarkable dependence on the primary electron beam energy. Loss spectra, obtained by using impinging electrons of low kinetic energy 7–8 eV, show both acoustic and nonlinear plasmons.

For primary energies below 6 eV, only a peak around 0.5 eV is observed. At $\mathcal{E}_p = 8$ eV, two clearly distinct plasmon modes may be resolved spectroscopically. The second plasmon mode around 1.65 eV increases its spectral weight and it is predominant in the spectrum recorded for $\mathcal{E}_p = 16$ eV. In addition to these peaks, graphene optical phonons are also clearly observed at 0.1 and 0.2 eV [51, 52]. This striking dependence on the primary energy explains why previous experimental studies were not able to resolve the two plasmon modes [32, 34].

The HREELS spectra for $\mathcal{E}_p = 7$ eV, recorded as a function of the scattering angle, are reported in Fig. 9.9. The two plasmon modes develop with variation of the scattering angle. The higher-energy mode emerges in off-specular geometries, for values of θ_s lower than 50° with respect to the sample normal. This mode is

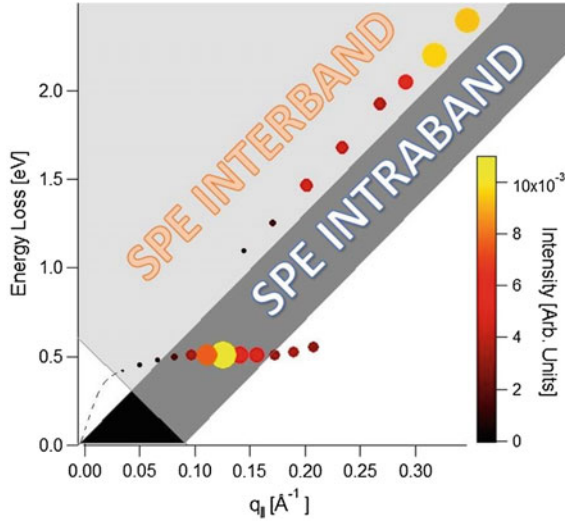


Fig. 9.10 Plasmon dispersion in MLG/Pt(111). The size of the markers reflects the intensity of the plasmon modes, normalized to that of the elastic peak. Markers are colored following a color scale reported on the right. The *light gray area* indicates the continuum of interband single-particle excitations (SPE), while the *dark gray* represents the area of the continuum of intraband SPEs. The plasmon mode enters the Landau damping regime of interband electron-hole excitations when its dispersion relation intercepts the boundary for the continuum of interband SPEs. The *dotted line* represents an interpolation of the nonlinear mode for small momenta, for which the intensity vanishes

characterized by a linear dispersion (acoustic plasmon). By contrast, the low-energy mode has a nonlinear dispersion relation (Fig. 9.10). Its intensity reaches a maximum and then decreases for higher momenta.

The dispersion curve of the nonlinear plasmon is well fitted by a power law:

$$\hbar\omega_{\text{nonlinear}} = \mathcal{B}q_{\parallel}^n \quad (9.45)$$

with $\mathcal{B} = (0.66 \pm 0.02) \text{ eV } \text{\AA}^n$ and $n = (0.12 \pm 0.01)$. In Fig. 9.10, we also show the electron-hole continuum or single-particle excitation (SPE) region.

The acoustic plasmon lies above the upper edge $v_F q_{\parallel}$ of the intraband excitations and it can be damped by interband transitions. In contrast to this, the nonlinear plasmon is located below $v_F q_{\parallel}$ and thus it is damped by intraband transitions.

The nonlinear plasmon was first predicted by Gao et al. [48] for doped free-standing graphene. It should be noted that theoretical work based on an oversimplified isotropic graphene band structure in the vicinity of the K point cannot explain it [53]. Its nonlinear dispersion, considered jointly with its high spectral weight and its lower energy, makes it a natural candidate for applications in graphene-based nonlinear plasmonics [54, 55].

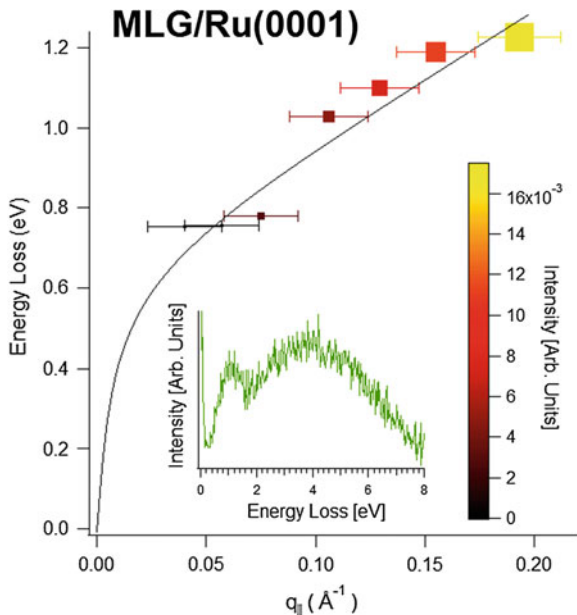


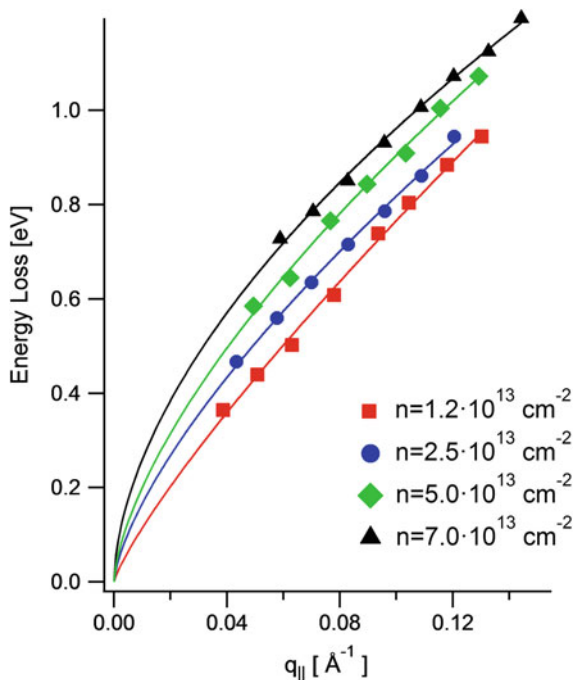
Fig. 9.11 Behavior of the nonlinear mode as a function of parallel momentum transfer $q_{||}$ for monolayer graphene (MLG) on Ru(0001). The intensity of the plasmon modes, normalized to that of the elastic peak, is indicated by the size of the markers, which are colored following a color scale reported on the *right*. The *solid line* represents the best fit. In the *inset*, a selected loss spectrum for MLG/Ru(0001) has been reported. Two distinct loss features are present

It should be borne in mind that in undoped graphene, both the acoustic plasmon and the nonlinear mode do not occur [48, 56]. In the case of MLG/Pt(111), the p-type doping of the graphene sheet at this interface is responsible for the existence of the above-mentioned two plasmon modes.

We also note that when the acoustic plasmon coexists in the EELS spectrum with the nonlinear mode, its intensity in the range 0.6–1.0 eV is particularly weak (as highlighted by the size of markers in Fig. 9.10) in comparison with the case of impinging energies above 12 eV, for which only the acoustic plasmon is recorded. Thus, the coexistence of the two modes induces a gap in the dispersion relation when considering the intensity of the plasmon modes.

Finally, in order to investigate whether the nonlinear mode depends on the graphene/metal interface, we have also studied its dispersion for MLG/Ru(0001) (Fig. 9.11). We find that two plasmon modes also exist in this case, even though the dispersion of the nonlinear mode is slightly different due to the dissimilar properties of this graphene/metal interface from those of MLG/Pt(111) (see [57–59] for more details).

Fig. 9.12 Dispersion relation of the intraband plasmon for several different values of electron density n . The intraband plasmon excitation arising from intrinsic electron density $n = 1.2 \times 10^{13} \text{ cm}^{-2}$ changes when more electrons are added by K doping. Adapted from [62]



9.6 Influence of Adsorbed and Intercalated Atoms

The tunability of graphene plasmons is a promising opportunity for the emerging field of graphene-based plasmonics. Both the conductivity of graphene and the dispersion of plasmons in graphene are related to the Fermi energy (or chemical potential at room temperature). The Fermi energy

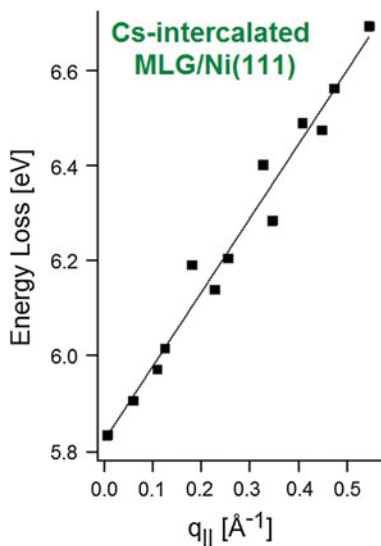
$$\mathcal{E}_F \approx \mu \approx \sqrt{\pi \hbar^2 v_F^2 n} \quad (9.46)$$

can be easily controlled (with a concomitant tuning of plasmon modes) by changing the charge concentration, which, in turn, can be realized by varying a gate potential or chemical doping [60]. In consideration of possible applications, it should be mentioned that distributing carriers into multiple graphene/insulator stacks efficiently enhances the plasmonic resonance frequency and its magnitude [61].

Highly electropositive or electronegative chemical species can easily dope graphene due to charge transfer. In Fig. 9.12, results for *K*-doped graphene/SiC [62] are reported. With the addition of more electrons in the π^* band, the Fermi level shifts upward away from the Dirac point E_D , causing a continuous blue-shift of the plasmon energy with increasing potassium coverage.

Similarly, a red-shift has been reported in water-exposed MLG/Pt(111) [31]. However, in the latter study, charge doping affects the frequency of the plasmon

Fig. 9.13 Plasmon dispersion for Cs-intercalated graphene/Ni(111). The linear dispersion of the π plasmon in pristine, free-standing graphene is recovered [63, 64]. Adapted from [65]



modes but not the acoustic behavior of the dispersion relation. Nevertheless, a reduction of the group velocity by about 11 % has been observed [31].

In contrast, experiments performed for Na-doped MLG/Ir (111) [34] and F4-TCNQ/MLG/SiC [65] by Pfnür's group show negligible influence of doping on the dispersion curve while noticeable changes in the damping mechanisms of the plasmon modes are reported.

Interesting results have been reported for plasmon dispersion in Cs-intercalated graphene/Ni(111) [66]. Without intercalation, the close surface proximity of graphene on Ni(111) exhibits a strong hybridization between Ni d bands and π states of graphene [67, 68]. The concomitant change of the energy spectrum results in the loss of the Dirac cones at the graphene K-points, and, with this, the linear plasmon spectrum is lost. However, intercalated chemical species may decouple such orbital interactions between the graphene sheet and the substrate [69, 70], restoring the integrity of graphene's Dirac cones. Correspondingly, the influence of intercalated atoms on the electronic structure of MLG/Ni(111) can manifest itself in the collective electronic properties of the intercalated system involving just the Coulomb coupling of the graphene sheet with a substrate Ni plasma (devoid of orbital hybridization effects with Ni). Thus, the dispersion curve of the π plasmon, in comparison with the analogous dispersion curve obtained in the absence of intercalated Cs, shows that alkali metal atoms make graphene effectively quasi-freestanding in this sense, with a much lowered charge transfer, and with the recovery of the linear plasmon dispersion typical of free-standing graphene whose interaction with the substrate Ni is solely Coulombic (Fig. 9.13).

9.7 Concluding Remarks

In summary, we have reviewed nonlocal plasmon dispersion in the Random Phase Approximation for monolayer graphene (as well as for a pair of layers) interacting with a conducting substrate: this was done for arbitrary wavelength, also examining dependence on layer-surface separation. We have emphasized some interesting effects of Landau damping of the plasmon modes which depend on the layer separation distance. Specifically, our numerical simulations exhibited in Fig. 9.2 indicate that, for a single layer whose distance is less than a critical value a_c , the lower acoustic plasmon mode is severely Landau damped; this behavior is consistent with reported experimental data. This may be seen in our plots when the plasmon energy lies below the main diagonal $\omega = v_F q$. The damping, as well as the critical distance, changes in the presence of an energy bandgap for graphene.

In regard to plasmon damping, we also discuss results for two graphene layers and a conductor, which are presented in Fig. 9.3. But, a major difference between the case in which there are two graphene layers rather than one is that, for separations between the closest layer and the conducting surface less than the critical value a_c , both the symmetric and antisymmetric modes become damped in different ranges of wave vector. It should be noted that the upper plasmon branch (symmetric mode) remains almost unaffected by Landau damping for all cases, involving either one or two graphene layers. Furthermore, the role of an energy gap in the graphene plasmon spectrum and damping has been addressed here in connection with Figs. 9.4 and 9.5. Moreover, considerable attention to the experimental studies of graphene plasmons is provided in Figs. 9.6–9.13. The reader should be aware that other materials having the Dirac-like energy spectrum of graphene are currently under investigation, in particular silicene and, most recently, topological insulators [71].

A fuller exposition of the theory and its bearing on experimental data is provided by Gumbs et al. [24]. The experimental data clearly confirm linear plasmon dispersion in the long wavelength limit. The simulations reviewed here go beyond this regime to a fully general numerical description of nonlocal effects in monolayer and bilayer graphene, including dependencies on spatial separation a and gap strength. Further developments of this work to include the roles of nonlocality of the substrate plasma and of its finite width may be facilitated by the results provided in Sect. 9.2.

Acknowledgments This work was supported in part by contract # FA 9453-13-1-0291 of AFRL.

Appendix 1: Dynamic Nonlocal Polarization Function for Free-Standing Graphene with no Bandgap; Brief Summary of the Results Derived in [19, 20]

The 2D RPA density perturbation response function, $\delta\rho/\delta V = R_{2D}^{(0)}(\mathbf{p}, \omega) \equiv D_0 \tilde{R}(x, \nu)$, for free-standing gapless Graphene in the $T = 0$

degenerate limit is given in terms of dimensionless frequency and wavenumber variables defined by $v = \omega/E_F = \omega/\mu$ and $x = p/p_F$, respectively, as (note that $D_0 \equiv \gamma^{-1} \sqrt{g_s g_v \rho_{2D}/\pi}$; g_s and g_v are spin and valley degeneracies, $g_s = g_v = 2$; $\hbar \rightarrow 1$):

$$\tilde{R}(x, v) = \tilde{R}^+(x, v) + \tilde{R}^-(x, v), \quad (9.47)$$

with $(\theta(z) \equiv \eta_+(z) = \text{Heaviside unit step function})$

$$\tilde{R}^+(x, v) = \tilde{R}_1^+(x, v)\theta(v - x) + \tilde{R}_2^+(x, v)\theta(x - v) \quad (9.48)$$

where (define $\tilde{\Pi} \equiv -\tilde{R}$)

$$\begin{aligned} -\text{Re}\tilde{R}_1^+(x, v) = \text{Re}\tilde{\Pi}_1^+(x, v) &= 1 - \frac{1}{8\sqrt{v^2 - x^2}} \{f_1(x, v)\theta(|2 + v| - x) \\ &+ \text{sgn}(v - 2 + x)f_1(x, -v)\theta(|2 - v| - x) \\ &+ f_2(x, v)[\theta(x + 2 - v) + \theta(2 - x - v)]\}, \end{aligned} \quad (9.49)$$

$$\begin{aligned} -\text{Re}\tilde{R}_2^+(x, v) = \text{Re}\tilde{\Pi}_2^+(x, v) &= 1 - \frac{1}{8\sqrt{x^2 - v^2}} \{f_3(x, v)\theta(x - |v + 2|) \\ &+ f_3(x, -v)\theta(x - |v - 2|) \\ &+ \frac{\pi x^2}{2} [\theta(|v + 2| - x) + \theta(|v - 2| - x)]\}, \end{aligned} \quad (9.50)$$

$$\begin{aligned} -\text{Im}\tilde{R}_1^+(x, v) = \text{Im}\tilde{\Pi}_1^+(x, v) &= \frac{-1}{8\sqrt{v^2 - x^2}} \{f_3(x, -v)\theta(x - |v - 2|) \\ &+ \frac{\pi x^2}{2} [\theta(x + 2 - v) + \theta(2 - x - v)]\}, \end{aligned} \quad (9.51)$$

$$-\text{Im}\tilde{R}_2^+(x, v) = \text{Im}\tilde{\Pi}_2^+(x, v) = \frac{\theta(v - x + 2)}{8\sqrt{x^2 - v^2}} [f_4(x, v) - f_4(x, -v)\theta(2 - x - v)] \quad (9.52)$$

and

$$-\tilde{R}^-(x, v) = \frac{\pi x^2 \theta(x - v)}{8\sqrt{x^2 - v^2}} + i \frac{\pi x^2 \theta(v - x)}{8\sqrt{v^2 - x^2}}. \quad (9.53)$$

The quantities $f_1(x, v)$, $f_2(x, v)$, $f_3(x, v)$, $f_4(x, v)$ are defined as

$$f_1(x, v) = (2+v) \sqrt{(2+v)^2 - x^2} - x^2 \ln \frac{\sqrt{(2+v)^2 - x^2} + (2+v)}{|\sqrt{v^2 - x^2} + v|}, \quad (9.54)$$

$$f_2(x, v) = x^2 \ln \frac{v - \sqrt{v^2 - x^2}}{x}, \quad (9.55)$$

$$f_3(x, v) = (2+v) \sqrt{x^2 - (2+v)^2} + x^2 \sin^{-1} \frac{2+v}{x}, \quad (9.56)$$

$$f_4(x, v) = (2+v) \sqrt{(2+v)^2 - x^2} - x^2 \ln \frac{\sqrt{(2+v)^2 - x^2} + (2+v)}{x}. \quad (9.57)$$

Appendix 2: Dynamic Nonlocal Polarization Function for Graphene with a Finite Energy Bandgap; Brief Summary of the Results Derived in [18]

The 2D RPA ring diagram polarization function for graphene with a gap Δ may be expressed as

$$\begin{aligned} \Pi_{2D}^{(0)}(q, \omega) &= \frac{g}{4\pi^2} \int d^2k \sum_{s,s'=\pm} \left(1 + ss' \frac{\mathbf{k} \cdot (\mathbf{k} + \mathbf{q}) + \Delta^2}{\varepsilon_k \varepsilon_{|\mathbf{k}+\mathbf{q}|}} \right) \\ &\times \frac{f(s\varepsilon_k) - f(s'\varepsilon_{\mathbf{k}+\mathbf{q}})}{s\varepsilon_{\mathbf{k}} - s'\varepsilon_{\mathbf{k}+\mathbf{q}} - \hbar\omega - i\hbar\delta} \end{aligned} \quad (9.58)$$

Since we limit our considerations to zero temperature, $T = 0$, the Fermi-Dirac distribution function is reduced to the Heaviside step function $f(\varepsilon, \mu; T \rightarrow 0) = \eta_+(\mu - \varepsilon)$, so (9.58) is simplified to

$$\Pi_{2D}^{(0)}(q, \omega) = -\chi_-^\infty(q, \omega) + \chi_+^\mu(q, \omega) + \chi_-^\mu(q, \omega), \quad (9.59)$$

where

$$\begin{aligned} \chi_\pm^\alpha &= \frac{g}{4\pi^2} \int d^2\mathbf{k} \eta_+(\alpha^2 - \Delta^2 - q^2) \sum_{s,s'=\pm} \left(1 \pm ss' \frac{\mathbf{k} \cdot (\mathbf{k} + \mathbf{q}) + \Delta^2}{\varepsilon_k \varepsilon_{|\mathbf{k}+\mathbf{q}|}} \right) \\ &\times \left(\frac{1}{\hbar\omega + \varepsilon_k \mp \varepsilon_{|\mathbf{k}+\mathbf{q}|} + i\hbar\delta} - \frac{1}{\hbar\omega - \varepsilon_k \pm \varepsilon_{|\mathbf{k}+\mathbf{q}|} + i\hbar\delta} \right) \end{aligned} \quad (9.60)$$

and

$$\Pi_{2D}^{(0)}(q, \omega) = \Pi_{(0)}(q, \omega) + \eta_+(\mu - \Delta)\Pi_{(1)}(q, \omega), \quad (9.61)$$

where

$$\begin{aligned} \Pi_{(0)}(q, \omega) &= -\chi_-^\infty(q, \omega) \\ \Pi_{(1)}(q, \omega) &= \chi_+^\mu(q, \omega) + \chi_-^\mu(q, \omega). \end{aligned} \quad (9.62)$$

The following notations are employed to specify the expressions involved in the polarization function:

$$\mathcal{F}(q, \omega) = \frac{g\hbar v_F^2}{16\pi} \frac{q^2}{\sqrt{|\omega^2 - (v_F q)^2|}} \quad (9.63)$$

and

$$\mathcal{X}_0 = \sqrt{1 - \frac{4\Delta^2}{\hbar^2(\omega^2 - (v_F q)^2)}}, \quad (9.64)$$

along with definitions of the following functions:

$$\mathcal{G}_<(x) = x\sqrt{\mathcal{X}_0^2 - x^2} - (2 - \mathcal{X}_0) \arccos(x/\mathcal{X}_0), \quad (9.65)$$

$$\mathcal{G}_>(x) = x\sqrt{x^2 - \mathcal{X}_0^2} - (2 - \mathcal{X}_0) \operatorname{arccosh}(x/\mathcal{X}_0), \quad (9.66)$$

$$\mathcal{G}_>(x) = x\sqrt{x^2 - \mathcal{X}_0^2} - (2 - \mathcal{X}_0) \operatorname{arcsinh}(x/\sqrt{-\mathcal{X}_0^2}). \quad (9.67)$$

Finally, the polarization function is given in the following form:

Real part

$$\text{Re}\Pi^{(0)}(q, \omega) = -\frac{g\mu}{\hbar^2 v_F^2} + \mathcal{F}(q, \omega)$$

$$\times \begin{cases} 0 & \Rightarrow Q_1 \\ \mathcal{G}_<\left(\frac{\mathcal{M}}{\hbar v_F q}\right) & \Rightarrow Q_2 \\ \mathcal{G}_<\left(\frac{\mathcal{P}}{\hbar v_F q}\right) + \mathcal{G}_<\left(\frac{\mathcal{M}}{\hbar v_F q}\right) & \Rightarrow Q_3 \\ \mathcal{G}_>\left(\frac{\mathcal{P}}{\hbar v_F q}\right) - \mathcal{G}_<\left(\frac{\mathcal{P}}{\hbar v_F q}\right) & \Rightarrow Q_4 \\ \mathcal{G}_>\left(\frac{\mathcal{P}}{\hbar v_F q}\right) - \mathcal{G}_<\left(\frac{\mathcal{M}}{\hbar v_F q}\right) & \Rightarrow \Omega_1 \\ \mathcal{G}_>\left(\frac{\mathcal{P}}{\hbar v_F q}\right) & \Rightarrow \Omega_2 \\ \mathcal{G}_>\left(\frac{\mathcal{P}}{\hbar v_F q}\right) - \mathcal{G}_>\left(\frac{\mathcal{M}}{\hbar v_F q}\right) & \Rightarrow \Omega_3 \\ \mathcal{G}_>\left(\frac{\mathcal{M}}{\hbar v_F q}\right) + \mathcal{G}_>\left(\frac{\mathcal{P}}{\hbar v_F q}\right) & \Rightarrow \Omega_4 \\ \mathcal{G}_0\left(\frac{\mathcal{P}}{\hbar v_F q}\right) - \mathcal{G}_0\left(\frac{\mathcal{M}}{\hbar v_F q}\right) & \Rightarrow \Omega_1 \end{cases} \quad (9.68)$$

Imaginary part

$$\text{Im}\Pi^{(0)}(q, \omega) = -\mathcal{F}(q, \omega)$$

$$\times \begin{cases} \mathcal{G}_>\left(\frac{\mathcal{P}}{\hbar v_F q}\right) - \mathcal{G}_>\left(\frac{\mathcal{M}}{\hbar v_F q}\right) & \Rightarrow Q_1 \\ \mathcal{G}_>\left(\frac{\mathcal{P}}{\hbar v_F q}\right) & \Rightarrow Q_2 \\ 0 & \Rightarrow Q_3 \\ 0 & \Rightarrow Q_4 \\ 0 & \Rightarrow \Omega_1 \\ -\mathcal{G}_<\left(\frac{\mathcal{M}}{\hbar v_F q}\right) & \Rightarrow \Omega_2 \\ \pi[2 - \mathcal{X}_0^2] & \Rightarrow \Omega_3 \\ \pi[2 - \mathcal{X}_0^2] & \Rightarrow \Omega_4 \\ 0 & \Rightarrow \Omega_5 \end{cases} \quad (9.69)$$

where

$$\mathcal{P} = 2\mu + \hbar\omega \text{ and } \mathcal{M} = 2\mu - \hbar\omega. \quad (9.70)$$

The analytic expressions provided in the left columns above for the real and imaginary parts of Π^0 pertain to the frequency wavenumber regions marked by the Q 's and Ω 's in the corresponding right column as indicated in Fig. 9.4. Specifically, these ω - q regions are defined as follows:

$$\begin{aligned}
Q_1 &\Rightarrow \hbar\omega < \mu - \sqrt{(\hbar v_F)^2(q - k_F^\mu)^2 + \Delta^2} \\
Q_2 &\Rightarrow \pm\mu \mp \sqrt{(\hbar v_F)^2(q - k_F^\mu)^2 + \Delta^2} < \hbar\omega < \\
&\quad -\mu + \sqrt{(\hbar v_F)^2(q + k_F^\mu)^2 + \Delta^2} \\
Q_3 &\Rightarrow \hbar\omega < -\mu + \sqrt{(\hbar v_F)^2(q + k_F^\mu)^2 + \Delta^2} \\
Q_4 &\Rightarrow -\mu + \sqrt{(\hbar v_F)^2(q + k_F^\mu)^2 + \Delta^2} < \hbar\omega < \hbar v_F q,
\end{aligned} \tag{9.71}$$

and

$$\begin{aligned}
\Omega_1 &\Rightarrow q < 2k_F \quad \& \quad \sqrt{(\hbar v_F)^2 q^2 + 4\Delta^2} < \hbar\omega < \\
&\quad < -\mu + \sqrt{(\hbar v_F)^2(q - k_F^\mu)^2 + \Delta^2} \\
\Omega_2 &\Rightarrow \mu + \sqrt{(\hbar v_F)^2(q - k_F^\mu)^2 + \Delta^2} < \\
&\quad \mu + \sqrt{(\hbar v_F)^2(q - k_F^\mu)^2 + \Delta^2} \\
\Omega_3 &\Rightarrow \hbar\omega > \mu + \sqrt{(\hbar v_F)^2(q + k_F^\mu)^2 + \Delta^2} \\
\Omega_4 &\Rightarrow q > 2k_F \quad \& \quad \sqrt{(\hbar v_F)^2 q^2 + 4\Delta^2} < \hbar\omega < \\
&\quad < -\mu + \sqrt{(\hbar v_F)^2(q - k_F^\mu)^2 + \Delta^2} \\
\Omega_5 &\Rightarrow \hbar v_F q < \hbar\omega < \sqrt{(\hbar v_F)^2 q^2 + 4\Delta^2}.
\end{aligned} \tag{9.72}$$

References

1. D.S.L. Abergel, V. Apalkov, J. Berashevich, K. Ziegler, T. Chakraborty, Properties of graphene: a theoretical perspective. *Adv. Phys.* **59**(4), 261482 (2010)
2. M.I. Katsnelson, *Graphene: Carbon in Two Dimensions* (Cambridge University Press, 2012)
3. H.S. Aoki, M. Dresselhaus, *Physics of Graphene* (Springer, 2014)
4. E.L. Wolf, *Graphene: A New Paradigm in Condensed Matter and Device Physics* (Oxford University Press, 2013)
5. G. Gumbs, D. Huang, O. Roslyak, Electronic and photonic properties of graphene layers and carbon nanoribbons. *Philos. Trans. R. Soc. A* **368**, 5351–5556 (2010)
6. A.H.C. Neto, F. Guinea, N.M.R. Peres, K.S. Novoselov, A.K. Geim, The electronic properties of graphene. *Rev. Mod. Phys.* **81**(1), 109 (2009)
7. S. Das Sarma, S. Adam, E.H. Hwang, E. Rossi, Electronic transport in two-dimensional graphene. *Rev. Mod. Phys.* **83**(2), 407 (2011)
8. L.E.F.F. Torres, S. Roche, J. Charlier, *Introduction to Graphene-Based Nanomaterials From Electronic Structure to Quantum Transport* (Cambridge University Press, 2014). ISBN:9781107030831
9. C.J. Tabert, E.J. Nicol, Dynamical polarization function, plasmons, and screening in silicene and other buckled honeycomb lattices. *Phy. Rev. B* **89**(19), 195410 (2014)

10. N.J. Horing, Quantum theory of electron gas plasma oscillations in a magnetic field. *Ann. Phys. (NY)* **31**, 1–63 (1965)
11. N.J.M. Horing, Aspects of the theory of graphene. *Philos. Trans. R. Soc. A* **368**(1932), 5525–5556 (2010)
12. N.J.M. Horing, Inhomogeneous structure of the nonlocal dynamic dielectric response of a quantum-well bound state and a three dimensional band of extended states. *Phys. Rev. B* **59**(8), 5648 (1999)
13. N.J.M. Horing, G. Gumbs, T. Park, Coupling of 2D plasmons to nonlocal bulk plasmons. *Phys. B* **299**(1), 165172 (2001)
14. N.J.M. Horing, Coupling of graphene and surface plasmons. *Phys. Rev. B* **80**(19), 193401 (2009)
15. P.C. Martin, J. Schwinger, Theory of many-particle systems. I *Phys. Rev.* **115**(6), 1342 (1959)
16. R.H. Ritchie, Surface plasmons in solids. *Surf. Sci.* **34**(1), 119 (1973)
17. H. Raether, Surface plasmons on smooth and rough surfaces and on gratings. *Springer Tracts Mod. Phys.* **111**, (1988)
18. H. Raether, Excitations of plasmons and interband transitions by electrons. *Springer Tracts Mod. Phys.* **88**, (1980)
19. P.J. Feibelman, Surface electromagnetic fields. *Prog. Surf. Sci.* **12**(4), 287407 (1982)
20. A. Liebsch, *Electronic Excitations at Metal Surfaces* (Plenum Pub Corp, New York, 1997). ISBN 10:0306455455
21. M.S. Kushwaha, Plasmons and magnetoplasmons in semiconductor heterostructures. *Surf. Sci. Rep.* **41**(1), 1416 (2001)
22. D.M. Newns, Dielectric response of a semi-infinite degenerate electron gas. *Phys. Rev. B* **1**(8), 3304 (1970)
23. N.J.M. Horing, E. Kamen, H. Cui, Inverse dielectric function of a bounded solid-state plasma. *Phys. Rev. B* **32**(4), 2184 (1985)
24. G. Gumbs, A. Iurov, N.J.M. Horing, Non-local plasma spectrum of graphene interacting with a thick conductor. *Phys. Rev. B* **91**(23), (2015)
25. P.K. Pyatkovskiy, Dynamical polarization, screening, and plasmons in gapped graphene. *J. Phys.: Condens. Matter* **21**(2), 025506 (2009)
26. B. Wünsch, T. Stauber, F. Sols, F. Guinea, Dynamical polarization of graphene at finite doping. *New J. Phys.* **8**(12), 318 (2006)
27. E.H. Hwang, S. Das Sarma, Dielectric function, screening, and plasmons in two-dimensional graphene. *Phys. Rev. B* **75**(20), 205418 (2007)
28. M. Rocca, Low-energy investigation of surface electronic excitations on metals. *Surf. Sci. Rep.* **22**(12), 171 (1995)
29. A. Politano, G. Chiarello, G. Benedek, E.V. Chulkov, P.M. Echenique, Vibrational spectroscopy and theory of alkali metal adsorption and co-adsorption on single-crystal surfaces. *Surf. Sci. Rep.* **68**(3–4), 305–389 (2013)
30. A. Politano, G. Chiarello, The influence of electron confinement, quantum size effects, and film morphology on the dispersion and the damping of plasmonic modes in Ag and Au thin films. *Prog. Surf. Sci.* **90**(2), 144–193 (2015)
31. A. Politano, A.R. Marino, G. Chiarello, Effects of a humid environment on the sheet plasmon resonance in epitaxial graphene. *Phys. Rev. B* **86**, 085420 (2012)
32. A. Politano, A.R. Marino, V. Formoso, D. Fariás, R. Miranda, G. Chiarello, Evidence for acoustic-like plasmons on epitaxial graphene on Pt(111). *Phys. Rev. B* **84**, 033401 (2011)
33. A. Politano, V. Formoso, G. Chiarello, Evidence of composite plasmonphonon modes in the electronic response of epitaxial graphene. *J. Phys.: Condens. Matter* **25**(34), 345303 (2013)
34. T. Langer, D.F. Förster, C. Busse, T. Michely, H. Pfnür, C. Tegenkamp, Sheet plasmons in modulated graphene on Ir(111). *New J. Phys.* **13**(5), 053006 (2011)
35. G. Giovannetti, P.A. Khomyakov, G. Brocks, V.M. Karpan, J. van den Brink, P.J. Kelly, Doping graphene with metal contacts. *Phys. Rev. Lett.* **101**, 026803 (2008)

36. C. Gong, D. Hinojos, W. Wang, N. Nijem, B. Shan, R.M. Wallace, K. Cho, Y.J. Chabal, MetalGrapheneMetal sandwich contacts for enhanced interface bonding and work function control. *ACS Nano* **6**(6), 5381–5387 (2012)
37. K.L. Grosse, B. Myung-Ho, L.E.P. Feifei, W.P. King, Nanoscale Joule heating, Peltier cooling and current crowding at graphene-metal contacts. *ACS Nano* **6**(6), 5381–5387 (2012)
38. C. Gong, S. McDonnell, X. Qin, A. Azcatl, H. Dong, Y.J. Chabal, K. Cho, R.M. Wallace, Realistic metalgraphene contact structures. *ACS Nano* **8**(1), 642649 (2013)
39. A. Politano, G. Chiarello, Quenching of plasmons modes in air-exposed graphene-Ru contacts for plasmonic devices. *Appl. Phys. Lett.* **102**(20), 201608 (2013)
40. J.T. Smith, A.D. Franklin, D.B. Farmer, C.D. Dimitrakopoulos, Reducing contact resistance in graphene devices through contact area patterning. *ACS Nano* **7**(4), 3661–3667 (2013)
41. C. Archambault, A. Rochefort, States Modulation in Graphene Nanoribbons through Metal Contacts. *ACS Nano* **7**(6), 5414–5420 (2013)
42. P. Janthon, F. Viñes, S.M. Kozlov, J. Limtrakul, F. Illas, Theoretical assessment of graphene-metal contacts. *J. Chem. Phys.* **138**(24), 244701 (2013)
43. F. Menges, H. Riel, A. Stemmer, C. Dimitrakopoulos, B. Gotsmann, Thermal transport into graphene through nanoscopic contacts. *Phys. Rev. Lett.* **111**(20), 205901 (2013)
44. A. Politano, G. Chiarello, Unravelling suitable graphene-metal contacts for graphene-based plasmonic devices. *Nanoscale* **5**(17), 8215–8220 (2013)
45. V.M. Silkin, A. García-Lekue, J.M. Pitarke, E.V. Chulkov, E. Zaremba, P.M. Echenique, Novel low-energy collective excitation at metal surfaces. *EPL (Europhys. Lett.)* **66**(2), 260 (2004)
46. V.M. Silkin, J.M. Pitarke, E.V. Chulkov, P.M. Echenique, Acoustic surface plasmons in the noble metals Cu, Ag, and Au. *Phys. Rev. B* **72**(11), 115435 (2005)
47. M. van Schilfgaarde, M.I. Katsnelson, First-principles theory of nonlocal screening in graphene. *Phys. Rev. B* **83**(8), 081409 (2011)
48. Y. Gao, Z. Yuan, Anisotropic low-energy plasmon excitations in doped graphene: An ab initio study. *Solid State Commun.* **151**(14), 1009–1013 (2011)
49. F.M.D. Pellegrino, G.G.N. Angilella, R. Pucci, Dynamical polarization of graphene under strain. *Phys. Rev. B* **82**(11), 115434 (2010)
50. A. Politano, G. Chiarello, Emergence of a nonlinear plasmon in the electronic response of doped graphene. *Carbon* **71**, 176–180 (2014)
51. A. Politano, A.R. Marino, D. Campi, D. Fariás, R. Miranda, G. Chiarello, Elastic properties of a macroscopic graphene sample from phonon dispersion measurements. *Carbon* **50**(13), 4903–4910 (2012)
52. A. Politano, A.R. Marino, G. Chiarello, Phonon dispersion of quasi-freestanding graphene on Pt (111). *J. Phys.: Condens. Matter* **24**(10), 104025 (2012)
53. E.H. Hwang, S. Das Sarma, Dielectric function, screening, and plasmons in two-dimensional graphene. *Phys. Rev. B* **75**(20), 205418 (2007)
54. A.V. Gorbach, Nonlinear graphene plasmonics: amplitude equation for surface plasmons. *Phys. Rev. A* **87**(1), 013830 (2013)
55. M. Kauranen, A.V. Zayats, Nonlinear plasmonics. *Nat. Photonics* **6**(11), 737748 (2012)
56. V. Despoja, D. Novko, K. Dekanić, M. Šunjić, L. Marušić, Two-dimensional and p plasmon spectra in pristine and doped graphene. *Phys. Rev. B* **87**(7), 075447 (2013)
57. B. Borca, S. Barja, M. Garnica, M. Minniti, A. Politano, J. M. Rodríguez-García, J. J. Hinarejos, D. Fariás, A. L. Vásquez de Parga, R. Miranda, Electronic and geometric corrugation of periodically rippled, self-nanostructured graphene epitaxially grown on Ru (0001). *New J. Phys.* **12**(9), 093018 (2010)
58. M. Batzill, The surface science of graphene: Metal interfaces, CVD synthesis, nanoribbons, chemical modifications and defects. *Surf. Sci. Rep.* **67**(3), 83115 (2012)
59. J. Winterlin, M.L. Bocquet, Graphene on metal surfaces. *Surf. Sci.* **603**(10), 1841–1852 (2009)
60. Z. Fang, Y. Wang, Z. Liu, A. Schlather, P.M. Ajayan, F.H.L. Koppens, P. Nordlander, N. J. Halas, Plasmon-induced doping of graphene. *ACS Nano* **6**(11), 10222–10228 (2012)

61. H. Yan, X. Li, B. Chandra, G. Tulevski, Y. Wu, M. Freitag, W. Zhu, P. Avouris, F. Xia, Tunable infrared plasmonic devices using graphene/insulator stacks. *Nat. Nanotechnol.* **7**(5), 330–334 (2012)
62. S.Y. Shin, N.D. Kim, J.G. Kim, K.S. Kim, D.Y. Noh, K.S. Kim, J.W. Chung, Control of the p plasmon in a single layer graphene by charge doping. *Appl. Phys. Lett.* **99**(8), 082110 (2011)
63. M.K. Kinyanjui, C. Kramberger, T. Pichler, J.C. Meyer, P. Wachsmuth, G. Benner, U. Kaiser, Direct probe of linearly dispersing 2D interband plasmons in a free-standing graphene monolayer. *EPL (Europhys. Lett.)* **97**(5), 57005 (2012)
64. C. Kramberger, R. Hambach, C. Giorgetti, M.H. Rummeli, M. Knupfer, J. Fink, B. Büchner, Reining Lucia, E. Einarsson, S. Maruyama et al., Linear plasmon dispersion in single-wall carbon nanotubes and the collective excitation spectrum of graphene. *Phys. Rev. Lett.* **100**(19), 196803 (2008)
65. C. Tegenkamp, H. Pfnür, T. Langer, J. Baringhaus, H.W. Schumacher, Plasmon electronhole resonance in epitaxial graphene. *J. Phys.: Condens. Matter* **23**(1), 012001 (2011)
66. A. Cupolillo, N. Ligato, L.S. Caputi, Plasmon dispersion in quasifreestanding graphene on Ni (111). *Appl. Phys. Lett.* **102**(11), 111609 (2013)
67. G. Bertonni, L. Calmels, A. Altibelli, V. Serin, First-principles calculation of the electronic structure and EELS spectra at the graphene/Ni(111) interface. *Phys. Rev. B* **71**(7), 075402 (2005)
68. M. Hasegawa, K. Nishidate, T. Hosokai, N. Yoshimoto, Electronic-structure modification of graphene on Ni(111) surface by the intercalation of a noble metal. *Phys. Rev. B* **78**, 085439 (2013)
69. D. Farias, A.M. Shikin, K.H. Rieder, YuS Dedkov, Synthesis of a weakly bonded graphite monolayer on Ni(111) by intercalation of silver. *J. Phys.: Condens. Matter* **11**(43), 8453 (1999)
70. A.M. Shikin, D. Farias, V.K. Adamchuk, K. Rieder, Surface phonon dispersion of a graphite monolayer adsorbed on Ni(111) and its modification caused by intercalation of Yb La and Cu layers. *Surf. Sci.* **424**(1), 155–167 (1999)
71. A. Politano, V.M. Silkin, I.A. Nechaev, M.S. Vitiello, L. Viti, Z.S. Aliev, M.B. Babanly, G. Chiarello, P.M. Echenique, E.V. Chulkov, Interplay of surface and Dirac plasmons in topological insulators: the case of Bi_2Se_3 . *Phys. Rev. Lett.* (2015)

Chapter 10

Semiconducting Carbon Nanotubes: Properties, Characterization and Selected Applications

Chiara Pintossi and Luigi Sangaletti

Abstract Carbon nanotubes are challenging materials from the point of view of nanotechnology, because of their peculiar electrical, mechanical and optical properties arising from their monodimensional geometry. Here, the properties of carbon nanotubes are discussed, starting from their crystalline structure, in order to understand their optical, electrical and vibrational behavior. In the second section, the most popular CNT synthesis mechanism are presented, while the last section is devoted to the CNTs applications, focusing on photovoltaic and gas sensor devices.

10.1 Introduction to Carbon Nanotubes

Even though carbon filaments with a diameter of less than 10 nm were produced in 70s and 80s [1–5], it was only after the first observation with high resolution transmission electron microscopy (HRTEM) of carbon nanotubes (CNTs) by Iijima in 1991 [6] that CNT field boosted up significantly, attracting more and more scientist all over the world.

The simplest and fundamental type of CNTs is represented by single walled carbon nanotubes (SWCNT) which can be schematized as a single graphene sheet, rolled up to form a cylinder.

Because the tube properties [7] are dramatically influenced by the way in which the graphene sheet is rolled, it's worth deeply understanding the geometry of nanotubes and their relationship with graphene crystalline structure.

Starting from the basis vectors of the graphene hexagonal honeycomb Bravais lattice, $\mathbf{a}_1$ and $\mathbf{a}_2$ represented in Fig. 10.1, it is possible to express the circumference

C. Pintossi · L. Sangaletti (✉)
I-LAMP, Interdisciplinary Laboratory for Advanced Materials Physics,
Università Cattolica del Sacro Cuore, Brescia, Italy
e-mail: sangalet@dmf.unicatt.it

C. Pintossi
e-mail: chiarapintossi@gmail.com

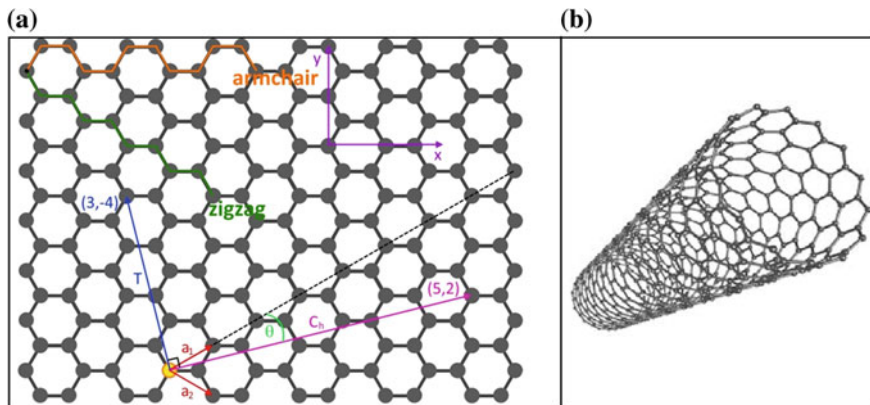


Fig. 10.1 **a** Schematic representation of the graphene Bravais lattice in which the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{C}_h$ and $\mathbf{T}$ are highlighted for a (5, 2) SWCNT. **b** A SWCNT obtained by rolling a graphene monolayer

of any carbon nanotube in term of the chiral vector $\mathbf{C}_h$, which is defined as the linear combination of the basis vectors: $\mathbf{C}_h = n\mathbf{a}_1 + m\mathbf{a}_2$. The pair of integer (n, m) uniquely defines the chiral vector and is often called simply nanotube chirality. Nanotubes with $m = 0$ are called *zigzag*, if $m = n$, the tube is denoted as *armchair*, while all the other CNTs are simply called *chiral*.

The other unit lattice vector $\mathbf{T}$, perpendicular to the chiral vector can be also expressed in terms of $n, m, \mathbf{a}_1$ and $\mathbf{a}_2$ as:

$$\mathbf{T} = \frac{(2m+n)}{b_R}\mathbf{a}_1 - \frac{(m+2n)}{b_R}\mathbf{a}_2$$

where b_R is equal to $3b$, if $m - n$ is a multiple of $3b$, or $b_R = b$, when $m - n$ is not a multiple of $3b$, with b being the greatest common divisor between m and n .

The tube diameter d is given by:

$$d = \frac{C_h}{\pi} = \frac{\sqrt{3}a_{C-C}}{\pi} \sqrt{m^2 + mn + n^2}$$

where a_{C-C} is the length of the carbon-carbon bond length (1.42 Å), i.e. the hexagon side in Fig. 10.1.

The chiral angle θ , defined as the angle between $\mathbf{C}_h$ and $\mathbf{a}_1$, is:

$$\theta = \tan^{-1} \left(\frac{\sqrt{3}n}{2m+n} \right)$$

and varies from $\theta = 0$ for *zigzag* nanotubes to $\theta = 30^\circ$ for the *armchair* ones.

The number of hexagons in each unit cell is given by:

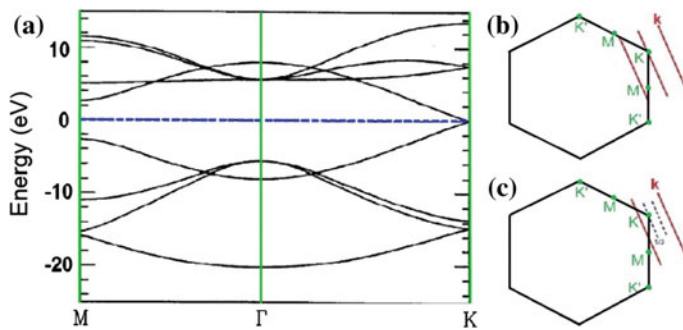


Fig. 10.2 **a** Tight-binding band structure of graphene. The wave vector $\mathbf{k}$ for a 1D CNT is shown in the 2D Brillouin zone of graphite, in red, for **b** metallic and **c** semiconducting tubes

$$N = \frac{2}{b_R} (m^2 + mn + n^2),$$

it follows that the CNT unit cell is N time larger than that of graphene.

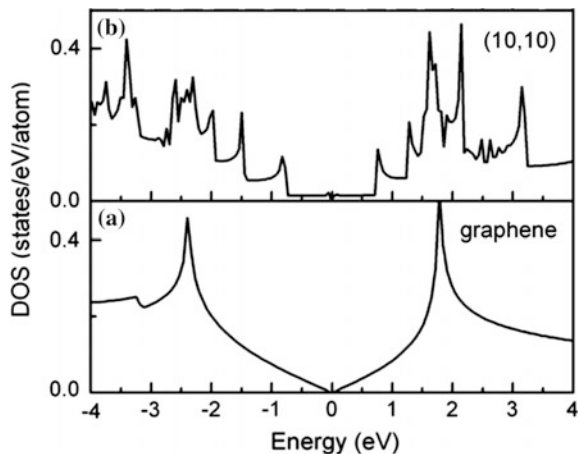
Theoretical studies predicted CNT electronic properties to be very sensitive to their geometrical structure. In fact, although graphene is a zero-gap semiconductor, single walled carbon nanotube can be either metallic or semiconducting, depending only on their chiral index (n, m) .

The electronic structure around the Fermi level of an isolated graphene sheet is given by an occupied π -band and an unoccupied π^* -band, characterized by having both linear dispersion and by crossing the Fermi level at the K point of the Brillouin zone (Fig. 10.2a). When a graphene sheet is rolled to form a tube, only some $\mathbf{k}$ states are allowed, due to the periodicity boundary conditions imposed in the circumferential direction. Whenever the set of allowed $\mathbf{k}$ states includes the K point (Fig. 10.2b), the system has a non-zero density of states (DOS) at the Fermi level and is then a metal with two linear dispersing bands; otherwise, when the K point is located at one-third of the distance between two adjacent allowed $\mathbf{k}$ (Fig. 10.2c), the CNT is a semiconductor with a variable energy band gap. Nanotubes with $n - m = 3b$, where b is an integer, are metallic, all the other are semiconductors. Each parallel $\mathbf{k}$ -line describes one pair of conduction and valence sub-bands in the nanotube. Transitions at the bandgaps of such sub-band pairs lead to strong optical resonances, and these transitions are traditionally labelled S_{ii} for semiconducting and M_{ii} for metallic nanotubes, where i is the sub-band index.

Multi-walled nanotubes (MWNTs) consist of multiple rolled layers (concentric tubes) of graphene in a Russian doll model. Its individual shells can be described as SWNTs, which can be metallic or semiconducting. Because of statistical probability and restrictions on the relative diameters of the individual tubes, a MWNT is usually a zero-gap metal.

The 1D density of state of CNT, derived straightforward from the graphite one [8], becomes large when the energy dispersion relation becomes flat as a function of

Fig. 10.3 The density of state of graphene (a) is compared to the CNTs one (b). The 1D confinement in CNTs gives rise to the characteristic van Hove singularities



k , giving rise to one-dimensional van Hove singularities (vHs) in the DOS, as shown in Fig. 10.3. VHs determine many solid state properties of carbon nanotubes, which affect the spectra observed by scanning tunneling spectroscopy (STS), optical absorption, and resonant Raman spectroscopy.

By tight binding calculation a simple relation between nanotube diameter d and the transition energies $E_{sc,m}$, the energy distance between the van Hove singularities, valid both for semiconducting (sc) and metallic (m) SWCNT can be defined as:

$$E_{sc,m} = 2i \frac{\gamma_0 a_{C-C}}{d}$$

with $i = 1$ for S_{11} , $i = 2$ for S_{22} and $i = 3$ for M_{11} ; $\gamma_0 = 2.9$ eV is the nearest neighbour carbon–carbon interaction energy and $a_{C-C} = 0.144$ nm is the nearest neighbour carbon–carbon distance. In UV/Vis/NIR absorption spectroscopy [9], the optical transitions, arising from vHs, are probed. By absorption of light, the electrons in the van Hove singularities of the valence band jumps to the corresponding energy levels in the conduction band (the S_{11} and S_{22} in semiconducting tubes). In principle, the same process happens in the case of metallic SWCNTs. However, the spacing between the van Hove singularities is larger so that only M_{11} transitions can be probed in the UV/Vis region (Fig. 10.4).

Also in the phonon DOS, carbon nanotubes exhibit other unusual aspects regarding their phonon dispersion relations [10], such as four acoustic branches. In addition to the longitudinal acoustic and transverse acoustic modes, there are two acoustic twist modes for rigid rotation around the tube axis, which are important for heat transport and charge carrier scattering. In SWNT Raman spectroscopy, the G-band spectra, present also in graphene, which is split into many features around 1580 cm^{-1} , and the lower frequency radial breathing mode (RBM) are usually the strongest features, and are both first-order Raman modes. The RBM is a unique phonon mode, appearing only in carbon nanotubes and its observation in the Raman

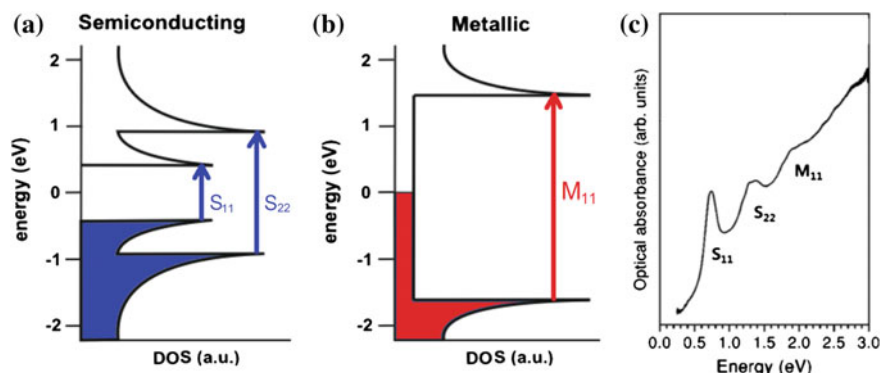
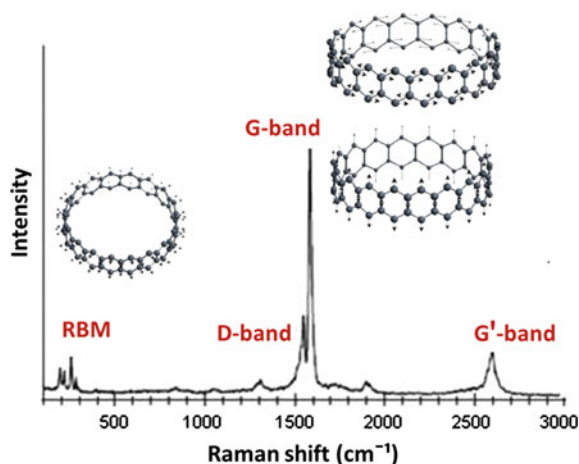


Fig. 10.4 The density of state of semiconducting (a) and metallic (b) are sketched, highlighting the S_{11} , S_{22} and M_{11} transitions. These transitions give rise to the three peaks in the optical absorption spectrum (c)

Fig. 10.5 Raman spectrum of a SWCNT. The vibrational modes associated to RBMs and the G-band are represented



spectrum provides direct evidence of the presence of SWNTs. The RBM is a bond-stretching out-of-plane phonon mode for which all the carbon atoms move coherently in the radial direction, and whose frequency is about 100–500 cm^{-1} . The RBM frequency is inversely proportional to the tube diameter.

Carbon materials presents also two relatively weak Raman signals whose phonon frequencies change with different laser excitation energy, which is called “dispersive” behavior. In graphite as well as in SWNTs, the D-band at 1350 cm^{-1} and the G'-band at 2700 cm^{-1} , excited by a laser photon energy of 2.41 eV, are, respectively, due to one- and two-phonon, second-order Raman scattering processes. In particular, because the D-band was firstly observed in the presence of defects in an otherwise perfect infinite graphite structure, it has been associated to structural defects in the sample (Fig. 10.5).

10.2 CNTs Synthesis

Since their discover, a lot of efforts have been made in the optimization of CNTs synthesis techniques, in order to produce large quantities of pure, diameter sorted tube, with low fabrication costs [11]. Even though most of these processes usually take place in vacuum, new techniques have recently been developed for synthesizing CNTs at atmospheric pressure in order to make nanotubes more commercially viable.

In this paragraph, arc discharge, laser ablation, plasma torch and chemical vapor deposition will be presented and discussed with more details.

- **Arc discharge** [12] has been the first CNT production method used since 1991. In fact, nanotubes were first observed in the carbon soot of graphite electrodes during an arc discharge that was aimed to produce fullerenes. In this method two carbon rods are placed end to end, separated by approximately 1 mm, in an enclosure that is usually filled with helium at relatively low pressure (between 50 and 700 mbar). A direct current of 50–100 A driven by approximately 20 V creates a high temperature (~ 4000 K) discharge between the two electrodes. The discharge vaporizes the anode carbon rod and forms a small rod shaped deposit on the cathode. The yield for this method is up to 30 % by weight and it can be used to produce both single- and multi-walled nanotubes with lengths of up to 50 μm with a low number of structural defects. For producing single-walled CNTs, metal catalysts, such as Fe, Co and Ni, are needed. The SWCNTs diameter distribution depends on the composition of the catalyst and on the growth conditions. If no catalysts are used, mostly MWCNTs are produced with a typical outer diameter of approximately 10 nm (Fig. 10.6)
- **Laser ablation** [13]: a pulsed laser vaporizes a graphite target, placed in a furnace at roughly 1200 $^{\circ}\text{C}$, containing small amounts of a metal catalyst. During laser ablation, a flow of inert gas flows through the growth chamber to carry the grown nanotubes to be collected on a cooler surface. The produced nanotubes are mostly in the form of ropes (bundles) consisting of tens of

Fig. 10.6 Schematic drawing of an arc discharge apparatus for CNTs production

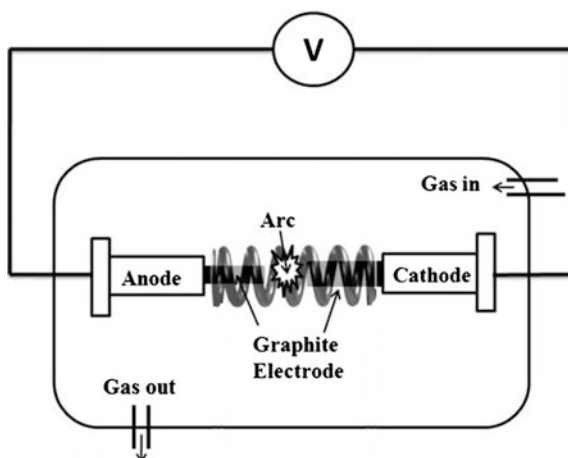
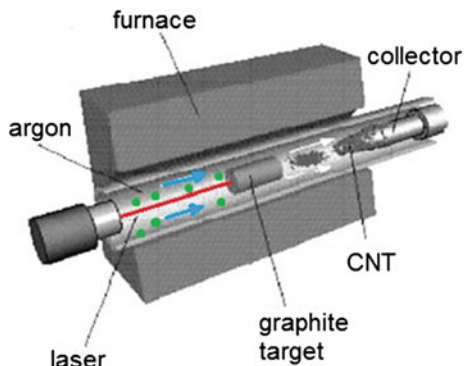


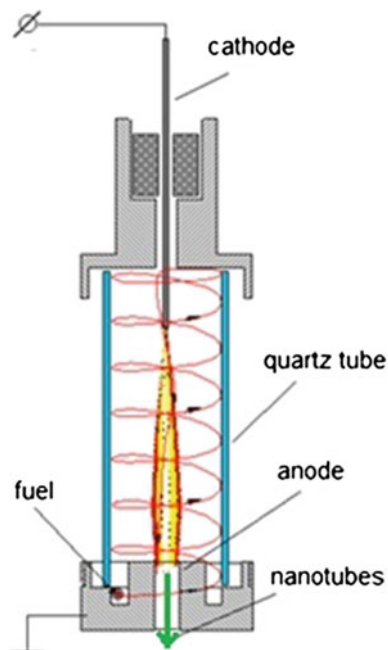
Fig. 10.7 Representation of a laser ablation system



individual nanotubes close-packed together via van der Waals forces. The yield of nanotube synthesis by this process is roughly 70 % (Fig. 10.7).

- Plasma torch** [14] the plasma torch operation conditions are similar to those used in the arc discharge and laser ablation approaches, but, instead of graphite vapors, a carbon-containing gas is used to supply the carbon necessary for the production of SWNT. In this way, the growth of SWNT is more efficient because decomposing a carbon containing gas can be 10 times less energy-consuming than graphite vaporization. To produce a continuous process, a gas mixture composed of argon, ethylene and ferrocene is introduced into a plasma torch, where it is atomized by the atmospheric pressure plasma, which has the form of an intense flame. The fumes created by the flame are found to contain SWNT, amorphous carbon as well as metallic and carbon nanoparticles (Fig. 10.8).
- Chemical vapor deposition (CVD)** [15] is nowadays the most popular method of producing CNTs. In this process, thermal decomposition of a hydrocarbon vapor is achieved in the presence of a metal catalyst. The process involves the passage of an hydrocarbon vapor through a tubular reactor in which a catalyst material is present at a temperature sufficiently high (600–1200 °C) to decompose the hydrocarbon. CNTs grow on the catalyst in the reactor and are collected upon cooling the system to room temperature. In the case of a liquid hydrocarbon (benzene, alcohol, etc.), the liquid is heated in a flask and an inert gas is purged through it, which in turn carries the hydrocarbon vapor into the reaction zone. If a solid hydrocarbon is used, it can be directly kept in the low-temperature zone of the reaction tube. Volatile materials (camphor, naphthalene, ferrocene etc.) directly turn from solid to vapor, and perform CVD while passing over the catalyst kept in the high-temperature zone. Like the CNT precursors, also the catalyst precursors in CVD may be used in any form: solid, liquid or gas, which may be suitably placed inside the reactor or fed from outside. Pyrolysis of the catalyst vapor at a suitable temperature liberates metal nanoparticles in situ (such a process is known as floating catalyst method). Alternatively, catalyst-coated substrates can be placed in the hot zone of the furnace to catalyze the CNT growth.

Fig. 10.8 Schematic drawing of a plasma torch system



Compared to arc-discharge and laser-ablation methods, CVD is a simple and economic technique for synthesizing CNTs at relatively low temperature and ambient pressure. In the case of MWCNTs, arc- and laser-grown tube possess better crystallinity compared to the CVD-grown ones, but the crystallinity of SWCNTs grown by CVD is close to that grown by arc or laser methods. CVD result to overcome arc and laser methods in yield and purity, reaching the best performances in terms of structure control and CNT architecture. CVD is versatile because it works with various type hydrocarbons in any state (solid, liquid or gas), enables the use of various substrates, and allows CNT growth in a variety of forms, such as powder, thin or thick films, aligned or entangled, straight or coiled nanotubes with the possibility of controlling the growth parameters (Fig. 10.9).

10.3 Carbon Nanotubes Applications

10.3.1 *Selected Applications of Semiconducting CNTs*

So far CNTs have been tested for a considerable number of applications spanning quite different areas from nanoelectronics to biosensing [16–24]. In particular, they have been considered in micro- and nanoelectronics, photonics, in the field of solar cells, for hydrogen and energy storage, as chemiresistors in gas sensing and

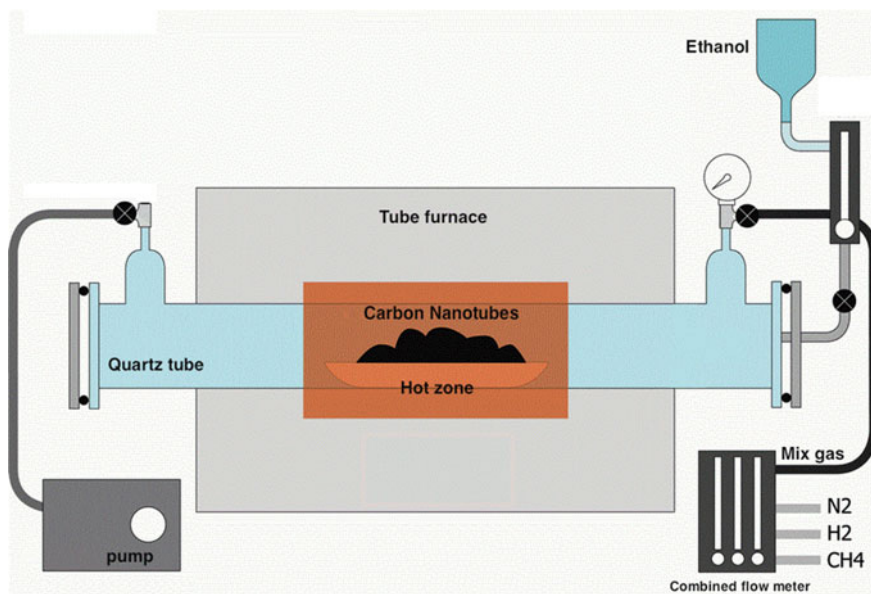


Fig. 10.9 Schematic representation of a chemical vapor deposition (CVD) chamber

environmental monitoring, and as structural materials in composites. Among these applications, nanoelectronics is most probably the one that merges issues related to basic physics (e.g. transport in 1D systems) to the capability to design, engineer, produce, and test novel architectures for electronic components at the nanoscale. A lot of effort was focused on the capability to produce single-wire FET devices. It is interesting to note that FET-based devices have also been proposed in the field of gas sensing [25], showing that cross-fertilization among different research areas is a process that CNT can enable due to their versatility as functional materials.

In the following, two applications are discussed, namely the use of SWCNT in hybrid solar cells, and the development of gas sensors based on SWCNT layers. These application share as a common feature the relatively easy preparation route of the active SWCNT layer and the possibility to achieve quite good results without resorting to control the layer at the nanoscale. Indeed even with CNT bundles, devices can be brought to operation, enabling the test of novel ideas and device architectures at relatively low cost.

10.3.2 CNTs for Photovoltaic Applications

In recent years, CNTs have played different roles in photovoltaic devices. Nanotubes are in fact used as a part of the photoactive layer, as a transparent electrode or in dye-sensitized solar cells, but among the most promising alternatives

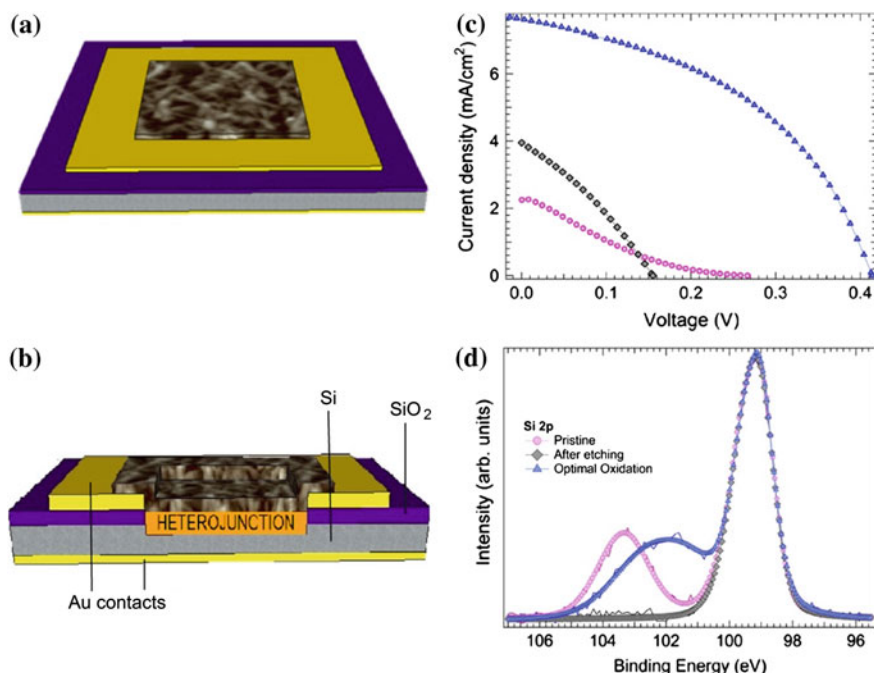


Fig. 10.10 A sketch of a hybrid CNT/Si solar cell (a) and its section view (b), with the CNT film on *top* (in gray), two Au contact (in yellow), n-type silicon in silver and silicon oxide in purple. J-V curves (c) and Si2p XPS peaks (d) are presented for different oxidation time: pristine sample (pink circles), without oxides (gray diamonds) and at the optimal oxidation state (blue triangles)

to silicon photovoltaics (PV) there are hybrid devices based on heterojunctions in which the p-type silicon is completely replaced by SWCNT, which are naturally p-doped in air by molecular oxygen (Fig. 10.10a, b).

The remarkable power conversion efficiencies (PCE) so far displayed by devices based on these junctions [26, 27] are strictly related to the peculiar physical properties of CNTs. On one side, single-SWCNT are excellent candidates for participating in photogeneration processes, because of their capability of absorbing light over a wide range of the solar spectrum, thanks to an energy bandgap variable from 0 to 3 eV, depending on the tube diameter and to the presence in their DOS of the vHs. On the other side, it has been recently proved [28] that SWCNTs do not merely act as a light absorber layer but, due to their extraordinary electrical conductivity ($\sim 10^6$ S/cm) [29], they are also able to provide efficient transport paths for carriers, which have a very high probability to reach electrodes and generate photocurrent in devices based on CNTs photoexcitation processes [30].

A considerable effort has been recently made to increase the efficiency of CNT-Si heterojunctions, until reaching a record PCE of about 15 % [31]. However, much has still to be done in order to identify the mechanisms behind the device operation and determine the parameters that mostly affect the cell behavior. In

addition to a number of technical issues related to the cell geometry, to contacting pads, and to the choice of the CNT layer, several studies have pointed out [32, 33] that a silicon oxide layer grows at the interface between monocrystalline silicon and CNTs not only during the manufacturing process but also once the junction is built and that the properties of this layer affect the overall cell performances.

The role played by the oxide layer is of fundamental importance for understanding the working mechanism behind these devices that could behave either as p-n junctions [34] or as a metal-insulator-semiconductor (MIS) devices [35]. In particular, in a p-n junction model, the best efficiency should be reached once all the insulating silicon oxides are removed and perfect junctions between CNTs and Si can be built. On the other hand, in MIS junctions a relatively low potential barrier between CNTs and silicon yields a high number of majority carriers recombinations, resulting in a saturation current larger than that of a p-n junction diode and, as a consequence, in a low open circuit voltage. In these devices a thin layer of insulating oxide is beneficial because it helps to confine electrons in silicon, avoiding them to fast recombine in CNTs [36] (Fig. 10.10c). Moreover, a thin insulating layer may also have the advantage to passivate interface states, hindering charge trapping, lowering recombinations and increasing the device open circuit voltage.

The choice between the two junction schemes (i.e. p-n vs. MIS) requires a specific investigation on the oxide role through direct evidence of the Si oxidation states during manufacturing and ageing processes to tailor the oxide layer properties.

Oxide removal from silicon upon HF exposure is expected, as it is a standard treatment to etch silicon wafers and, on this basis, similar results have been claimed when CNT/Si heterojunctions were HF etched [37]. The possibility to access the buried interface layer by angle resolved X-ray photoemission spectroscopy (AR-XPS) has recently been demonstrated (Fig. 10.10d) and the existence of an oxide layer with optimal thickness was inferred by several groups [32, 33, 38] to rationalize the behavior of the cells in the presence or absence of oxides. Therefore, in spite of a seemingly straightforward preparation route, the CNT/SiO_x/Si interface represents a quite complex system from the point of view of materials, as it matches a relatively low-density CNT bundle layer with a nanostructured and chemically inhomogeneous SiO_x layer grown on a n-doped silicon wafer.

In particular, in devices based on bundles of randomly aligned SWCNT the presence of a thin layer of oxides has always lead to better cell performances [32, 33, 38].

As shown by Jia et al. [32], the electrical parameters of a CNT/Si solar cell can be modulated by changing the thickness of the SiO_x layer, either by removing oxides with HF or by letting these oxides regrow, through either oxidation in ambient air or exposure to nitric acid.

For understanding the operation mechanism, a combination of suitable spectroscopic tools has been used recently [39]: AR-XPS for a direct probe of the buried oxide interface, Raman spectroscopy to obtain information on CNT properties, and

the acquisition of current density-voltage (J-V) characteristics for monitoring the behavior of the photovoltaic device after exposure to acid.

Based on this experimental approach, It has been demonstrated that a relatively mild HF etching does not change the CNT film properties, acting only as silicon oxides etcher. This finding addresses a long-standing question on the unaligned CNT-Si hybrid junctions, probing the role of the complex silicon oxide interface to tailor the PV cell performances. The build-up of an interface Si oxide layer is required to increase the efficiency. This is assumed as a strong indication that the junction can be described with a MIS model rather than a p-n junction. These results indicate that the SiO_x buried interface can be regarded as a passivating and inversion layer, which improves the cell performances by hindering the e-h pair recombination.

10.3.3 CNT Interaction with Gases: From Surface Chemistry to Devices

Up to now CNT have been tested against many polluting gases, as they are known to interact with many gas molecules that, depending on their reducing or oxidizing properties, may inject or extract electrons from the CNTs, resulting in a detectable electrical signal. Indeed, a large body of studies reports the response of CNTs to many gases, in particular to ammonia and nitrogen dioxide, as prototypes of reducing and oxidizing gas, respectively. Many CNT-based gas sensors devices have been tested so far, showing that CNTs are a promising class of materials for the development of gas sensors [40–53].

The gas sensing capability of CNT is based on the possible interactions of molecules with the CNT walls. The basic interaction mechanism can be clearly evidenced in semiconducting CNT. It has been shown by Collins et al. [54] that the presence of oxygen induces significant effects on the conductivity and the DOS of CNTs. Indeed, CNTs readily become p-doped when exposed to the oxygen molecules present in air and therefore the interaction with molecules capable to transfer charge to and from CNTs can increase or decrease the number of carries (holes), depending on the reducing or oxidizing properties of the target gas molecules. These interactions have been explored in experimental [41, 54–57] and theoretical works using ab initio calculations and density functional theory (DFT) [58–60]. The distribution of electrons around the CNT make it electrochemically active. Therefore when oxidizing or reducing gas molecules, e.g. NO_2 or NH_3 respectively, are adsorbed onto carbon nanotubes surface, the charge transfer affects the electronic properties in terms of band structure, DOS, and electrical conductivity of CNTs. Many studies have been inspired by the experimental evidences presented in [55], where the variation of the electrical properties of CNTs upon exposure to NH_3 and NO_2 was clearly evidenced. On the theoretical side, one of the first works based on DFT was performed by Zhao et al. [58] who

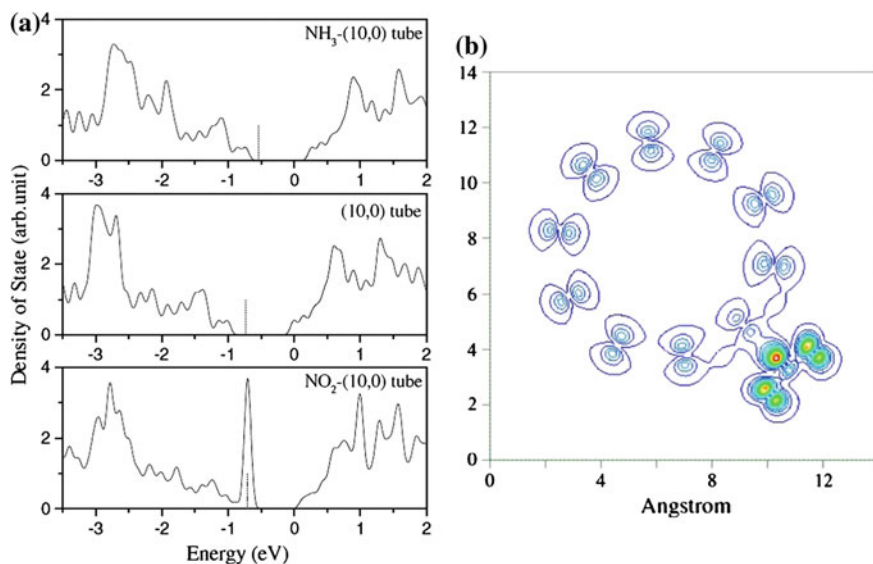


Fig. 10.11 Electronic DOS of (10, 0) carbon nanotubes after adsorption of NH_3 (top panel), and NO_2 (bottom panel) compared with the DOS of (10, 0) pristine CNTs. Dashed lines denote the Fermi level. **b** Contour plot of the electron charge density (slice passing through SWCNT and NO_2 molecule) [58]

studied the interaction of a large set of molecules (NO_2 , O_2 , H_2O , NH_3 , CH_4 , CO_2 , H_2 , N_2 , Ar) and SWCNTs with different chirality, clarifying the mechanisms of charge transfer, the adsorption sites, and the CNT-molecule minimum distance. The gas molecule adsorption on the SWCNT surface may induce a significant variation of the electronic properties due to the charge transfer. In Fig. 10.11, the density of electronic states of pure (10, 0) carbon nanotubes and upon absorption of NH_3 and NO_2 (respectively in middle, top, and bottom panel) is shown. The NH_3 effect is predicted to be weaker with respect to NO_2 interaction which shifts the Fermi level into the valence band. Nevertheless, detection of NH_3 is still a benchmarking measure to test the capability of CNT to interact with small reducing molecules.

In turn, the strong interaction of NO_2 with CNTs may result in poorly reversible interactions, leaving the CNT layer poisoned by NO_2 and therefore, to some extent, less capable to restore the pristine electrical parameters and to proceed with the detection of further exposure to this target molecule.

According to Zhao et al. [58] for molecules like H_2O , NH_3 , CO_2 , which are donors, i.e. they inject electrons on the nanotube surface, the charge transfer towards the tube is low (0.01–0.035 electron per molecule) and the adsorption energy defined as $E_a(d) = E_{\text{tot}}(\text{tube} + \text{molecule}) - E_{\text{tot}}(\text{tube}) - E_{\text{tot}}(\text{molecule})$ is weak (0.2 eV), d being the equilibrium distance between the tube and the molecule, while for the NO_2 acceptor which induces high charge transfer from the tube (0.06–0.14 electron per molecule), the adsorption energies are also greater (0.4 eV).

In order to exemplify the interaction between gas molecules and bundle of SWCNTs, in Fig. 10.12 we report the case of an H_2 molecule adsorbed on (10, 10) SWCNTs bundle considered by Zhao et al. [58]. As one can observe from the table on the right side of Fig. 10.12, the adsorption energy and charge transfer in the interstitial and groove sites are greater with respect to that on the surface sites. This can be explained by considering the greater number of carbon nanotubes that interact with the molecule. It is important to highlight that the possibility to be adsorbed in interstitial sites is reduced for large molecules.

From the experimental point of view, reference measurements have been presented by Goldoni et al. [41], based on X-ray photoemission spectroscopy (XPS). In particular, the interaction between SWCNTs and gas molecules like SO_2 , NH_3 , O_2 or NO_2 has been studied. The authors show that the reaction with oxygen is not due to carbon nanotube interaction, but to the presence of residual contaminants of the purification process (especially the species containing Na). Once these contaminants are completely removed, SWCNTs appear not very sensitive to O_2 , CO , H_2O e N_2 , while a strong sensitivity to NO_2 , SO_2 , and NH_3 is still observed. In Fig. 10.13 XPS C 1s spectra of the SWCNT clean sample, i.e. without

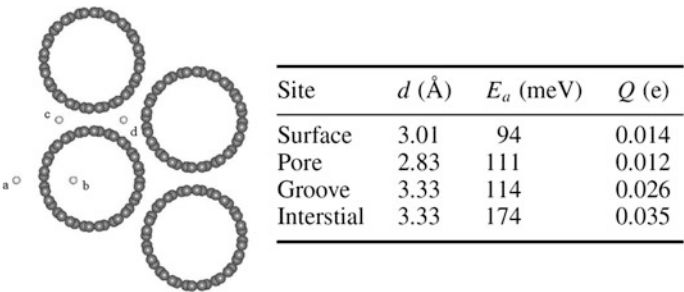


Fig. 10.12 Illustration of possible adsorption sites for the H_2 molecule in the (10, 10) SWCNTs bundle. **a** Surface; **b** pore; **c** groove; **d** interstitial. In the table on the right, the equilibrium tube-molecule distance (**d**), adsorption energy (E_a) and charge transfer (Q) on the different adsorption sites are reported [58]

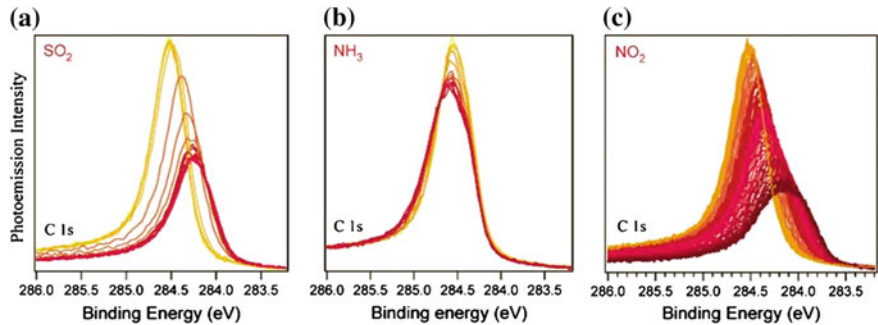


Fig. 10.13 C 1s spectra of the clean SWCNT bucky-paper measured while exposing the sample at 150 K to SO_2 (a), NH_3 (b), and NO_2 (c) at partial pressures of about 10^{-8} mbar [41]

contaminants, during gas exposure of SO_2 , NH_3 and NO_2 are shown. The shift of C 1s peak during gas exposure is toward lower binding energies in the case of SO_2 and NO_2 (Fig. 10.13a, c), while during NH_3 exposure (Fig. 10.13b) the shift is smaller and toward higher binding energies. If these binding energy shifts are ascribed to shifts of the nanotube chemical potential, the NO_2 and SO_2 molecules bound to SWCNTs act as charge acceptors, while the NH_3 molecules act as charge donors, in full agreement with the experimental observations reported by Kong et al. [55]. and theoretical results [58]. The results of [41], clearly address one typical problem of CNT synthesis, i.e. the capability to control contaminations related to the CNT growth and sample preparation, mostly in terms of residual impurities from metal catalysts, as well as surfactants used to disperse CNTs and prepare thin layers.

In addition to the above mentioned capability to interact with target gas molecules, SWCNTs soon resulted to be appealing systems to be used as gas sensors for three other reasons at least: the huge surface to volume ratio, the remarkable transport properties that favor charge drift towards the electrodes, and their one-dimensional nature that makes them very sensitive to very small external perturbations [51]. Furthermore, gas sensors based on SWCNTs are able to operate at room temperature, coping with constraints on power consumption in portable devices. In fact, the use of nanostructured metal oxide layers in gas sensing (so far the most diffuse materials in chemiresistors) requires to operate at relatively high temperatures (few hundreds of $^\circ\text{C}$) by coupling the sensing layer with suitable resistive heaters.

As discussed above, gas molecules adsorbed on the carbon nanotube surface could act as dopants (donors or acceptors), they could shift the Fermi level or change the band structure of SWCNT. Such a change may be easily detected, once the nanotubes are inserted within an electronic device. The different types of electrochemical sensors based on SWCNTs electrical properties (e.g. resistance [44, 55, 61–63], capacitance [64] or frequency [65] variation, ionization chamber with CNTs [66]) can be grouped in: sensors based on impedance, piezoelectric transducer-based sensors and solid state miniaturized sensors [67]. The chemiresistor and the chemFET (chemical Field Effect Transistor) belong to the solid state miniaturized sensors category. In Fig. 10.14, the CNT-based devices and the electronic circuit schemes in the chemiresistor and chem-FET configuration are shown.

In the chemiresistor configuration, the CNTs create a bridge between two electrodes (to which a certain potential is applied), allowing the current flow. When a reactive gas molecule is adsorbed on the nanotube surface, the charge transfer alters the electrical resistance of the CNTs.

The chemFET configuration is similar to the chemiresistor configuration, except for the fact that the current flowing through the CNTs between source (S) and drain (D) electrodes is modulated by a third electrode, the gate (G), separated from the other two by a thin layer of insulating material (typically SiO_2). This metal/CNT semiconductor/metal system has the characteristics of a p-type transistor [55]. An

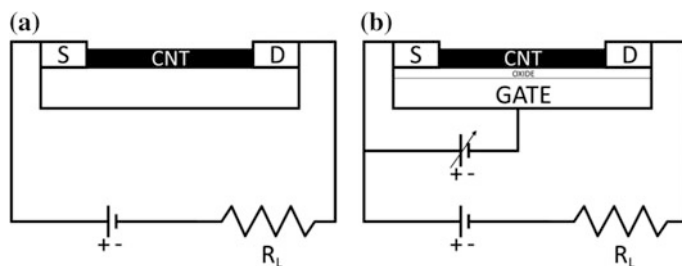


Fig. 10.14 CNT device and electronic circuit scheme in the chemiresistor (a) and chem-FET (b) configuration

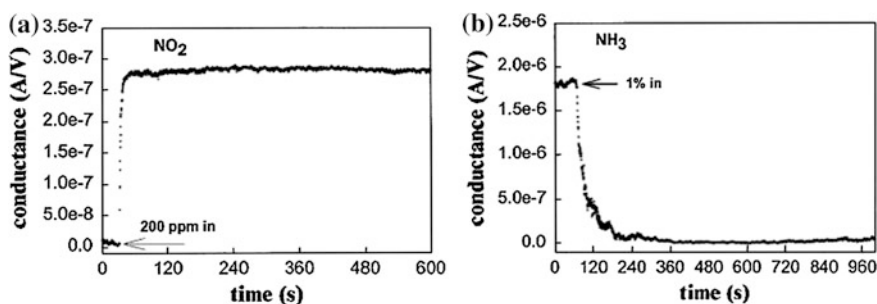


Fig. 10.15 Electrical response of a semiconducting SWCNT to gas molecules in a chem-FET configuration, but at fixed gate voltage (Figure adapted from [55]). **a** Conductance (under $V_{\text{gate}} = +4$ V, in an initial insulating state) versus time in a 200 ppm NO_2 flow. **b** Conductance ($V_{\text{g}} = 0$ V, in an initial conducting state) versus time recorded with the same SWCNT sample in a flow of Ar containing 1 % NH_3 . These conductance variations to oxidizing (i.e. NO_2) or reducing (i.e. NH_3) molecules indicate the p-type character of the SWCNT response

example of response to NO_2 and NH_3 exposure of a CNT-based chemFET is shown in Fig. 10.15.

In Table 10.1 a summary of pristine CNT chemiresistors and ChemFETs sensing performance is shown [50]. The chemFET configuration tends to be more sensitive to gas because of the possibility to modulate the conductance of CNTs via the gate voltage. However it requires more sophisticated electronics and much higher costs [44]. Chemiresistors are often chosen for the gas sensing measurements because they are simpler to produce and to manage. For this type of sensors interdigitated electrodes (IDE) are often used to improve charge collection from the CNT layer. The main techniques to deposit CNTs are: direct CVD on the substrate; deposition of a solution in which the CNTs are dispersed, e.g. by drop-casting method or ink-jet printing. Finally, in order to align CNTs between the electrodes, a method called dielectrophoresis, which consists on applying an AC electric field between the electrodes during the deposition process, has been used by Suheiro et al. [63] to build a MWCNT-based ammonia gas sensor. Usually the signal is processed as

Table 10.1 Summary of selected sensing performance of pristine CNT chemiresistors and chemFETs (N/S = Not-stated) [50]

CNT type	Sensor configuration	Targeted analytes	Detection limit	Response time (s)	Reversibility
Single SWNT	ChemFET	NO ₂ , NH ₃	2 ppm (NO ₂) 0.1 % (NH ₃)	<600	Irreversible
SWNTs	ChemFET	Alcoholic vapors (methanol, ethanol, 1-propanol, 2-propanol and tert-butanol.)	N/S	5–150	Reversible (–20 V gate bias potential)
SWNTs	Chemiresistor	O ₂	N/S	N/S	Reversible
SWNTs	ChemFET	DMMP	<1 ppb	1000	Reversible (3 V gate bias)
MWNTs	Chemiresistor	NO ₂	10 ppb	N/S	Reversible (165 °C)
SWNTs	Chemiresistor	NO ₂ , Nitrotoluene	44 ppb (NO ₂) 262 ppb (Nitrotoluene)	600	Reversible (UV)
MWNTs	Chemiresistor	NH ₃	10 ppm	~ 100	Reversible
SWNTs	Chemiresistor	SOCl ₂ , DMMP	100 ppm	10	Irreversible
SWNTs	Chemiresistor	O ₃	6 ppb	<600	Reversible
MWNTs	Chemiresistor	NO ₂	5–10 ppb	–600 (165 °C)	Reversible (165 °C)
SWNTs	Chemiresistor	Methanol, acetone	N/S	~ 100	N/S
SWNTs	Chemiresistor	H ₂ O	N/S	10–100	Reversible
Carboxylated SWNT	Chemiresistor	CO	1 ppm	~ 100	Reversible

$\Delta R/R_0$, (where $\Delta R = R - R_0$, R being the sensor resistance measured during exposure to gas and R_0 the baseline resistance before gas exposure) and the capability of the sensor to track variations in the target gas concentration is described by the calibration curve, where the $\Delta R/R_0$ values are plotted vs the target gas concentration. A schematic drawing of these measurements is shown in Fig. 10.16. The figure has been drawn for the case of a reducing molecule (i.e. ammonia) and the resistance increase is determined by the injection of electrons in the p-type CNTs.

Novel perspectives in the development of chemiresistor gas sensors (CGS) can be determined by the capability to meet the requirements of environmental monitoring, in particular the sensitivity in the low ppb range.

Recently, low-cost CNT-based layers prepared on plastic substrates have been produced to monitor the presence of ammonia in the environment, where the NH₃ concentrations in the low-ppb range are expected [68]. A systematic monitoring of ammonia concentrations [69] is mandatory to reduce the hazard for human health

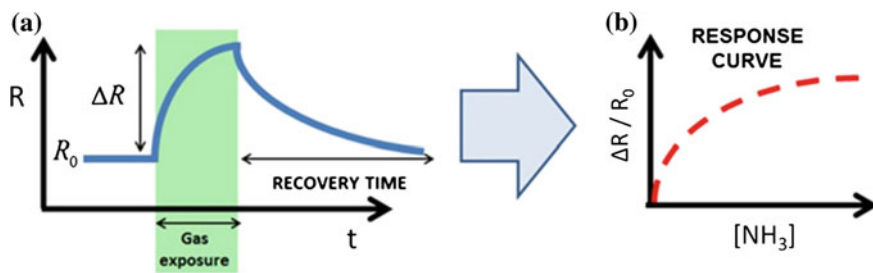


Fig. 10.16 **a** Sketch of sensor response to gas exposure and **b** response curve

and vegetation determined by the widespread use of ammonia and its derivatives as agricultural nitrogen fertilizers, which is known to determine a severe environmental problems, such as the eutrophication of semi-natural ecosystems, acidification of soils, formation of fine particulate matter in the atmosphere, and alteration of the global greenhouse balance. In spite of this urgency, the detection of ammonia atmospheric concentrations in urban areas has been so far widely overlooked, since its average levels are usually low, i.e. in the 20–30 ppb range [70], while novel ammonia CGS are not usually tested at these low levels.

Ammonia is one of the molecules most frequently used to test the capability of these devices to operate as CGS. However, little attention has been paid to the capability of these sensors to measure sub-ppm $[\text{NH}_3]$ in air. In fact, ppt sensitivity of pristine CNTs to ammonia has been demonstrated, in inert Ar atmosphere under UV irradiation [71], a detection limit (DL) of 50 ppb has been achieved for ammonia diluted in Ar using CNTs functionalized with polyaniline [72], and finally a detection limit of 30 ppb has been demonstrated for CNT layers on plastic substrates operating in ambient air [68].

Another emerging issue in the development of high sensitivity CNT-based devices is represented by hybrid systems where the CNTs are coupled with oxide nanoparticles. When decorated with oxide nanoparticles, the CNT layers can be regarded as hybrid systems, displaying properties that can be either ascribed to the superposition of the parent materials, or to novel properties emerging from the junction between the CNT wall and the nanoparticle. In the field of gas sensing, hybrid systems with CNT have been mostly created through decoration with SnO_2 , WO_3 or TiO_2 nanoparticles [73, 74]. These oxides alone, in the form of thin or thick films, are known as the most diffused materials for gas sensing, with remarkable sensitivity towards hazardous such gas as NO_2 , NO_x , etc. It is therefore a straight choice to create hybrids where CNT can provide superior charge transport and eventually gas sensing capability to be combined to that of oxide. Here, the p-n heterojunction formed between, e.g., an n-type metal oxide and a p-type carbon support is supposed to play an important role in the sensing mechanism of MOs coated CNT heterostructures [75–78].

References

1. M. Endo. Mecanisme de croissance en phase vapeur de bres de carbone (The growth mechanism of vapor-grown carbon fibers). Ph.D. thesis, University of Orleans, Orleans, France, 1975. (in French)
2. M. Endo. Ph.D. thesis, Nagoya University, Japan, 1978. (in Japanese)
3. S. Iijima, J. Cryst. Growth **55**, 675–683 (1980)
4. A. Oberlin, M. Endo, T. Koyama, Carbon **14**, 133 (1976)
5. A. Oberlin, M. Endo, T. Koyama, J. Cryst. Growth **32**, 335–349 (1976)
6. S. Iijima, Nature **354**, 56–58 (1991)
7. M.S. Dresselhaus, G. Dresselhaus, Ph Avouris, *Carbon Nanotubes. Synthesis, Structure Properties, and Applications* (Springer, 2001)
8. P. Castrucci, C. Scilletta, S. Del Gobbo, M. Scarselli, L. Camilli, M. Simeoni, B. Delley, A. Continenza, M. De Crescenzi, Nanotechnology **22**, 115701 (2011)
9. I. Stemmler, C. Backes, *Absorption Spectroscopy as a Powerful Technique for the Characterization of Single-Walled Carbon Nanotubes* (White paper, PerkinElmer, 2013)
10. M.S. Dresselhaus, G. Dresselhaus, R. Saito, A. Jorio, Raman spectroscopy of carbon nanotubes. Phys. Rep. **409**, 47–99 (2005)
11. J. Prasek, J. Drbohlavova, J. Chomoucka, J. Hubalek, O. Jasek, V. Adam, R. Kizek, Methods for carbon nanotubes synthesis—review. J. Mater. Chem. **21**, 15872 (2011)
12. N. Arora, N.N. Sharma, Arc discharge synthesis of carbon nanotubes: Comprehensive review. Diam. Relat. Mater. **50**, 135–150 (2014)
13. T. Guo, P. Nikolaev, A. Thess, D.T. Colbert, R.E. Smalley, Catalytic growth of single-walled nanotubes by laser vaporization. Chem. Phys. Lett. **243**, 49–54 (1995)
14. M. Pacheco, J. Pacheco, M. Valdivia, L. Bernal, R. Valdivia, A. Huczko, H. Lange, A. Cruz, R. Lopez-Callejas, Synthesis of Carbon Nanostructures by Using Thermal Plasma Torch. Braz. J. Phys. **34**, 4B (2004)
15. J. Kong, A.M. Cassell, Hongjie Dai, Chemical vapor deposition of methane for single-walled carbon nanotubes. Chem. Phys. Lett. **292**, 567–574 (1998)
16. M.F.L. De Volder, S.H. Tawfik, R.H. Baughman, A.J. Hart, Carbon Nanotubes: Present and Future Commercial Applications. Science **339**(6119), 535–539 (2013)
17. M. Meyyappan (ed.), *Carbon Nanotubes: Science and Applications* (CRC Press, 2005)
18. A. Javey, J. Kong (eds.), *Carbon Nanotube Electronics* (Springer Science + Business Media, 2009)
19. Dirk M. Guldi and Nazario Martín (eds.), *Carbon Nanotubes and Related Structures, Synthesis, Characterization, Functionalization, and Applications* (Wiley-VHC, 2010)
20. S. Yamashita, Y. Saito, J.H. Choi (eds.) *Carbon Nanotubes and graphene for photonics applications* (Woodhead Publishing, 2013)
21. C. Hierold (ed.), *Advanced Micro & Nanosystems*, vol. 8, Carbon Nanotube Devices: Properties, Modeling, Integration and Applications (Wiley-VHC, Weinheim, 2008)
22. Rüdiger Klingeler Robert, B. Sim (eds.), *Carbon Nanotubes for Biomedical Applications* (Springer, Berlin, Heidelberg, 2011)
23. Tony McNally, Petra Potschke (eds.), *Polymer Carbon Nanotube Composites: Preparation: Properties and Applications* (Woodhead Publishing, London, 2011)
24. Amin Salehi-Khojin, Fatemeh Khalili-Araghi, Marcelo A. Kuroda, Kevin Y. Lin, Jean-Pierre Leburton, Richard I. Masel, On the sensing mechanism in carbon nanotube chemiresistors. ACS Nano **5**(1), 153–158 (2011)
25. J. Kong, N.R. Franklin, C. Zhou, M.G. Chapline, S. Peng, K. Cho, H. Dai. Science **287**, 622 (2000)
26. D.D. Tune, B.S. Flavel, R. Krupke, J.G. Shapter, Carbon nanotube-silicon solar cells. Adv. Energy Mater. **2**, 1043–1055 (2012)
27. P. Castrucci, Carbon nanotube/silicon hybrid heterojunctions for photovoltaic devices. Adv. Nano Res. **2**, 23–56 (2014)

28. D.D. Tune, F. Hennrich, S. Dehm, M.F.G. Klein, K. Glaser, A. Colsmann, J.G. Shapter, U. Lemmer, M.M. Kappes, R. Krupke, B.S. Flavel, The role of nanotubes in carbon nanotube-silicon solar cells. *Adv. Energy Mater.* **3**, 1091–1097 (2013)
29. P.L. Mc Euen, J.Y. Park, Electron transport in single-walled carbon nanotubes. *MRS Bull.* **29**, 272–275 (2004)
30. S. Ponzoni, G. Galimberti, L. Sangaletti, P. Castrucci, S. Del Gobbo, M. Morbidoni, M. Scarselli, S. Pagliara, Selective optical switching of interface-coupled relaxation dynamics in carbon nanotube–Si heterojunctions. *J. Phys. Chem. C* **118**, 24110–24116 (2014)
31. E. Shi, L. Zhang, Z. Li, P. Li, Y. Shang, Y. Jia, J. Wei, K. Wang, H. Zhu, D. Wu, S. Zhang, A. Cao, TiO₂-coated carbon nanotube-silicon solar cells with efficiency of 15 %. *Sci. Rep.* **2** (884), 1–5 (2012)
32. Y. Jia, A. Cao, F. Kang, P. Li, X. Gui, L. Zhang, E. Shi, J. Wei, K. Wang, H. Zhu, D. Wu, Strong and reversible modulation of carbon nanotube-silicon heterojunctions solar cells by an interfacial oxide layer. *Phys. Chem. Chem. Phys.* **14**, 8391–8396 (2012)
33. D.D. Tune, A.J. Blanch, R. Krupke, B.S. Flavel, J.G. Shapter, Nanotube film metallicity and its effect on the performance of carbon nanotube-silicon solar cells. *Phys. Status Solidi A* **211**, 1479–1487 (2014)
34. Y. Jung, X. Li, N.K. Rajan, A.D. Taylor, M.A. Reed, Record high efficiency single-walled carbon nanotube/silicon p–n junction solar cells. *Nano Lett.* **13**, 96–99 (2013)
35. P. Wadhwa, B. Liu, M.A. McCarthy, Z. Wu, A.G. Rinzler, Electronic junction control in a nanotube-semiconducting schottky junction solar cell. *Nano Lett.* **10**, 5001–5005 (2010)
36. H.C. Card, Photovoltaic properties of MIS-schottky barriers. *Solid State Electron.* **20**, 971–976 (1977)
37. X. Li, J.S. Huang, S. Nejati, L. McMillon, S. Huang, C.O. Osuji, N. Hazari, A.D. Taylor, Role of HF in oxygen removal from carbon nanotubes: implications for high performance carbon electronics. *Nano Lett.* **14**, 6179–6184 (2014)
38. C. Pintossi, G. Salvinelli, G. Drera, S. Pagliara, L. Sangaletti, S. Del Gobbo, M. Morbidoni, M. Scarselli, M. De Crescenzi, P. Castrucci, Direct evidence of chemically inhomogeneous, nanostructured, Si–O buried interfaces and their effect on the efficiency of carbon nanotube/Si photovoltaic heterojunctions. *J. Phys. Chem. C* **117**, 18688–18696 (2013)
39. C. Pintossi, S. Pagliara, G. Drera, F. De Nicola, P. Castrucci, M. De Crescenzi, M. Crivellari, M. Boscardin, L. Sangaletti, Steering the efficiency of carbon nanotube-silicon photovoltaic cells by acid vapor exposure: a real-time spectroscopic tracking. *ACS Appl. Mater. Interfaces* **7**(18), 9436–9444 (2015)
40. K.A. Mirica, J.G. Weis, J.M. Schnorr, B. Esser, T.M. Swager, Mechanical drawing of gas sensors on paper. *Angew. Chem. Int. Ed.* **51**, 10740–10745 (2012)
41. A. Goldoni, L. Petaccia, L. Gregoratti, B. Kaulich, A. Barinov, S. Lizzit, A. Laurita, L. Sangaletti, R. Larciprete, *Carbon* **42**, 2099 (2004)
42. S. Chopra, K. McGuire, N. Gothard, A.M. Rao, A. Pham, *Appl. Phys. Lett.* **83**, 2280 (2003)
43. T. Someya, J. Small, P. Kim, C. Nuckolls, J.T. Yardley, *Nano Lett.* **3**, 877 (2003)
44. J. Li, Y. Lu, Q. Ye, M. Cinke, J. Han, M. Meyyappan, *Nano Lett.* **3**, 929 (2003)
45. J.P. Novak, E.S. Snow, E.J. Houser, D. Park, J.L. Stepnowski, R.A. McGill, *Appl. Phys. Lett.* **83**, 4026 (2003)
46. F. Picaud, C. Girardet, A.M. Rao, *J. Appl. Phys.* **105**, 014315 (2009)
47. P. Vichchulada, P.Q. Zhang, M.D. Lay, *Analyst* **132**, 719 (2007)
48. D.R. Kauffman, A. Star, *Angew. Chem.* **47**, 6550 (2008)
49. M. Penza, R. Rossi, M. Alvisi, M.A. Signore, E. Serra, *J. Phys. D Appl. Phys.* **42**, 072002 (2009)
50. T. Zhang, S. Mubeen, N.V. Myung, M.A. Deshusses, *Nanotechnology* **19**, 332001 (2008)
51. A. Goldoni, L. Petaccia, S. Lizzit, R. Larciprete, *J. Phys.: Condens. Matter* **22**, 013001 (2010)
52. M. Penza, R. Rossi, M. Alvisi, E. Serra, *Nanotechnology* **21**, 105501 (2010)
53. M. Penza, P.J. Martin, J.T.W. Yeow, Carbon Nanotube Gas sensors, in *Gas Sensing Fundamentals. Springer Series on Chemical Sensors and Biosensors*, ed. by C.-D. Kohl, T. Wagner (Springer, Berlin, Heidelberg, 2014)

54. P.G. Collins, K. Bradley, M. Ishigami, A. Zettl, *Science* **287**, 1801 (2000)
55. J. Kong, N.R. Franklin, C. Zhou, M.G. Chapline, S. Peng, K. Cho, H. Dai, *Science* **287**, 622 (2000)
56. G.U. Sumanasekera, C.K.W. Adu, S. Fang, P.C. Eklund, *Phys. Rev. Lett.* **85**, 1096 (2000)
57. X.P. Tang, A. Kleinhammes, H. Shimoda, L. Fleming, K.Y. Bennoun, S. Sinha, C. Bower, O. Zhou, Y. Wu, *Science* **288**, 492 (2000)
58. J. Zhao, A. Buldum, J. Han, J.P. Lu, Gas molecule adsorption in carbon nanotubes and nanotube bundles. *Nanotechnology* **13**, 195–200 (2002)
59. H. Chang, J.D. Lee, S.M. Lee, Y.H. Lee, *Appl. Phys. Lett.* **79**, 3863 (2001)
60. B.B. Shirvani, J. Beheshtian, G. Parsafar, N.L. Hadipour, *Comput. Mater. Sci.* **48**, 655 (2010)
61. L. Valentini, C. Cantalini, I. Armentano, J.M. Kenny, L. Lozzi, S. Santucci, **21**, 1996 (2003)
62. C.Y. Lee, S. Baik, J.Q. Zhang, R.I. Masel, M.S. Strano, *J. Phys. Chem. B* **110**, 11055 (2006)
63. J. Suehiro, G.B. Zhou, M. Hara, *J. Phys. D Appl. Phys.* **36**, 109 (2003)
64. E.S. Snow, F.K. Perkins, E.J. Houser, S.C. Badescu, T.L. Reinecke, *Science* **307**, 1942 (2005)
65. S. Chopra, A. Pham, J. Gaillard, A. Parker, A.M. Rao, *Appl. Phys. Lett.* **80**, 4632 (2002)
66. A. Modi, N. Koratkar, E. Lass, B.Q. Wei, P.M. Ajayan, *Nature* **424**, 171 (2003)
67. C. Brett, Other types of sensors: Impedance-based sensors, fet sensors, acoustic sensors, in *Environmental Analysis by Electrochemical Sensors and Biosensors, Nanostructure Science and Technology*, eds. by L.M. Moretto, K. Kalcher (Springer, 2014), pp. 351–370
68. F. Rigoni, S. Tognolini, P. Borghetti, G. Drera, S. Pagliara, A. Goldoni, L. Sangaletti, *Analyst* **138**, 7392–7399 (2013)
69. B. Timmer, W. Olthuis, A. van den Berg, *Sens. Actuators B* **107**, 666 (2005)
70. M. Chiesa, F. Rigoni, M. Paderno, P. Borghetti, G. Gagliotti, M. Bertoni, A. Ballarin, Denti, L. Schiavina, A. Goldoni and L. Sangaletti. *J. Environ. Monit.* **14**, 1565 (2012)
71. G. Chen, T.M. Paronyan, E.M. Pigas, A.R. Harutyunyan, *Sci. Rep.* **2**, 343 (2012)
72. T. Zhang, M.B. Nix, B.-Y. Yoo, M.A. Deshusses, N.V. Myung, *Electroanalysis* **18**, 1153 (2006)
73. T. Kerdcharoen, C. Wongchoosuk, Carbon nanotube and metal oxide hybrid materials for gas sensing. in *Semiconductor Gas Sensors, Woodhead Publishing Series in Electronic and Optical Materials*, eds. by R. Jaaniso, O.K. Tan, pp. 386–407 (2013)
74. D. Eder, Carbon nanotube-inorganic hybrids. *Chem. Rev.* **110**, 1348–1385 (2010)
75. C. Marichy, P.A. Russo, M. Latino, J.P. Tessonier, M.G. Willinger, N. Donato, G. Neri, N. Pinna, Tin dioxide–carbon heterostructures applied to gas sensing: structure-dependent properties and general sensing mechanism. *J. Phys. Chem. C* **117**, 19729–19739 (2013)
76. C. Marichy, N. Donato, M.-G. Willinger, M. Latino, D. Karpinsky, S.-H. Yu, G. Neri, N. Pinna, Tin dioxide sensing layer grown on tubular nanostructures by a non-aqueous atomic layer deposition process. *Adv. Funct. Mater.* **21**, 658–666 (2011)
77. E. Llobet, R. Ionescu, E.H. Espinoza, R. Leghrib, A. Felten, R. Erni, *Sens. Actuators B* **131**, 174 (2008)
78. B.-Y. Wei, M.-C. Hsu, P.-G. Su, H.-M. Lin, R.-J. Wu, H.-J. Lai, *Sens. Actuators B Chem.* **101**, 81–89 (2004)

Chapter 11

Effects of Charging and Perpendicular Electric Field on Graphene Oxide

H. Hakan Gürel, M. Topsakal and S. Ciraci

Abstract We present a first-principles study of the effects of charging, electric field on the oxidation/deoxidation of graphene oxide consisting of epoxy and hydroxyl groups. We first determined the proper basis set which prevents the spurious spilling of electrons of graphene oxide, when negatively charged or exerted by perpendicular electric field and treated with periodic boundary conditions. Applied perpendicular electric field to graphene surface is provided side specific functionalization. We showed that the bonds between oxygen and graphene are weakened under applied electric field. For specific values of excess charge or perpendicular electric field, oxygen atom moves to the top site from bridge site which is normally absorbed in equilibrium. Individually adsorbed oxygen atoms cannot form oxygen molecules due to an energy barrier. This energy barrier is dramatically weakened, when negatively charged or exerted by an electric field. Beside the epoxy groups, hydroxyl groups have an important role of oxidation/deoxidation of graphene oxide. Charging and perpendicular electric field mediates the reduction of graphene oxide through the formation of H_2O and H_2O_2 . Our results explain the role of external effects to the reduction of graphene oxide.

H. Hakan Gürel
Technology Faculty, Information Systems Engineering Department,
Kocaeli University, 41380 Kocaeli, Turkey
e-mail: hhakan.gurel@kocaeli.edu.tr

M. Topsakal
UNAM-National Nanotechnology Research Center, Bilkent University,
06800 Ankara, Turkey
e-mail: metosa@bilkent.edu.tr

S. Ciraci (✉)
Department of Physics, Bilkent University, 06800 Ankara, Turkey
e-mail: ciraci@fen.bilkent.edu.tr

11.1 Introduction

Graphene oxide (GOX) has been an important material because it allows the production of large scale graphene sheets through the reduction of oxidized multilayer graphene [1, 2]. Despite the great amount of experimental and theoretical research carried out recently, [1–16] a thorough understanding of the interaction of epoxy and hydroxyl groups with graphene resulting in oxidation/deoxidation process are not fully understood yet. Active research to date has concluded that the oxidation and reduction of graphene are, in fact, rather complex and comprise the interplay of various molecules and atoms, such as O, O₂, CO₂, OH, H, H₂O, as well as external agents. GOX has been also a subject of interest because the electronic properties of graphene, in particular its linear π and π^* bands, [17] which cross at the Fermi level undergo dramatic changes upon oxidation [18–21]. Introducing a band gap, which varies with oxygen coverage and hence changes semimetallic graphene into a semiconductor, has been an active field of study in graphene based nanoelectronics. Furthermore, it has been reported that the oxidized surface of graphene can be changed reversibly between dark and light spots with the applied lateral electric field [21]. To tune the band gap, the extent of reduction can also be controlled by applied perpendicular electric field, whereby deoxidized spots are attained underneath an STM tip under a specific bias [21]. These results show that the properties of graphene can be modified by the controlled and reversible reduction/oxidation of GOX, which are realized by charging it or by applying perpendicular electric field.

11.2 Methodolgy

We carried out spin-polarized and spin-unpolarized calculations within density functional theory (DFT) using projector-augmented wave potentials [22]. The exchange-correlation potential is approximated by generalized gradient approximation using Perdew, Burke and Ernzerhof (PBE) functional [23]. Initially, we used two different basis sets, namely plane waves (PW) and local basis (AO) sets for reasons explained below, which are applied by using first-principles simulation packages VASP [24] and SIESTA, [25] respectively. Based on the results derived from the test calculations, all of our conclusions concerning the effect of the charging and electric field are obtained from AO calculations using SIESTA [25].

A plane-wave basis set with kinetic energy cutoff of 500 eV is used in PW calculations [24]. Atomic positions and lattice constants are optimized by using the conjugate gradient method, where the total energy and atomic forces are minimized. The convergence for energy is chosen as 10^{-5} eV between two steps. Oxygen-atom adsorbed to one side of graphene breaks inversion symmetry and a net electric-dipole moment is generated perpendicular to the graphene surface. Dipole corrections [26] are applied in order to remove spurious dipole interactions between periodic images for the neutral calculations.

In AO calculations [25], we use double ζ polarized basis set and the eigenstates of the Kohn-Sham Hamiltonian are expressed as linear combinations of numerical atomic orbitals. A 200 Ryd mesh cut-off is chosen and the self-consistent calculations are performed with a mixing rate of 0.1. The Brillouin zone is sampled with a Monkhorst-Pack mesh [27] with $(5 \times 5 \times 1)$ $\mathbf{k}$ -points, whereas we use $(45 \times 45 \times 1)$ $\mathbf{k}$ -points in specific systems. Core electrons are replaced by norm-conserving, nonlocal Trouiller-Martins pseudopotentials [28]. The vacuum spacing between graphene layers in adjacent supercells is taken 15 Å. The convergence criterion for the density matrix is taken as 10^{-4} .

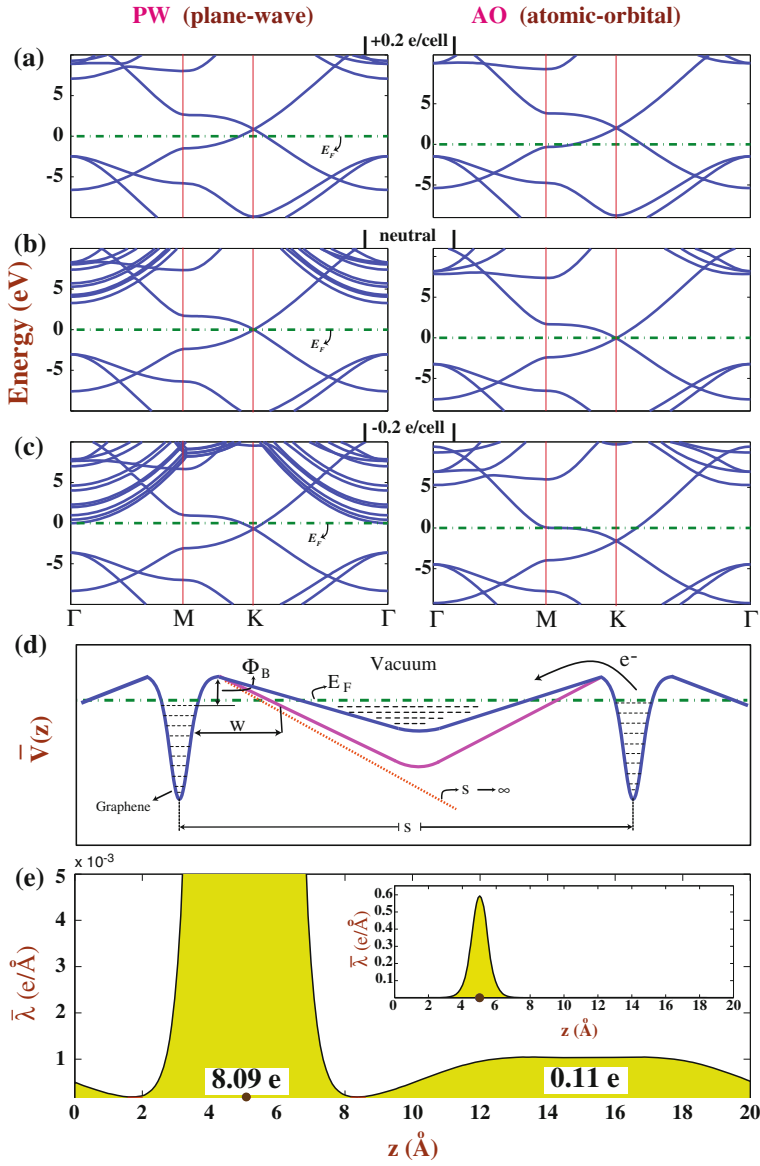
11.3 Theoretical Investigations of Charged Nanosystems

The theoretical investigations of charged nanosystems have been an active field of study. Developing appropriate formalisms to provide reliable predictions on the effects of charging and applied electric field have been the subject matter of several studies [26, 29–43].

The aim of this study is to investigate the effects of electric field and charging on OH, H and O which are adsorbed to graphene. Here, we performed the calculations of charged GOX by adding desired amount of excess electrons for the case of negative charging or by removing electrons for the case of positive charging, where both cases are treated using periodic boundary conditions (PBC). Throughout the study, $Q > 0$ (or surface charge density $\bar{\sigma} = Q/A$ in Coulomb/m², A being the area of the cell) indicates the positive charging, namely number of depleted electrons per cell; $Q < 0$ indicates the negative charging, namely the number of excess electrons per cell, and $Q = 0$ is the neutral cell. The bare (super)cell is made up of $(n \times n)$ primitive unit cells of graphene; each supercell comprises $2n^2$ carbon atoms and $8n^2$ valence electrons. We assume that graphene planes, which are repeating periodically along the z -axis are parallel to the (x, y) -plane. The electric field $\mathbf{E}$, which is applied perpendicularly to the graphene plane, is specified as positive i.e. $E_{field} > 0$, if it is along z -direction (or it is pointing towards adsorbates, i.e. H, OH, O). This electric field induces electronic charge transfer from the adsorbate to graphene and the case is vice versa if the direction of $\mathbf{E}$ is reversed, i.e. $E_{field} < 0$. Electric field induced charge transfer modifies the charge distribution and hence affects the physical and chemical properties.

We first examine the limitations of PBC method, where two-dimensional graphene layers separated by large spacing s are repeated periodically along z -direction. We carried out first-principles PW calculations as well as AO calculations using linear combinations of numerical atomic orbitals and investigated the effect of charging on the graphene layer.

In Fig. 11.1a–c we present electronic band structures of graphene calculated by PW and AO for $Q = +0.2$ ($\bar{\sigma} = +0.61$ C/m²), 0, -0.2 electrons/cell ($\bar{\sigma} = -0.61$ C/m²) using the primitive cell ($n = 1$). Further details about the changes in the band



◀ **Fig. 11.1 a–c** Energy band structures of graphene calculated for $Q = +0.2$ e/per primitive unit cell ($\bar{\sigma} = +0.61$ C/m²), 0 and -0.2 e/per primitive unit cell ($\bar{\sigma} = -0.61$ C/m²). *Left (right)* panels are calculated using PW basis set using VASP code (AO basis set using SIESTA code). **d** Schematic description of the plane-averaged potential of negatively charged graphene, $\bar{V}(z)$ dipping at the center of vacuum spacing and forming a quantum well like structure. The vacuum spacing between two adjacent graphenes is denoted by s . The *triangular* potential barrier and its width are Φ_B and w , respectively. The quantum well confining electrons, which are spilled from charged graphene of PW calculations is illustrated by *blue lines*; the well without bound electrons shown by *red lines* corresponds to AO calculations. For one single graphene layer corresponding to the “actual case”, $\bar{V}(z \rightarrow \pm\infty) \rightarrow -\infty$ for $s \rightarrow \infty$. **e** Plane-averaged (*linear*) charge density, $\bar{\lambda}(z)$ obtained from PW calculations for $Q = -0.2$ e/cell ($\bar{\sigma} = -0.61$ C/m²) and $s = 20$ Å. $\bar{\lambda}(z)$ in full y -axis range is shown by *inset*. Here “cell” indicates the primitive unit cell, and $s = 20$ Å. This figure is taken from [29]

structure of graphene upon charging have been discussed in our recent studies [42, 43]. In the case of $Q = +0.2$ ($\bar{\sigma} = +0.61$ C/m²), the Fermi level shifts down or the Dirac point raises for both PW and AO calculations. Accordingly, semimetallic graphene turns into a hole-doped metallic system.

For neutral case ($Q = 0$), band structures calculated using PW and AO are similar. The work function is the difference of the reference vacuum energy and the Fermi level. The value of work function extracted from Fig. 11.1b is 4.26 eV and this is in good agreement with earlier study [44]. Briefly, the results calculated using PW and AO are similar. Minute differences originate from different pseudopotentials used in calculations. Moreover, the perfect convergences of different basis sets can be achieved only by using very large cutoff values.

As for the case of negative charging with $Q = -0.2$ e/cell ($\bar{\sigma} = -0.61$ C/m²), the Fermi level shifts up significantly in AO results. In the case of PW results the Fermi level shifts up only slightly and then is pinned by the free electron-like parabolic bands touching to the Fermi level; thereafter the position of the Fermi level is practically unaltered with increasing negative charging as shown in Fig. 11.1c. The dramatic difference for the case of $Q < 0$ between PW and AO results can be understood through the plane-averaged electronic potential along the z -direction, i.e. $\bar{V}(z)$ as described schematically in Fig. 11.1d. In the actual case, $\bar{V}(z)$ passes through a maximum near the surface and goes to $-\infty$ as $z \rightarrow \pm\infty$ at both sides of one single graphene layer, which is charged with $Q < 0$. Then the electrons in the graphene can spill to vacuum only if they tunnel through tunneling across a wide triangular barrier. The width of the barrier, w decreases with increasing negative charging. Accordingly, the tunneling current would be negligibly low and excess electrons would practically be trapped in graphene if the value of $|Q|$ is not very high. However, when treated within PBC, $\bar{V}(z)$ of periodically repeating graphene layers with a vacuum spacing s between them can make a dip reminiscent of a quantum well at the center of the spacing as shown in Fig. 11.1d. Under these circumstances, Kohn-Sham Hamiltonian using PW can acquire solutions in this quantum well, which are localized along the z -direction, but free-electron like in the (x, y) -plane parallel to graphene. These states form the parabolic bands (k_x, k_y) plane

as shown in Fig. 11.1c. If the quantum well dips below the Fermi level with increasing s or Q , excess electrons in graphene start to be accommodated in these 2D free-electron like bands. This ends up with the spilling of electrons into the vacuum region.

In Fig. 11.1e we present the plane-averaged charge density, $\bar{\lambda}(z)$ calculated within PW. By integrating $\bar{\lambda}(z)$ along the z -direction, it is concluded that graphene can keep only -0.09 e/cell ($\bar{\sigma} = -0.28$ C/m²) out of -0.2 e/cell ($\bar{\sigma} = -0.61$ C/m²) excess electrons. The remaining -0.11 e/cell ($\bar{\sigma} = -0.33$ C/m²) are spilled out into the vacuum, namely to the region between two adjacent graphene planes. Clearly, the spilling of excess electron, which can be normally trapped in graphene, is an artifact of PBC. The amount of excess electrons spilled to the vacuum spacing increases with increasing s and increasing negative charging. While wide s is not convenient because of excessive charge spilling, small s may prevent excess electrons from spilling into the vacuum at very low negative charging and can yield reasonable solutions as long as the coupling of the adjacent layers is negligible. However, the charge spilling can set in when the negative charging exceeds a threshold value corresponding to the given s . The dependency of electronic structure and hence the band gap opening on s in PW calculations of silicene (graphene like Si) under perpendicular electric field has been also reported recently [45]. The perpendicular electric field breaks the mirror symmetry of $\bar{V}(z)$ and gives rise to a sawtooth like variation. At high $\mathbf{E}$ and large s , electrons of silicene spilling to low potential site result in a situation discussed above and in [30] above for $Q < 0$.

The following conclusions have been arrived regarding the features and limitations of first-principles methods in treating these external effects using plane wave (PW) and local basis (AO) sets: (i) When negatively charged, the electronic potential between periodically repeating graphene layers makes a dip at the center of the vacuum spacing s . This dip forms a quantum well like structure and its depth from the Fermi level increases with s and with excess charge $Q < 0$. (ii) PW calculations result in charge spilling to the vacuum region for large s , which is the consequence of the artifact of PBC. (iii) On the other hand, such a spilling does not occur in AO calculations even if PBC is used, since local basis orbitals at carbon sites fail to represent states in the quantum well like potential at the middle of the vacuum spacing. (iv) This is the artifact of AO calculations, which turned out to be its advantage, whereby the artifact of PBC is tacitly canceled out. (v) The electronic potential of one single graphene layer trapping the excess charge in the actual case is close to that obtained by AO calculations using PBC. Hence the results concerning charging obtained from AO calculations are expected to be very close to the results of the actual case and are physical. (vi) This analysis can also be extended to graphene exerted by an perpendicular electric field $\mathbf{E}$. While the excess electrons can spill to the vacuum region at the lower energy side of the saw tooth like electronic potential treated by PW and by using PBC, AO calculations provide predictions close to the actual case of single graphene under the perpendicular $\mathbf{E}$.

According to the above analysis concerning the artifact of PW basis set using PBC, our results in this study are obtained by performing first-principles spin polarized calculations within Density Functional Theory (DFT) [46] using linear combination of numerical atomic orbitals (LCNAO). Therefore, present results are expected to be very close to the results of the actual case corresponding to single graphene layer under $\mathbf{E}$ or containing excess $Q < 0$.

11.4 Interaction of H_2O , OH, O and H with Graphene

The interactions of H_2O , OH, O and H with graphene are essential for the reduction of GOX through desorption of H_2O , O_2 and H_2O_2 . Here we characterize their interactions in equilibrium conditions by calculating the optimized binding geometry, the corresponding binding energy, the minimum energy barrier in their diffusion and the path of diffusion with least energy barrier.

11.4.1 Binding of H_2O to Graphene

H_2O , an indispensable ingredient in the oxidation/reduction process of GOX, has chemical as well as van der Waals (vdW) interaction with graphene [47]. Since the generalized gradient approximation (GGA) does not include vdW interaction, it underestimates the binding energy between H_2O and graphene. The binding energy calculated by using GGA lies in the range of 18–47 meV depending on its orientation and position [48]. On the other hand, the local density approximation (LDA) overestimates the binding energy to be 151 meV. Therefore, the binding energy of H_2O including vdW interaction is estimated to be between 18 and 151 meV. It is really a weak interaction and a water molecule physisorbed to graphene can be desorbed near the room temperature.

11.4.2 Binding of OH to Graphene

It is demonstrated experimentally that OH is one of the critical functional groups existing in GOX [49, 50]. The binding energies and magnetic moments of OH adsorbed to graphene are calculated by AO for hollow (H), top (T) and bridge (B) sites on graphene layer. As shown in Fig. 11.2a, the top (T) site is most favorable energetically with a binding energy $E_b = 0.97$ eV and the magnetic moment $\mu = 0.51 \mu_B$. Carbon atoms underlying OH are displaced slightly upwards from the plane of graphene. Earlier calculations found the binding energy of OH in the energy range between 0.54 and 0.86 eV [51–56]. The energy landscaping shown in Fig. 11.2b is calculated by relaxing all the position of carbon atoms, as well as the height z of

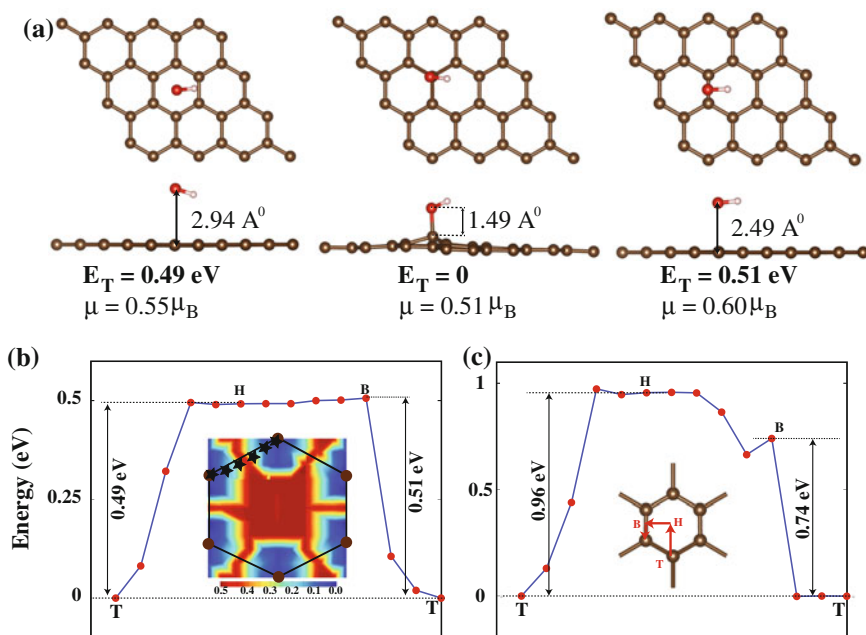


Fig. 11.2 Interaction and binding between OH/H and graphene surface. **a** Top and side views of atomic configurations, total energy E_T and magnetic moments μ in Bohr magneton μ_B , of OH adsorbed to hollow (H), top (T) and bridge (B) sites on graphene. The zero of energy is set to E_T at the top site. Large brown, large red and small yellow balls represent carbon, oxygen and hydrogen atoms, respectively. **b** The energy landscaping of OH adsorbed to different sites in the honeycomb structure and the variation of energy of OH migrating along symmetry sites on a hexagon, i.e. $T \rightarrow H \rightarrow B \rightarrow T$. The minimum energy barrier between T and H-sites is $E_B = 0.49$ eV. A possible path of minimum energy barrier for the diffusion of adsorbed OH is shown by stars. **c** The variation of energy of H adatom migrating along the symmetry sites on a hexagon, i.e. $T \rightarrow H \rightarrow B \rightarrow T$. The minimum energy barrier occurs between T and B-sites is $E_B = 0.74$ eV. Accordingly, H adatom migrates above the C–C bonds. Equilibrium binding energies of OH and H occur at the top site as 0.97 and 0.76 eV, respectively. Calculations are performed using a system, where a single OH or H is adsorbed to each (4×4) supercell of graphene. This figure is taken from [29]

single adsorbed OH, while its x and y coordinates are fixed in the (4×4) supercell of graphene. These calculations are repeated for 36×36 (x, y)-grid points on a hexagon. One can deduce the minimum energy barrier to the diffusion to be $E_B = 0.49$ eV from the variation of the calculated total energies along the symmetry directions, $T \rightarrow H \rightarrow B \rightarrow T$ of the hexagon presented in Fig. 11.2b. This comparatively low energy barrier allows easy migration of OH on graphene at elevated temperatures. We expect that E_B calculated at low coverage is modified at very high coverage due to increased hydroxyl-hydroxyl interaction [15, 57].

11.4.3 Binding of O to Graphene

When neutral, single oxygen is adsorbed at the bridge site of graphene, namely above the center of the C–C bonds of honeycomb structure. Calculations yield that the bridge (B-)site is energetically the most favorable adsorption site; the top (T-) sites (on top of carbon atoms) are not favorable. PW calculations [15] predicted the binding energies of a single oxygen at the B-site as 2.35, 2.40, 2.43 and 2.43 eV using (2×2) , (3×3) , (4×4) and (5×5) supercells of graphene, respectively. The binding energy is practically unaltered for supercells larger than (5×5) [15]. Using AO we found that the binding energy of oxygen on (5×5) graphene supercell is 2.34 eV.

The formation energy related with the oxidation of graphene is negative and hence graphene cannot be oxidized through oxygen molecules [15, 58]. Actually, O_2 is physisorbed to bare graphene surface with a very weak binding energy of 58 meV calculated without van der Waals correction; but it raises to 115 meV when the van der Waals correction [59] is included. Only at defect sites like holes and vacancies, O_2 can dissociate [60] and its constituent oxygen atoms become adsorbed to carbon atoms having a lower coordination numbers [58]. In contrast to O_2 , the bonding of free oxygen to graphene is rather strong and changes between 2.43 and 3.20 eV depending on the coverage [15]. The crucial question to be addressed is how oxygen atoms can remain strongly bound to graphene despite their negative formation energy and why bound oxygen atoms are prevented from desorption through the formation of oxygen molecule. Recent studies have provided the energy barrier for the diffusion of adsorbed O on graphene and for the desorption of O through the formation of O_2 and CO_2 [15, 16].

11.4.4 Binding of H and H_2 to Graphene

The binding of atomic hydrogen H on graphene has been treated in earlier studies [61–63]. Binding energies reported by different authors vary in a wide range of energy [64]. They lie between 0.47 and 1.44 eV with a majority of data being between 0.6 and 0.85 eV. This variability can be ascribed to differences in structures, optimization procedure, and computational methodology used to calculate the chemisorption of H on graphite [64]. Here, for the sake of completeness, we calculate the binding energy using the same calculation parameters and local basis set used throughout the present work. In agreement with previous studies, we found that the strongest binding of H atom occurs at the top site with a binding energy of $E_b = 0.76$ eV. The variation of the total energies of H adatom moving along specific directions of the honeycomb structure are also calculated and the minimum energy barrier to the diffusion is found to be $E_B = 0.74$ eV as shown in Fig. 11.2c.

Similar to water molecule, the binding of hydrogen molecule H_2 , to graphene is weak. Calculations with the LDA VWN [65] functional result in a binding energy

of 93 meV for molecular hydrogen. Due to the lack of vdW interaction, GGA calculations using PW91 and PBE functional predict relatively lower binding energies of 23 and 13 meV, respectively [66].

11.5 Effects of an Electric Field and Charging

11.5.1 *Effects of an Electric Field and Charging on Adsorbed O*

The effects of applied electric field and charging are interrelated. Upon the adsorption of oxygen atom, 0.78 electrons are transferred from carbon atoms to oxygen in neutral case. The electric field applied perpendicularly to the graphene $\mathbf{E}$, which is specified as positive if it is along z -direction (or it is pointing towards oxygen adatom). This electric field induces electronic charge transfer from the adsorbed oxygen to graphene or vice versa if its direction is reversed. Electric field induced charge transfer modifies the charge distribution and hence affects the physical and chemical properties. Results obtained from this section will enlighten recent experimental studies performed for similar systems [20, 21].

Interesting effects of electric field on oxygen adsorbed graphene are summarized in Fig. 11.3. The binding energy of the oxygen adatom increases with applied negative electric field, which is perpendicular to the surface and pointing the direction opposite to oxygen. The height h of oxygen adatom from the graphene plane also increases, even if one expects the opposite trend. This paradoxical situation originates from the definition of the binding energy, E_b , of oxygen adsorbed to graphene, which is given by $E_b = E_T[\text{O}] + E_T[\text{graphene}, \mathbf{E}] - E_T[\text{O} + \text{graphene}, \mathbf{E}]$, in terms of the total energies of single neutral oxygen atom $E_T[\text{O}]$, of bare graphene under $\mathbf{E}$, $E_T[\text{graphene}, \mathbf{E}]$ and single oxygen adsorbed to graphene under applied $\mathbf{E}$, $E_T[\text{O} + \text{graphene}, \mathbf{E}]$ all calculated using (5×5) supercell. According to this definition, increasing E_b does not mean that the bond between oxygen and graphene has become stronger. In fact, that h increases with increasing negative $\mathbf{E}$ implies the opposite situation.

Before we explain this paradoxical situation, we first consider charge rearrangements caused by $\mathbf{E}$. Induced charge transfer depends on the direction of $\mathbf{E}$. In Fig. 11.4 we present the effects of applied perpendicular electric field on the charge distribution and potential energy $\bar{V}(z)$. An electric field $\mathbf{E} = -1.0 \text{ V/\AA}$ transfers electrons from the bottom side to the upper side of the graphene and also causes to further transfer of carbon electrons to oxygen adatom as shown in Fig. 11.4b. At the end, the total energy of whole system increases. This effect creates an interesting situation also for the bare graphene by breaking the projection symmetry between its two sides. The direction of electron transfer is reversed when the direction of the applied electric field is reversed. Accordingly, one can monitor various properties of bare graphene, in particular its chemical activity by applying perpendicular electric

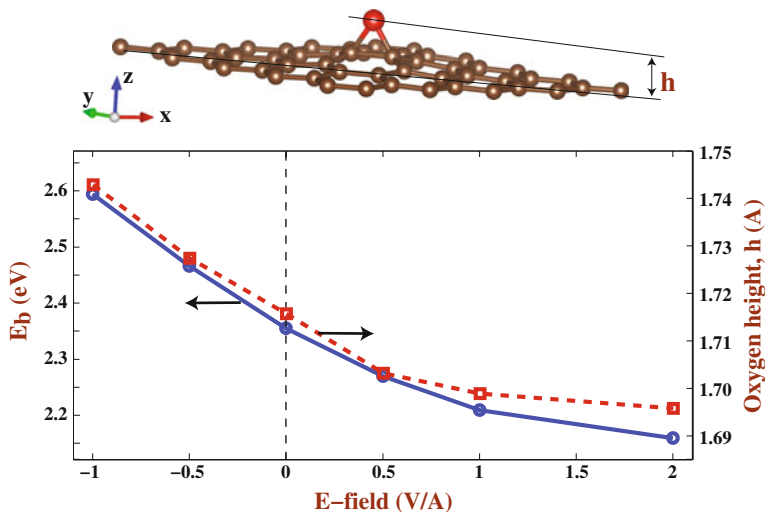


Fig. 11.3 Variation of the binding energy E_b , of oxygen adatom, and height h , from the graphene plane with the electric field $\mathbf{E}$, applied perpendicular to the graphene plane. $\mathbf{E}$ is taken positive, if it is along z -direction (or it is pointing towards oxygen adatom) and vice versa. (Results are obtained from AO calculations using PBC.) This figure is taken from [29]

field and hence by modifying electron concentration at both sides as seen in Fig. 11.4c.

The self-consistent potential energy averaged on the planes perpendicular to z -axis, $\bar{V}(z)$ is presented in Fig. 11.4d. Within PBC $\bar{V}(z)$ displays a saw teeth behavior. The form of the potential suggests interesting situations for both sides of graphene. For example, one can monitor the work function on both sides. Side specific effects induced by perpendicular electric field are pronounced in single layer honeycomb structures consisting of two or three atomic planes, such as silicene [45, 67, 68], single layer transition metal dichalcogenides [69] and graphene bilayer.

The effect of applied perpendicular electric field on the calculated total energies of oxygen adsorbed at different sites on the NEB path between two adjacent B-sites through the T-site is presented in Fig. 11.5. The difference of total energy ΔE , between oxygen adsorbed at the T-site and B-site is as large as ~ 1.2 eV for $\mathbf{E} = +2.0$ V/Å. However, when the direction of $\mathbf{E}$ is reversed, ΔE decreases and becomes even negative for $\mathbf{E} = -2.0$ V/Å. This is a dramatic effect and is expected to bear on the reduction/oxidation of graphene surfaces. Since the energy barrier between the B-site and T-site can be modified by the applied electric field, the monitoring of oxygen diffusion at the graphene surface will be possible through applied electric field. We note that the magnitudes of the electric fields of 1–2 V/Å are high but they are in the range, which can be applied at least for short times [70].

We now resolve the above paradoxical situation, namely that the equilibrium oxygen graphene distance h increases with E_b increases as shown in Fig. 11.3 and

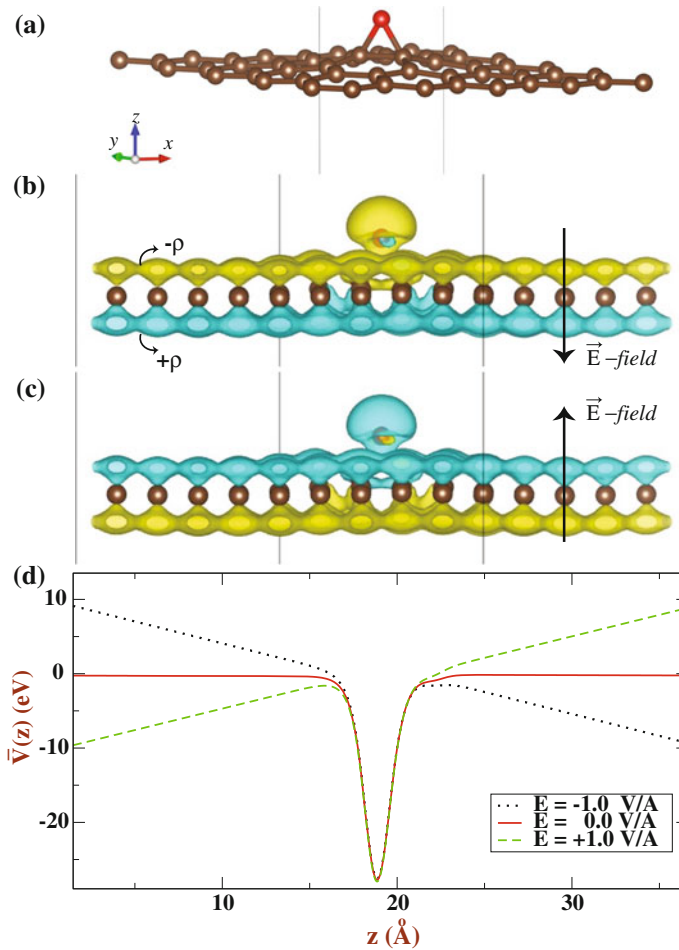


Fig. 11.4 **a** Atomic configuration showing a single oxygen atom adsorbed at the B-site of a (5×5) graphene supercell with *red* and *brown* balls indicating oxygen adatom and carbon atoms of graphene, respectively. The (xy) -plane coincides with the atomic plane of bare graphene. z -direction is perpendicular to graphene. **b** and **c** The difference charge density $\Delta\rho$ for the electric field antiparallele ($\vec{E} < 0$) and parallel ($\vec{E} > 0$) to z -direction, respectively. Yellow (blue) isosurfaces indicate electron accumulation (depletion) induced by $\vec{E}$ applied in different directions. **d** Self-consistent field potential energy averaged on the planes perpendicular to z -axis throughout the supercell, $\bar{V}(z)$, for the cases $\vec{E} = 0$, $\vec{E} < 0$, and $\vec{E} > 0$. (Results are obtained from AO calculations using PBC.) This figure is taken from [29]

develop a criterion for the strength of bond between graphene and oxygen adatom. To this end, we consider the system consisting of a (5×5) graphene supercells with a single oxygen adatom, which are either charged by Q or under a perpendicular electric field, $\vec{E}$ as shown in Fig. 11.6a. We calculated the optimized total energies of this system while oxygen adatom is pulled along z -direction perpendicular to the

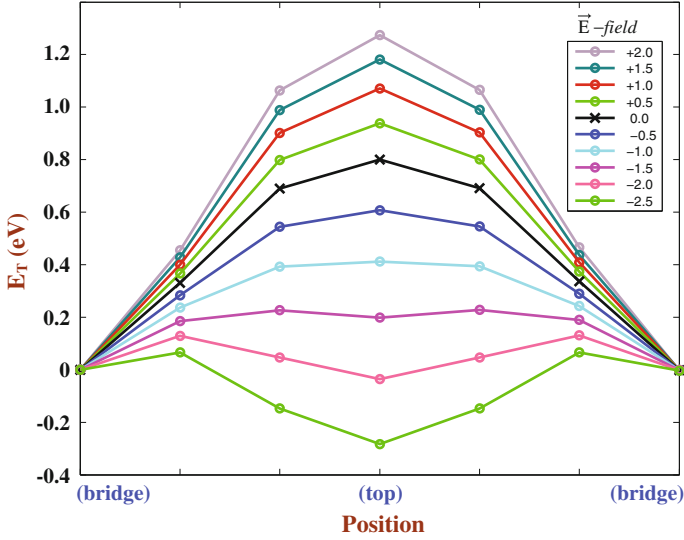
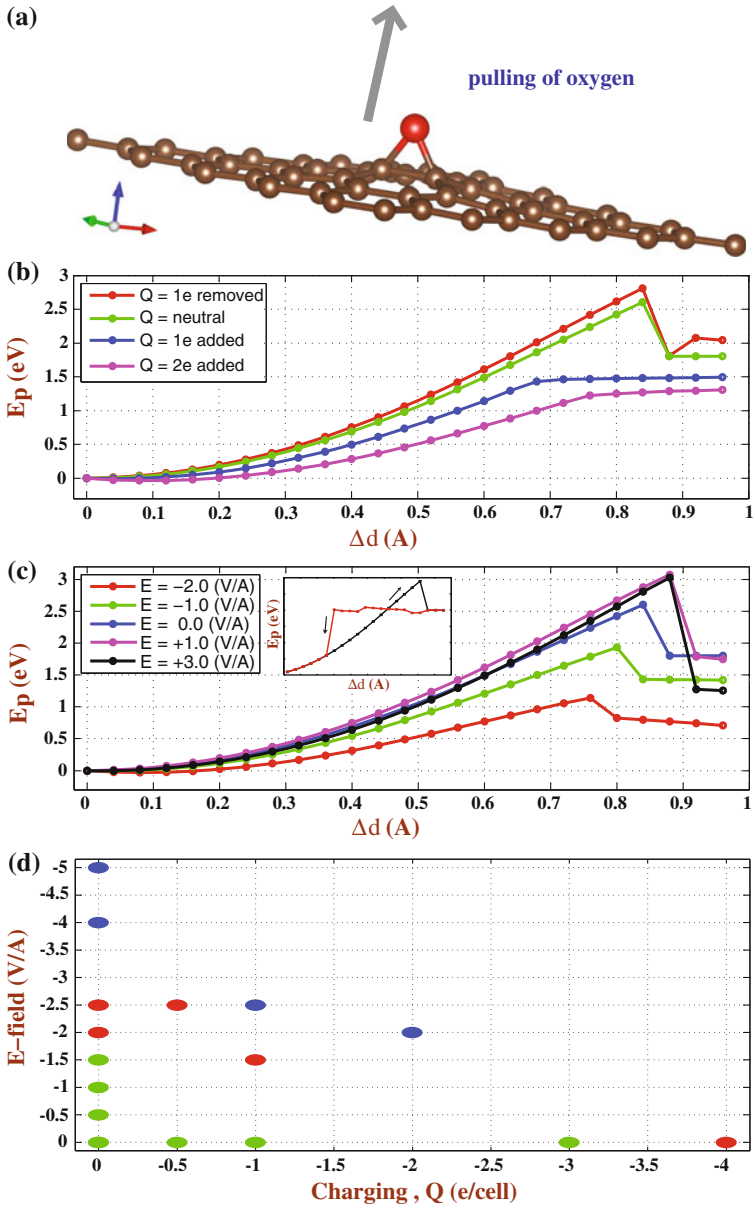


Fig. 11.5 Variation of the total energy E_T , of single oxygen atom adsorbed to each (5×5) graphene supercell under applied perpendicular electric fields $\mathbf{E}$ on a NEB path between two adjacent B-sites through a T-site. The total energy at the B-site is set to zero for all cases. The unit of $\mathbf{E}$ is V/Å. (Results are obtained from AO calculations using PBC.) This figure is taken from [29]

plane of graphene. For each fixed value Δd of oxygen adatom from its equilibrium height h , carbon atoms around oxygen adatom are relaxed, while the rest of the atoms are fixed to prevent graphene from displacement. These analyses are continued by varying Δd for different values of Q [e per (5×5) cell] and for different values of $\mathbf{E}$ perpendicular to the graphene plane.

The variation of the total energies with pulling Δd are plotted in Fig. 11.6b, c for different values of Q and $\mathbf{E}$, respectively. For the sake of comparison, we also included the pulling curves for $Q = 0$ and $\mathbf{E} = 0$. These figures convey interesting features regarding the effects of either charging or applied perpendicular electric field on the strength of the bond between graphene and oxygen adatom: Normally, the energy associated with pulling $E_p = E_T[\text{O} + \text{graphene}, Q, \mathbf{E}, \Delta d] - E_T[\text{O} + \text{graphene}, Q, \mathbf{E}, \Delta d = 0]$ increases with increasing Δd , since the system is strained and pulled upwards. Eventually it passes through a maximum value denoted by E_p^* and drops suddenly at about $0.8 < \Delta d < 0.9$ Å. Our analysis suggests that E_p^* can be taken as a measure for the strength of the bond between oxygen adatom and graphene. We note that E_p^* , namely the energy barrier to pull out the adsorbed oxygen adatom from graphene surface is ~ 2.60 eV for both $Q = 0$ and $\mathbf{E} = 0$. This energy consistent with the binding energy calculated for the bridge site. An interesting feature of the present analysis is that E_p^* is strongly dependent on charging and electric field. While E_p^* increase with $Q > 0$, it decreases dramatically for $Q < 0$. For example, $E_p^* = 1.46$ eV for $Q = -1.0$ e/cell (or -0.02 e per carbon atom or $\bar{\sigma} =$



◀ **Fig. 11.6** Variation of the pulling energy E_p , of a single oxygen adatom adsorbed to each (5×5) supercell of graphene charged by Q or exerted by $\mathbf{E}$ with pulling distance along the z -axis, Δd from its equilibrium height h . **a** Atomic configuration. *Red (large) and brown (small) balls* indicate oxygen and carbon atoms, respectively. **b** Variation of E_p with Δd , along z -axis for different charging values Q . **c** Variation of E_p with Δd , along z -axis for different applied electric field $\mathbf{E}$. All cases have their own equilibrium heights h . The maximum value of the pulling energy for each case is specified by E_p^* . As shown by *inset* for the case of $Q = 0$ and $\mathbf{E} = 0$, a bistability occurs when oxygen atom approaches to graphene from a long distance. **d** Positions of adsorbed oxygen atom under different values of electric field $\mathbf{E}$ and/or charging Q . *Green, red and blue dots* indicate respectively, oxygen atom adsorbed to the bridge-site, top-site and desorbed by moving away from graphene plane. (Results are obtained from AO calculations using PBC.) This figure is taken from [29]

-0.12 C/m^2). Similarly, E_p^* can be as low as $\sim 1.14 \text{ eV}$ under the perpendicular electric field $\mathbf{E} = -2 \text{ V/\AA}$. As shown by inset in Fig. 11.6c, a bistability occurs if oxygen atom moves in the reverse direction and hence approaches to the graphene from $\Delta d > 1.0 \text{ \AA}$. Notably, for $Q < 0$ E_p curve doesn't experience a sharp fall passing its maximum value. This is related with the excess charge on oxygen atom. These are important results and demonstrate that it is easier to desorb oxygen adatoms from graphene by negatively charging or by applying perpendicular electric field. Under high local charging and local electric field, which can be attained by the sharp tip of a Scanning Tunneling Microscope or by a gate voltage the reduction of GOX through the desorption of O adatoms can be achieved.

Since the negatively charging or negative perpendicular electric field both weaken the bond between oxygen and graphene and hence lower E_p^* , we next explore their effects on the bond between graphene and oxygen adatom when they both coexist. Thus the negative $\mathbf{E}$ of different magnitudes exerts on a negatively charged system consisting of single oxygen adatom adsorbed to each (5×5) supercell of graphene. We found that the effects same as in Fig. 11.6b, c are attained by applying relatively lower electric field when the system is negatively charged. In particular, adsorbed oxygen moves high above graphene and becomes weakly bound under the electric field $\mathbf{E} = -2.5 \text{ V/\AA}$, if the whole system is charged by $Q = -1.0 \text{ e/cell}$ (or $\bar{\sigma} = -0.12 \text{ C/m}^2$). We note that this value is much smaller than the electric field required for desorption of oxygen from a neutral system. In Fig. 11.6d we schematically show different positions of oxygen adatom (the bridge-, top-site and moved away from graphene for desorption) occurring under different values of negative $\mathbf{E}$ and negative Q .

11.5.2 Effects of an Electric Field and Charging on Adsorbed OH

The effects of perpendicular electric field $\mathbf{E}$ and charging Q on OH adsorbed to graphene occur through the modification of the equilibrium charge distribution. In Fig. 11.7 we present atomic geometry of adsorption, the isosurfaces of difference

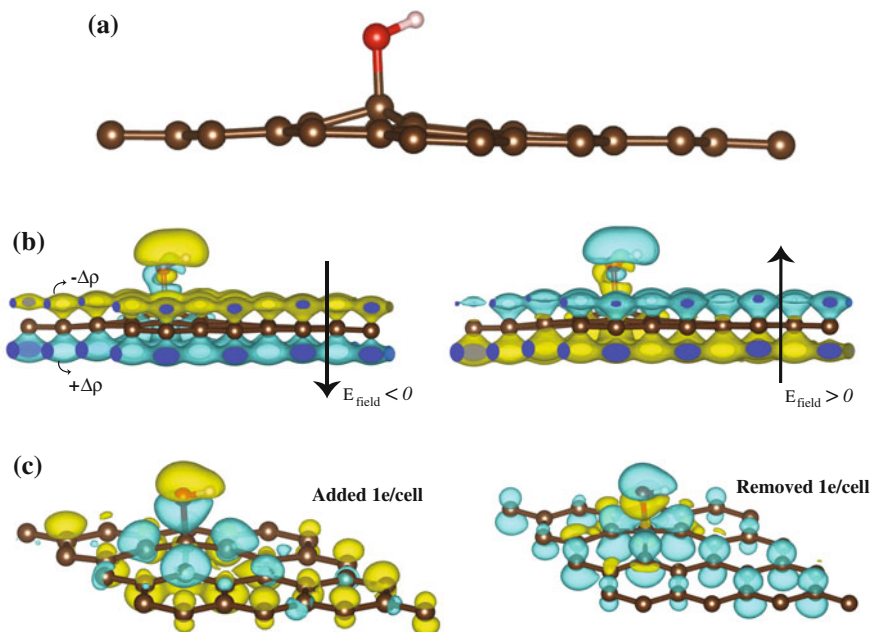


Fig. 11.7 **a** Equilibrium atomic configuration of one OH adsorbed to the top site of the (4×4) supercell of graphene. **b** Isosurfaces of the difference charge density $\Delta\rho$ under perpendicular electric field $E_{field} = \pm 1.0$ V/Å. Negative and positive $\Delta\rho$ are shown by yellow and turquoise isosurfaces, respectively. **c** Isosurfaces of the difference charge density for $Q = \pm 1.0$ e/supercell. This figure is taken from [31]

charge density, $\Delta\rho = \rho[\mathbf{E} \text{ or } Q] - \rho[\mathbf{E} = 0; Q = 0]$. For $E_{field} < 0$, electronic charge is transferred from the bottom site of graphene towards OH and upper site leading to accumulation of more charge on OH. The charge transfer is reversed when the direction of the perpendicular electric field is reversed i.e. $E_{field} > 0$. The negative charging $Q < 0$ realized by adding electrons in the system gives rise to electron accumulation at OH. This situation is, however, reversed for positive charging, $Q > 0$ realized by removing electrons from the OH+graphene system. Clearly, the interaction of OH with graphene, hence the strength of the bond is modified depending on the magnitude and direction of $\mathbf{E}$, as well as the sign and magnitude of charging.

Here we further investigate the strength of the bond by calculating the optimized total energy of OH+graphene system as OH is pulled in a perpendicular direction (z -direction) gradually for different values of Q [e per (4×4) cell] and for different values of E_{field} . For each fixed value Δz of OH from its equilibrium height, all carbon atoms in the supercell are relaxed, while carbon atoms at the corners of supercell are fixed to prevent graphene from displacement. This analysis is continued by varying Δz . The variation of the total energies with pulling Δz are plotted in Fig. 11.8b, c for different values of Q and E_{field} , respectively. For the sake of comparison, we also included the pulling curves for $Q = 0$ and $\mathbf{E} = 0$. These figures

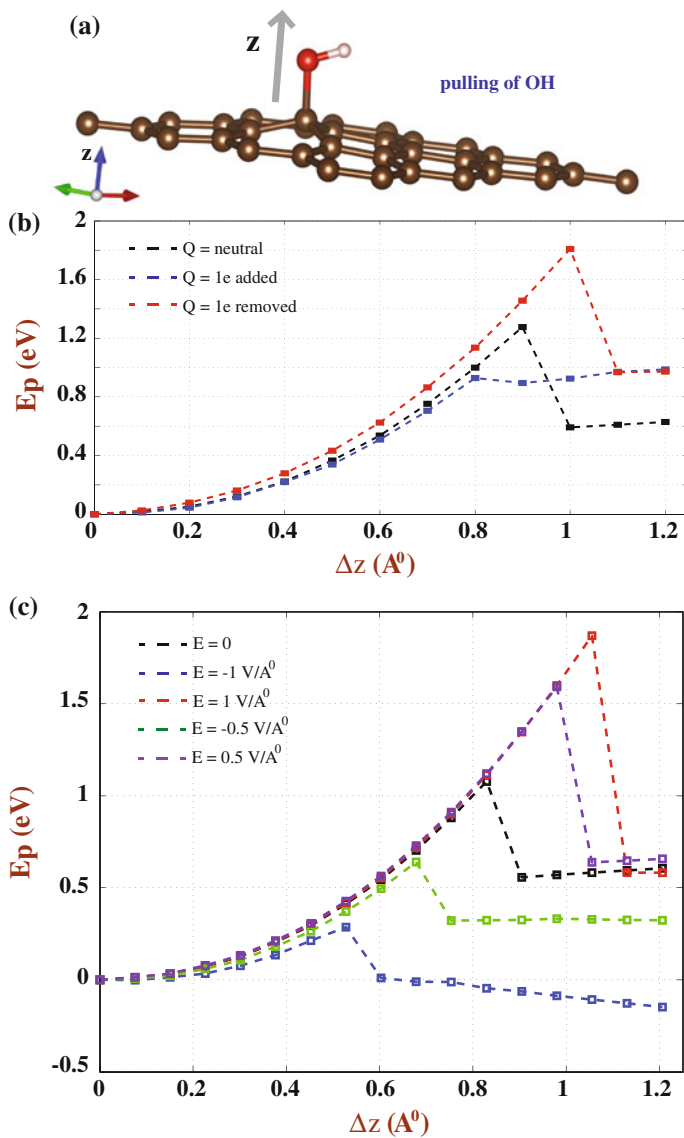


Fig. 11.8 **a** Atomic configuration consisting of one OH adsorbed to each (4×4) supercell of graphene, where OH is pulled in the perpendicular direction by the distance Δz from its equilibrium height on graphene surface. *Large brown, large red and small yellow balls stand for C, O and H atoms, respectively.* **b** The total energy E_p , versus the distance Δz for different values of Q . **c** E_p , versus the pulling distance Δz for different values of E_{field} . The zero of energy is set to the total energies corresponding to $\Delta z = 0$. This figure is taken from [31]

clarify the effects of either Q or $\mathbf{E}$ on the strength of the OH-graphene bond as discussed in the previous sections: Normally, the energy associated with pulling $E_p = E_T[\text{OH} + \text{graphene}; Q, \mathbf{E}; \Delta z] - E_T[\text{OH} + \text{graphene}; Q, \mathbf{E}; \Delta z = 0]$ increases with increasing Δz , since the system is strained and pulled upwards. Eventually it passes through a maximum value denoted by E_p^* and drops suddenly at about $0.5 < \Delta z < 1.0$ Å. Our analysis suggests that E_p^* can be taken as a measure for the strength of the bond between OH and graphene. We note that E_p^* , namely the energy barrier to pull out the adsorbed OH from graphene surface is ~ 1.3 eV for both $Q = 0$ and $\mathbf{E} = 0$. This energy is 0.33 eV larger than the equilibrium binding energy E_b of OH, since it corresponds to the strained configuration of underlying graphene, where carbon atoms at the corners of the supercell are fixed. An interesting feature of the present analysis is that E_p^* is strongly dependent on charging and electric field. While E_p^* increases with $Q > 0$, it decreases dramatically for $Q < 0$. For example, $E_p^* \sim 1.8$ eV for $Q = +1.0$ e/cell, but it decreases to $E_p^* \sim 0.9$ eV for $Q = -1.0$ e/cell. Notably, for $Q < 0$ E_p^* curve doesn't experience a sharp fall passing its maximum value. This is related with the excess charge on OH. Similarly, E_p^* can be as low as ~ 0.35 eV under the perpendicular electric field $E_{field} = -1.0$ V/Å. We note that a bistability [29] may occur if OH moves in the reverse direction and hence approaches to the graphene from $\Delta z > 1.2$ Å. These results also confirm that it is easier to desorb OH and to achieve the reduction of GOX by negatively charging or by applying $E_{field} < 0$. Under high local charging and local electric field, which can be attained by the sharp tip of a Scanning Tunneling Microscope or by a gate voltage the reduction of GOX can be achieved easily.

11.6 Desorption of Oxygen from GOX

In the above sections we discussed interaction of single O, H, H₂, H₂O and OH with graphene surface and also revealed how the binding and related properties of O and OH are affected with applied $\mathbf{E}$ and Q . The interaction of oxygen atom with graphene surface was investigated thoroughly in earlier studies [15, 29, 58]. In this section we investigate the binary interactions among H, O and OH. Our objective is to reveal whether oxygen atoms can desorb from GOX via hydroxyl groups and how the reduction process is affected by Q and $\mathbf{E}$.

11.6.1 Formation of Oxygen Molecule

Having examined the binding energy, binding site and diffusion of single oxygen atom and their variation with charging and applied electric field, we next consider the formation of O₂, which can be essential in deoxidation. Since single oxygen

adatom adsorbed to graphene has negative formation energy, one cannot expect that O_2 molecule dissociates into two atomic oxygens, both adsorbed to graphene. However, O_2 molecules can dissociate at the edges of vacancy defects or holes and subsequently the constituent atomic oxygens are adsorbed to carbon atoms with lower coordination [58]. In view of the negative formation energy, one normally expects that two oxygen adatom adsorbed individually to graphene can readily form O_2 in an exothermic process when two atoms are at close proximity. Therefore, the interaction of two oxygen adatom on graphene is essential for the formation of O_2 . In Fig. 11.9a we examined O–O interaction for $Q < 0$ and $\mathbf{E} = 0$ while one O is at the position identified as “ p ” and marked by the arrow as shown by inset, the other one is diffusing from a distant bridge-site identified as “ a ”. Initially, the interaction between them is weak, but develops as one O atom is migrating from a towards p through b , c and d positions. Here a and c correspond to the bridge- and b and d to the top-sites. Here we consider three situation, which are illustrated by three curves denoted by I, II and III. For the curve I, the (x, y) -coordinates of all carbon atoms of graphene are kept fixed, but their z -coordinates are relaxed as the first O adatom migrates from a to the second O at p . While the first O is forced to migrate through the path of the minimum energy barrier by optimizing its height from graphene, the second O is fully relaxed. Overcoming the barrier of ~ 3.3 eV corresponding to configuration K , O_2 is desorbed at G . Curve I is similar to the curve in Fig. 11.5c of [15] calculated using PW method. For the curve II, only one carbon atom out of 72 in the supercell is fixed to prevent underlying graphene from displacing in the course of forced migration of the first O adatom. A relatively smaller barrier of ~ 2.3 eV develops between d and H , and eventually one O adatom is desorbed once H is overcome at L . Apparently curve I and curve II are associated with high energy barrier for the desorption of O_2 or single oxygen atom from GOX. A different migration path is followed for the curve III: Once the first O adatom arrived at d , it is fixed there while the second O adatom at p is forced to the top site at e . To attain the final configuration J , an energy barrier of 1.3 eV from c has to be overcome. After J , two O adatoms form O_2 and desorb from graphene. Clearly, this path described by curve III has much lower barrier than those F and H . The energy barriers of curve I–III are lowered with negative charging.

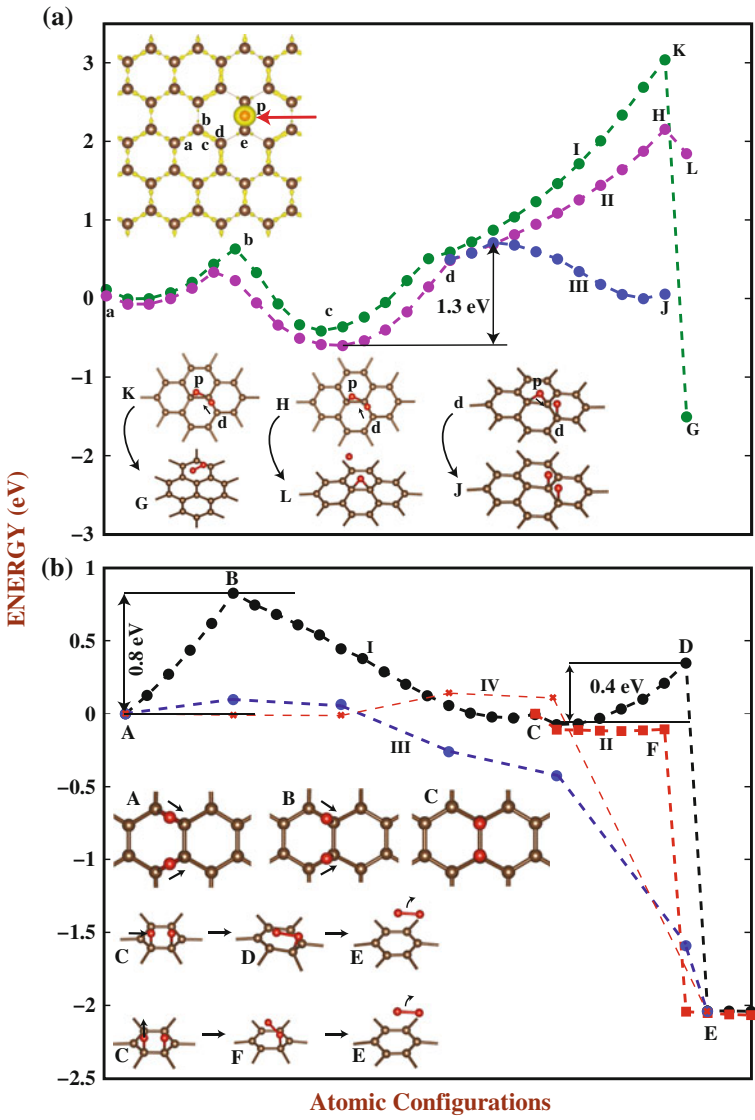
Larciprete et al. [16] pointed out a dual path described by inset A in Fig. 11.9b, which allows two O adatoms at the bridge sites to move along C–C bonds towards to adjacent top sites. They calculated the energy barrier to be 1.13 eV using PW method and found that it is in good agreement with temperature programmed desorption (TPD) data. They observed significant O_2 desorption at ~ 500 K during the thermal annealing of low oxygen density GOX. Here, we first examine the low-energy-barrier dual path proposed by Larciprete et al. [16] and calculate the energy barrier for $Q < 0$ and $\mathbf{E} = 0$ using AO method by relaxing all atoms in the (5×5) supercell except one carbon atom to hinder the displacement of graphene layer. In Fig. 11.9b we present the energy variation of the dual path corresponding to curve I through A, B, C and D configurations shown by inset. The energy barrier at B configuration is rather low as compared to F and H configurations in Fig. 11.9a. Overcoming an energy barrier of 0.8 eV,

two O adatoms become attached to two adjacent top sites of graphene described as the configuration C. The energy barrier calculated by AO is ~ 0.3 eV lower than predicted by Larciprete et al. [16] perhaps due to the different parameters of calculations and different basis set. If one of these O atoms is forced towards the other one an energy barrier of 0.4 eV develops between C and D. The formation of O_2 and desorption is achieved at E when D is overcome. A different path specified as curve II in Fig. 11.9b goes to F from C, where one of O adatoms of the configuration C is raised. This way, the formation of O_2 attained without an energy barrier after the configuration C. The atomic configurations corresponding to various stages are described by insets.

Having discussed two paths with low energy barrier leading to the formation of O_2 from GOX, we next examine how these barriers are affected by external agents such as negative charging and applied electric field. When charged by $Q = -4$ e per supercell ($\bar{\sigma} = -0.332$ C/m²), the energy variation of curve I and curve II is modified as presented by curve III in Fig. 11.9b. Two oxygen adatoms migrating from the configuration A towards the configuration C encounter a small energy barrier of ~ 0.1 eV. This is really small barrier, which can be further lowered or completely suppressed with increasing excess electrons. The application of the electric field of $E = -2$ V/Å perpendicular to the plane of graphene induces a similar effect shown by curve IV: While the barrier of 0.8 eV at B is suppressed, a small barrier of ~ 0.2 eV near C appears. The energy barrier in curve III and curve IV are rather small and explain why GOX can easily be deoxidized under external effects [20, 21].

11.6.2 Interaction Between Adsorbed H and O

Formation of OH from adsorbed O and H is another process contributing to deoxidation of GOX. Here we examine the interaction between coadsorbed H and O atoms to see how OH can form. We consider three different paths for H atom approaching the adsorbed O atom. (i) Both H and O are initially coadsorbed; O is adsorbed at the bridge site and initially H is adsorbed at the top site at the close proximity of O adatom as shown in Fig. 11.10a. Here we move adsorbed H atom on its migration path with minimum energy barrier on graphene by fixing its x - and y -coordinates, but fully relaxing its z -coordinate, as well as all the coordinates of adsorbed O atom and of all carbon atoms of graphene. As the coadsorbed H approaches the adsorbed O, the energy falls suddenly by ~ 1.4 eV when OH forms. This exothermic process occurs without any barrier. At this moment, owing to the constraints in the approach of H adatom, O adatom desorbs to form strong O–H bond. Eventually, OH becomes detached from graphene. Apparently, the bond energy of OH compensates the binding energies of O and H atoms with graphene, as well as the energy lowering of ~ 1.4 eV of whole system. However, as shown in Fig. 11.10a, an energy barrier of 1.3 eV can develop as the adsorbed H approaches the adsorbed O, if the graphene is fixed from corner atoms, rather than it is fully relaxed. (ii) Along the second path shown in



◀ **Fig. 11.9** Interaction between two oxygen adatoms (epoxy groups) at close proximity. **a** Variation of the interaction energy between two oxygen adatom are presented for three different case represented by curves I–III associated with single path. Calculations are carried out for a rectangular supercell consisting of 72 carbon atoms and two oxygen adatoms. Details are given in the text. The *top inset* shows the isosurface charge densities of C–C bonds and positions of two oxygen adatoms in the course of migration. *Bottom insets* describes how O₂ and O are desorbed from various configurations. **b** Variation of the interaction energy between two oxygen adatoms involving dual paths are presented by curve I and II. Relevant atomic configurations A–F associated with dual paths and ending with O₂ desorption are shown by *insets*. The modification of curve I and II under excess electronic charge ($Q = -4$ e per supercell or $\bar{\sigma} = -0.332$ C/m²) and applied electric field ($E = -2$ V/Å) (Results are obtained from AO calculations using PBC.) This figure is taken from [29]

Fig. 11.10b, where O is adsorbed at the bridge site, free H is approaching it from the top vertically. At a specific distance overcoming an energy barrier of 0.1 eV, OH is formed in an exothermic process by lowering the energy of the system by 2.6 eV. Because of the strategy of approach, O adatom, which is detached from graphene forms weakly bound OH. The gain of energy through the formation of free OH compensate the desorption of O adatom. Upon the adsorption of OH the total energy can be further lowered to ~ 3.3 eV. (iii) Along the third path shown in Fig. 11.10c, whereby O is adsorbed, free H is approaching it horizontally to form OH. The formation of OH occurs without any barrier, and the total energy is first lowered 2.6 eV once weakly bound OH is formed. Thereafter, the energy is further lowered to ~ 3.3 eV upon the adsorption of OH. It appears that the formation of OH through the interaction between O adatom and H atom, which is either free in the environment or coadsorbed to graphene can occur easily by the release of significant energy. Formation of OH adsorbed on graphene is a crucial step towards the formation of H₂O from hydroxyl groups.

11.6.3 Interaction Between Adsorbed H and OH

Here we consider also three different cases to investigate the interaction between H and adsorbed OH. (i) As described in Fig. 11.11a, initially one H and OH are coadsorbed at close proximity and occupy the top sites of two outer carbon atoms of two adjacent C–C bonds. When a weakly bound H₂O is formed, the energy gained from this process compensates the sum of the binding energies of H and OH and further lowers the total energy by 2.65 eV. The cases shown in Fig. 11.11b, c are similar to the above analysis and involve the interaction between free H atom and adsorbed OH. In the case described in Fig. 11.11b, the barrier is only 30 meV between adsorbed OH and free H atom. Once this small barrier is overcome, a weakly bound H₂O forms. In this exothermic process, an energy of 4.5 eV is released. This is a significant energy which can trigger other reactions. The reaction described in Fig. 11.11c takes place in two stages: First, free H atom is bonded to graphene temporarily, eventually an energy of 4.5 eV is released, when a weakly bound H₂O is formed.

Fig. 11.10 Variation of the total energy E_T as H atom is approaching the adsorbed oxygen from different directions by a displacement Δs . **a** H atom on graphene is approaching oxygen atom adsorbed at the bridge site. Energy variation presented by black (red) dashed lines corresponds to graphene substrate, which is fully relaxed (fixed at the corner atoms). Δs follows the energy path with minimum barrier described in the text. **b** Free H atom is approaching the adsorbed O atom vertically from the top and is forming weakly bound OH. Here $\Delta s = -\Delta z$. **c** H atom is approaching the adsorbed O atom horizontally. Weakly bound OH corresponds to an intermediate state, which is exited by ~ 0.9 eV relative to adsorbed OH. Calculations are performed using (4×4) supercell of graphene. Large-brown, large-red and small-yellow atoms are carbon, oxygen and hydrogen atoms, respectively. This figure is taken from [31]

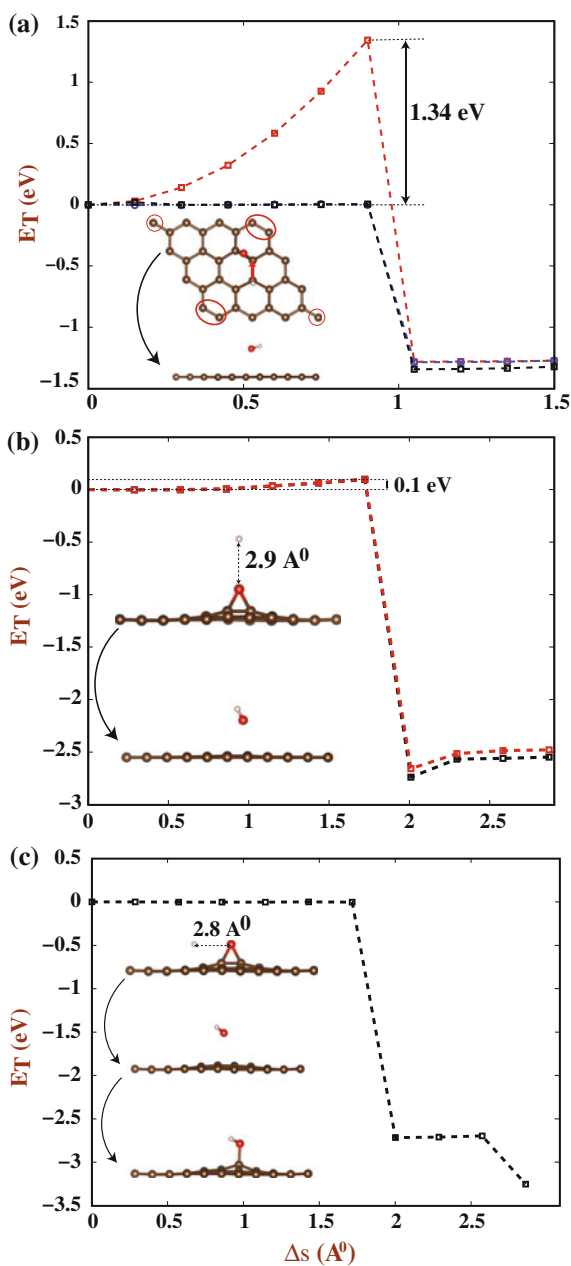
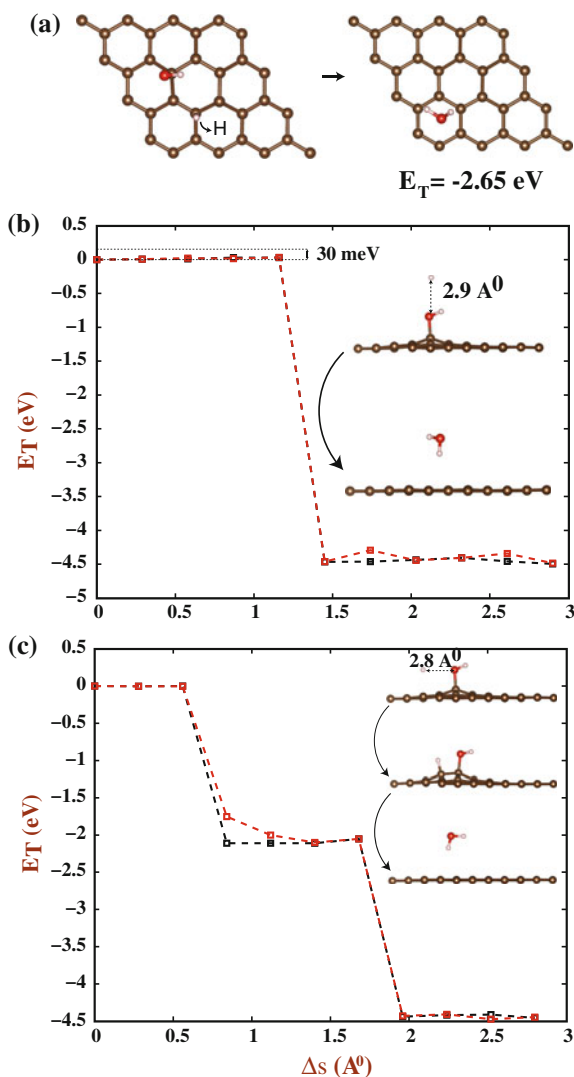


Fig. 11.11 a Initially, H and OH are coadsorbed and are at the close proximity occupying outer top sites of two adjacent C–C bonds. When H₂O is formed the total energy is lowered by -2.65 eV.

b Variation of the total energy E_T as a free H atom approaches to OH from the top by $\Delta s = -\Delta z$. Formation of H₂O occurs upon overcoming an energy barrier of 30 meV. **c** As a free H atom approaching horizontally first it is adsorbed to graphene and then form H₂O. Whole process is exothermic and release ~ 4.5 eV. H₂O by itself is weakly bound to graphene and can desorb easily. Calculations are performed using (4×4) supercell of graphene. This figure is taken from [31]



11.6.4 Interaction Between Two OH Co-Adsorbed in Close Proximity

Finally, we examine the interaction and chemical processes, when two OH coadsorbed at the close proximity approach each other. The interaction between two coadsorbed OH's is relevant for the deoxidation of GOX for the reasons pointed out at the beginning. Here, we move one OH along the path of minimum energy barrier towards the other OH. While moving one OH, its x - and y -coordinates are fixed

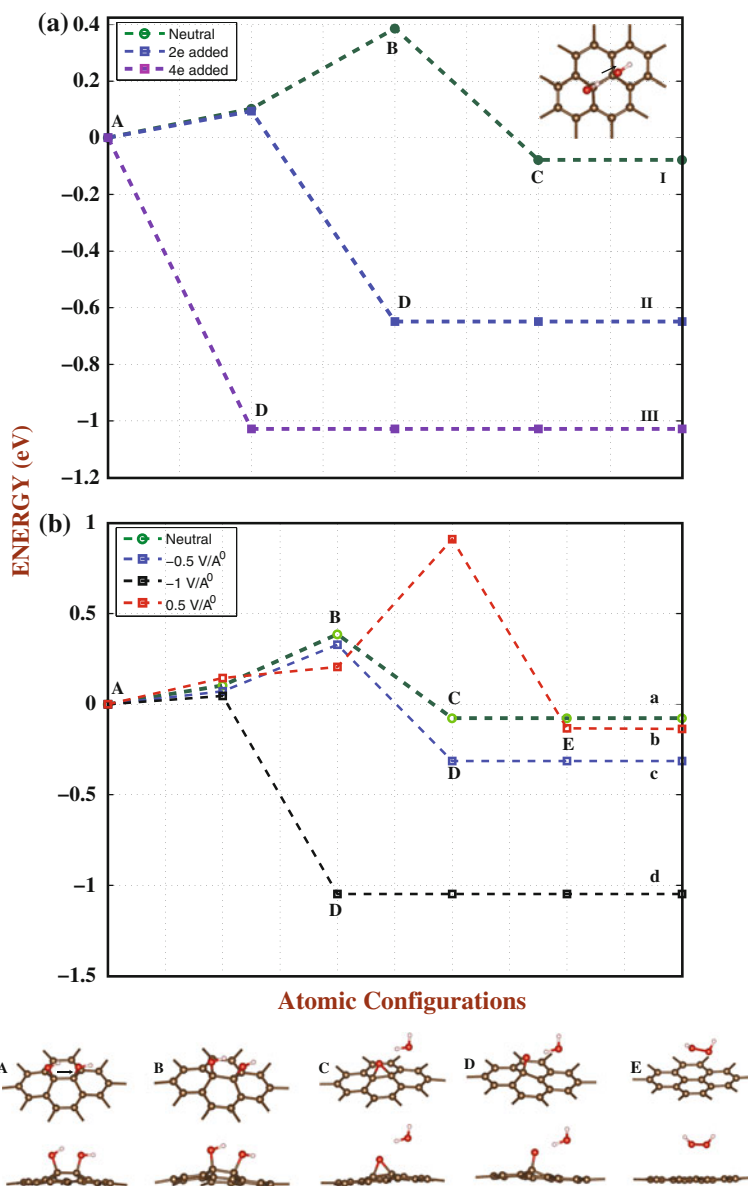


Fig. 11.12 Variation of the total energy E_T as two OH atoms adsorbed at close proximity are approaching each other. **a** The system is charged by $Q = -2$ and $Q = -4$ electrons per supercell. **b** The system is under an electric field $E_{field} = +0.5, -0.5$ and -1.0 V/Å. The neutral system (i.e. $Q = 0$ and $E_{field} = 0$) is also shown by green dashed lines for the sake of comparison. Various atomic configurations starting from A going through B and ending in C or D or E are described at the bottom of the figure. Large brown, large red and small yellow balls represent C, O and H atoms, respectively. The zero of energy is set to the total energies of the initial configuration-A i.e. two adsorbed OH's are widely separated. Calculations are performed using (6×6) supercell of graphene. This figure is taken from [31]

along the migration path, but its z -coordinate, all the coordinates of the second OH, as well as all the coordinates of graphene atoms are fully relaxed. In this case, an energy barrier of ~ 0.4 eV prevents these two OH from engaging in a chemical reaction. Note that due to OH–OH interaction this barrier is smaller than E_B in Fig. 11.2. When the energy barrier is overcome, the chemical reaction sets in to form one weakly bound H_2O molecule and one O atom bound to the bridge site. The former is prone to desorb easily and hence to remove one adsorbed O atom from GOX. The relevant processes, $A \rightarrow B \rightarrow C$, and corresponding energy variation are shown in Fig. 11.12.

Since a significant barrier is involved in OH–OH interaction, here we examine how the energy barrier at B is modified externally by charging or by applying perpendicular electric field. First, we consider the case, where the system is negatively charged by implementing two excess electrons, i.e. $Q = -2$ electrons/cell. Under these circumstances, the energy barrier is dramatically lowered to ~ 0.1 eV. Overcoming this small barrier, the configuration-A proceeds to the configuration-D, whereby two coadsorbed OH form one H_2O molecule and one O atom adsorbed to the top site. Interestingly, for $Q = -4$ electrons/cell, the energy barrier completely disappears and the system directly passed from the configuration-A to the configuration-D forming again weakly bound H_2O and O atom adsorbed at the bridge site.

As shown in Fig. 11.12b, the OH–OH interaction under applied perpendicular E is reminiscent of the above charged cases. While the energy barrier of neutral system under $E_{field} = 0$ corresponds to the configuration-B, it increases to 0.95 eV under $E_{field} = +0.5$ V/Å. However, it decreases to 0.35 and to 0.05 eV under $E_{field} = -0.5$ V/Å and $E_{field} = -1.0$ V/Å, respectively. Under $E_{field} = +0.5$ V/Å, the system transforms to the configuration-E, whereby H_2O_2 is released.

The variations of interaction energies of two coadsorbed OH under excess charge and/or perpendicular electric field demonstrate the critical role played by hydroxyl groups in the reduction process of GOX. Since epoxy and hydroxyl groups coexist in GOX, desorption of oxygen adatom can take place through H_2O or H_2O_2 almost spontaneously under appropriate Q or E .

11.7 Conclusion

In order to understand recent experimental works reporting reduction of GOX by charging, by heating the system, or by applying electric field, we investigated the effects of these external agents on graphene. First, we discussed the limitations and artifacts of methods using plane wave (PW), and local orbital (AO), basis sets within periodic boundary conditions in treating the systems which are either charged or exerted by perpendicular electric field. While both methods have been successful in treating the atomic and electronic structure of various neutral or positively charged adatom-surface complexes, AO calculations hindering the

spurious vacuum charging within PBC are found to be appropriate to treat the systems, which are either negatively charged or are exerted by a perpendicular electric field. It appears that an artifact of AO calculations cancels out another artifact of PBC occurring in PW calculations. Our related analyses are believed to be very valuable for future theoretical studies including excess charge and electric field.

Determining that AO calculations can provide the proper treatment of one single graphene layer, which is charged negatively or exerted by a perpendicular electric field, we carried out structure optimized, self-consistent field calculations under specific constraints and NEB calculations between well defined initial and final configurations. We found that the strength of the bond between graphene and oxygen adatom is weakened with negative charging and/or with perpendicular electric field applied in the direction opposite to oxygen. Under negative perpendicular electric field and negative charging, both coexisting in the system the desorption and hence reduction of GOX can take place much easier. Despite the negative formation energy, the formation of oxygen molecule from individual oxygen adatoms adsorbed to graphene are hindered by the energy barriers of 0.8–1.3 eV. The calculated energy barriers are significantly lower than the sum of the binding energies of two single oxygen adatom due to the concerted action. In this respect, we note the possibility of another reaction path even with lower energy barrier.

We also considered interactions among coadsorbed H, O, and OH. Adsorbed or free hydrogen atoms can easily interact with oxygen adatom to form OH. Moreover, free or adsorbed hydrogen atoms can also interact with adsorbed OH to form H_2O . Adsorbed OH by themselves can diffuse relatively easy and interact with each other. However, an energy barrier of 0.4 eV hinders them to engage in a chemical process. We showed that by negatively charging the system or by applying perpendicular field one can suppress this energy barrier and promote the chemical reaction to form H_2O . H_2O by itself is very weakly bound to graphene and can desorb in ambient temperature. Each desorbed H_2O removes one oxygen atom from graphene oxide.

The calculated energy barriers on various reaction paths are shown to be lowered by charging the system with excess electrons or by applying a perpendicular electric field. This explains why the reduction of GOX through desorption of oxygen or O_2 is facilitated as reported by earlier experimental studies. Notably, perpendicular electric field can be used also for side specific functionalization of single and multi layered nanostructures. Finally, we note that the levels of charging or applied electric field comparable to those used in this study can be achieved locally in experiments through the tip of STM or femtosecond laser systems.

Acknowledgments The Authors acknowledge useful and stimulating discussions with Professor A. Dana and Mr O. Ekiz. Dr H.H. Gürel acknowledges the support of TUBITAK-BIDEB. Part of the computational resources has been provided by TUBITAK ULAKBIM, High Performance and Grid Computing Center (TR-Grid e-Infrastructure) and UYBHM at Istanbul Technical University through Grant No. 2-024-2007. The content of this review has been published in [29, 31, 32, 42, 43]

References

1. H.C. Sniepp, J.-L. Li, M.C. McAllister, H. Sai, M. Herrera-Alonso, D.H. Adamson, R.K. Prud'homme, R. Car, D.A. Saville, I.A. Aksay, Functionalized single graphene sheets derived from splitting graphite oxide. *J. Phys. Chem. B* **110**, 8535–8539 (2006)
2. J.-L. Li, K.N. Kudin, M.J. McAllister, R.K. Prudhomme, I.A. Aksay, R. Car, Oxygen-driven unzipping of graphitic materials. *Phys. Rev. Lett.* **96**, 176101–176105 (2006)
3. D.W. Boukhvalov, M.I. Katsnelson, Modeling of graphite oxide. *J. Am. Chem. Soc.* **130**, 10697–10701 (2008)
4. J. Yan, L. Xian, M.Y. Chou, Structural and electronic properties of oxidized graphene. *Phys. Rev. Lett.* **103**, 086802–086806 (2009)
5. J.-A. Yan, M.Y. Chou, Oxidation functional groups on graphene: structural and electronic properties. *Phys. Rev. B* **82**, 125403–125413 (2010)
6. L. Wang, Y.Y. Sun, K. Lee, D. West, Z.F. Chen, J.J. Zhao, S. Zhang, Stability of graphene oxide phases from first-principles calculations. *Phys. Rev. B* **82**, 161406–161409 (2010)
7. A. Bagri, C. Mattevi, M. Acik, Y.J. Chabal, M. Chhowalla, V.B. Shenoy, Structural evolution during the reduction of chemically derived graphene oxide. *Nat. Chem.* **2**, 581–587 (2010)
8. H.J. Xiang, S.-H. Wei, X.G. Gong, Structural motifs in oxidized graphene: a genetic algorithm study based on density functional theory. *Phys. Rev. B* **82**, 035416–5 (2010)
9. E.C. Mattson, H. Pu, S. Cui, M.A. Schofield, S. Rhim, G. Lu, M.J. Nasse, R.S. Ruoff, M. Weinert, M. Gajdardziska-Josifovska, J. Chen, C.J. Hirschmugl, Evidence of nanocrystalline semiconducting graphene monoxide during thermal reduction of graphene oxide in vacuum. *ACS Nano* **5**, 9710–9717 (2011)
10. S. Pei, H. Cheng, The reduction of graphene oxide. *Carbon* **50**, 3210–3228 (2012)
11. S. Mao, H. Pu, J. Chen, Graphene oxide and its reduction: modeling and experimental progress. *RSC Adv.* **2**, 2643–2662 (2012)
12. H. Huang, Z. Li, J. She, W. Wang, Oxygen density dependent band gap of reduced graphene oxide. *J. Appl. Phys.* **111**, 054317–4 (2012)
13. S. Kim, S. Zhou, Y. Hu, M. Acik, Y.J. Chabal, C. Berger, W. de Heer, A. Bongiorno, E. Riedo, Room-temperature metastability of multilayer graphene oxide films. *Nat. Mater.* **11**, 544–549 (2012)
14. T. Sun, S. Fabris, Mechanisms for oxidative unzipping and cutting of graphene. *Nano Lett.* **12**, 17–21 (2012)
15. M. Topsakal, S. Ciraci, Domain formation on oxidized graphene. *Phys. Rev. B* **86**, 205402–205413 (2012)
16. R. Larciprete, S. Fabris, T. Sun, P. Lacovig, A. Baraldi, S. Lizzit, Dual path mechanism in the thermal reduction of graphene oxide. *J. Am. Chem. Soc.* **133**, 17315–17321 (2011)
17. K. Novoselov, A. Geim, S. Morozov, D. Jiang, Y. Zhang, S. Dubonos, I. Grigorieva, A. Firsov, Electric field effect in atomically thin carbon films. *Science* **306**, 666–669 (2004)
18. D.A. Dikin, S. Stankovich, E.J. Zimney, R. Piner, G.H.B. Dommett, G. Evmenenko, S.T. Nguyen, R.S. Ruoff, Preparation and characterization of graphene oxide paper. *Nature* **448**, 457–460 (2007)
19. P.P. Yao, P.L. Chen, L. Jiang, H.P. Zhao, H.F. Zhu, D. Zhou, W.P. Hu, B.H. Han, M.H. Liu, Electric current induced reduction of graphene oxide and its application as gap electrodes in organic photoswitching devices. *Adv. Mater.* **22**, 5008 (2010)
20. Z. Wei, D. Wang, S. Kim, S.Y. Kim, Y. Hu, M.K. Yakes, A.R. Laracuente, Z. Dai, S.R. Marder, C. Berger, W.P. King, W.A. de Heer, P.E. Sheehan, E. Riedo, Nanoscale tunable reduction of graphene oxide for graphene electronics. *Science* **328**, 1373–137 (2010)
21. O.O. Ekiz, M. Urel, H. Guner, A.K. Mizrak, A. Dana, Reversible electrical reduction and oxidation of graphene oxide. *ACS Nano* **5**, 2475–2482 (2011)
22. P. Blochl, Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994)
23. J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996)

24. G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996)
25. J.M. Soler, E. Artacho, J.D. Gale, A. Garcia, J. Junquera, P. Ordejón, D. Sánchez-Portal, The SIESTA method for Ab initio order-N materials simulation. *J. Phys.: Condens. Matter* **14** (11), 2745–2779 (2002)
26. G. Makov, M.C. Payne, Periodic boundary conditions in ab initio calculations. *Phys. Rev. B* **51**, 4014–4022 (1995)
27. H.J. Monkhorst, J.D. Pack, Special points for Brillouin-zone integrations. *Phys. Rev. B* **13**, 5188 (1976)
28. N. Troullier, J.L. Martins, Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B* **1991**, 43 (1993)
29. M. Topsakal, H.H. Gürel, S. Ciraci, Effects of charging and electric field on graphene oxide. *J. Phys. Chem. C* **117**, 5943 (2013)
30. H.H. Gürel, V.O. Özcelik, S. Ciraci, Effects of charging and perpendicular electric field on the properties of silicene and germanene. *J. Phys.: Condens. Matter* **25**, 305007 (2014)
31. H.H. Gürel, S. Ciraci, Enhanced reduction of graphene oxide by means of charging and electric fields applied to hydroxyl groups. *J. Phys.: Condens. Matter* **25**, 435304 (2014)
32. H.H. Gürel, S. Ciraci, Effects of charging and electric field on graphene functionalized with titanium. *J. Phys.: Condens. Matter* **25**, 275302 (2014)
33. R. Poloni, A.S. Miguel, M.V. Fernandez-Serra, A first principles study of the effect of charge doping on the 1D polymerization of C60. *J. Phys.: Condens. Matter* **24**, 095501–095507 (2012)
34. C. Attaccalite, L. Wirtz, M. Lazzeri, F. Mauri, A. Rubio, Doped graphene as tunable electron-phonon coupling material. *Nano Lett.* **10**, 1172–1176 (2010)
35. B. Aradi, B. Hourahine, T. Frauenheim, DFTB+, a sparse matrix-based implementation of the DFTB method. *J. Phys. Chem. A* **111**, 5678–5684 (2007)
36. A.S. Bernard, I.K. Snook, Transformation of graphene into graphane in the absence of hydrogen. *Carbon* **48**, 981–987 (2010)
37. B. Delley, An all-electron numerical method for solving the local density functional for polyatomic molecules. *J. Chem. Phys.* **92**, 508–517 (1990)
38. Z.M. Ao, A.D. Hernandez-Nieves, F.M. Peeters, S. Lia, The electric field as a novel switch for uptake/release of hydrogen for storage in nitrogen doped graphene. *Phys. Chem. Chem. Phys.* **14**, 1463–1467 (2012)
39. K.T. Chan, H. Lee, M.L. Cohen, Gated adatoms on graphene studied with first-principles calculations. *Phys. Rev. B* **83**, 035405–035410 (2011)
40. V.W. Brar, R. Decker, H.M. Solowan, Y. Wang, L. Maserati, K.T. Chan, H. Lee, C.O. Girit, A. Zettl, S.G. Louie, M.L. Cohen, M.F. Crommie, Gate-controlled ionization and screening of cobalt adatoms on a graphene surface. *Nat. Phys.* **7**, 43 (2011)
41. A.M. Suarez, L.R. Radovic, E. Bar-Ziv, J.O. Sofo, Gate voltage control of oxygen diffusion on graphene. *Phys. Rev. Lett.* **106**, 146802–146804 (2011)
42. M. Topsakal, S. Ciraci, Static charging of graphene and graphite slabs. *Appl. Phys. Lett.* **98**, 131908–131911 (2011)
43. M. Topsakal, S. Ciraci, Effects of static charging and exfoliation of layered crystals. *Phys. Rev. B* **85**, 045121–045127 (2012)
44. K.T. Chan, J.B. Neaton, M.L. Cohen, First-principles study of metal adatom adsorption on graphene. *Phys. Rev. B* **77**, 235430–12 (2008)
45. N.D. Drummond, V. Zolyomi, V.I. Falko, Electrically tunable band gap in silicene. *Phys. Rev. B* **85**, 075423–7 (2012)
46. P. Hohenberg, W. Kohn, Inhomogeneous Electron Gas *Phys. Rev.* **136**, B864 (1964); W. Kohn, L.J. Sham, Self-consistent equations including exchange and correlation effects. *ibid.* **140**, A1133 (1965)
47. X. Lin, J. Ni, C. Fang, Adsorption capacity of H₂O, NH₃, CO, NO₂ on pristine graphene. *J. Appl. Phys.* **113**, 034306 (2013)

48. O. Leenaerts, B. Partoens, F.M. Peeters, Adsorption of H₂O, NH₃, CO, NO₂, and NO on graphene: a first-principles study. *Phys. Rev. B* **77**, 125416 (2007)
49. W.W. Cai, R.D. Piner, F.J. Stadermann, S. Park, M.A. Shaibat, Y. Ishii, D.X. Yang, A. Velamakanni, S.J. An, M. Stoller, Synthesis and solid-state NMR structural characterization of C-13-labeled graphite oxide. *Science* **321**, 1815–1817 (2008)
50. T. Szabo, O. Berkesi, P. Forgo, K. Josepovits, Y. Sanakis, D. Petridis, I. Dekany, Evolution of surface functional groups in a series of progressively oxidized graphite oxides. *Chem. Mater.* **18**, 2740–2749 (2006)
51. A. Jelea, F. Marinelli, Y. Ferro, A. Allouche, C. Brosset, Quantum study of hydrogen oxygen graphite interactions. *Carbon* **42**, 3189–3198 (2004)
52. R.J.W.E. Lahaye, H.K. Jeong, C.Y. Park, Y.H. Lee, Density functional theory study of graphite oxide for different oxidation levels. *Phys. Rev. B* **79**, 125435 (2009)
53. M.C. Kim, G.S. Hwang, R.S.J. Ruoff, Epoxide reduction with hydrazine on graphene: a first principles study. *Chem. Phys.* **131**, 064704 (2009)
54. S.C. Xu, S. Irle, D.G. Musaev, M.C.J. Lin, Quantum chemical study of the dissociative adsorption of OH and H₂O on pristine and defective graphite (0001) surfaces: reaction mechanism and kinetics. *Phys. Chem. C* **111**, 1355–1365 (2007)
55. P.A.J. Denis, Density functional investigation of thioepoxidated and thiolated graphene. *Phys. Chem. C* **113**, 5612–5619 (2009)
56. Q. Zhu, Y.H. Lu, J.Z. Jiang, Stability and properties of two-dimensional graphene hydroxide. *J. Phys. Chem. Lett.* **2**, 1310–1314 (2011)
57. H. Khantha, N.A. Cordero, L.M. Molina, D.A. Alonso, L.A. Grifalco, Interaction of lithium with graphene: an ab initio study. *Phys. Rev. B* **70**, 125422 (2004)
58. M. Topsakal, H. Sahin, C. Ciraci, Graphene coatings: an efficient protection from oxidation. *Phys. Rev. B* **85**, 155445–155447 (2012)
59. S. Grimme, Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **27**, 1787 (2006)
60. H.H. Gürel, V.O. Özcelik, S. Ciraci, Dissociative adsorption of molecules on graphene and silicene. *J. Phys. Chem. C* **118**(47) (2014)
61. L. Jelaoua, V. Sidis, DFT investigation of the adsorption of atomic hydrogen on a cluster-model graphite surface. *Chem. Phys. Lett.* **300**, 157 (1999)
62. X. Sha, B. Jackson, First-principles study of the structural and energetic properties of H atoms on a graphite (0001) surface. *Surf. Sci.* **496**, 318 (2002)
63. S. Casolo, O.M. Lrvik, R. Martinazzo, G.F. Tantardini, Understanding adsorption of hydrogen atoms on graphene. *J. Chem. Phys.* **30**, 054704 (2009)
64. V.V. Ivanovskaya, A. Zobelli, D. Teillet-Billy, N. Rougeau, V. Sidis, P.R. Briddon, Hydrogen adsorption on graphene: a first principles study. *Eur. Phys. J. B* **76**, 481 (2010)
65. S.H. Vosko, L. Wilk, M. Nusair, Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **58**, 1200 (1980)
66. D. Henwood, J.D. Carey, Ab initio investigation of molecular hydrogen physisorption on graphene and carbon nanotubes. *Phys. Rev. B* **75**, 245413 (2007)
67. S. Cahangirov, M. Topsakal, E. Akturk, H. Sahin, S. Ciraci, Two- and one-dimensional honeycomb structures of silicon and germanium. *Phys. Rev. Lett.* **102**, 236804–236811 (2009)
68. H. Sahin, S. Cahangirov, M. Topsakal, Bekaroglu, E.; Akturk, E.; Senger R.T.; Ciraci, S. Monolayer honeycomb structures of group-IV elements and III-V binary compounds: first-principles calculations. *Phys. Rev. B* **80**, 155453 (2009)
69. C. Ataca, H. Sahin, S. Ciraci, Stable, single-layer MX₂ transition-metal oxides and dichalcogenides in a honeycomb-like structure. *J. Phys. Chem. C* **116**, 8983–8999 (2012)
70. P.P. Yao, P.L. Chen, L. Jiang, H.P. Zhao, H.F. Zhu, D. Zhou, W.P. Hu, B.H. Han, M.H. Liu, Electric current induced reduction of graphene oxide and its application as gap electrodes in organic photoswitching devices. *Adv. Mater.* **22**, 5008 (2010)

Chapter 12

Structural and Optical Properties of Tungsten Oxide Based Thin Films and Nanofibers

E.O. Zayim and A. Tabatabaei Mohseni

Abstract Tungsten oxide nanomaterials confined to one and two dimensions can be prepared with tungsten metal powder, tungsten chloride, peroxotungstic acid and acetylated peroxotungstic acid precursors by sol-gel, evaporation and electrodeposition techniques. Nanofibers and nanowires of tungsten oxide are synthesized by organic/inorganic blend of tungsten hexachloride, tungsten metal powder and polyvinylpyrrolidone with electrospinning technique. Standard and mesoporous tungsten oxide thin films are prepared from an ethanolic solution of tungsten hexachloride. Several polymers were employed as a template to generate the mesoporous structure. Additionally a detailed systematic study of the evaporated tungsten oxide thin films has been carried out at progressively increasing temperatures. Overall, the optical, electrochemical and structural properties of the deposited films were examined in both liquid and solid electrolytes. All solid electrochromic devices were fabricated using tungsten oxide active electrochromic layers. The fabrication and evaluation of a prototype solid-state electrochromic device are also described.

12.1 Introduction

Tungsten oxide (WO_3), also known as tungsten trioxide or tungstic anhydride, is an n-type semiconductor with a wide band gap of 2.5–3.2 eV [1]. Tungsten trioxide is a thermally stable and water insoluble. Tungsten oxides are unique materials that have been extensively studied because of their unique physical and chemical properties and widespread applications. Industrial applications of tungsten oxide

E.O. Zayim (✉)

Physics Department, Istanbul Technical University, Maslak, Istanbul, Turkey
e-mail: ozesra@itu.edu.tr

A. Tabatabaei Mohseni

Nanoscience & Nanotechnology Program, Istanbul Technical University, Maslak, Istanbul, Turkey

coatings have encouraged many research groups to investigate this specific material from different perspectives.

Thin films and nanomaterials of tungsten trioxide are rigorously investigated for chromogenic devices, photo- and electrochromic “smart” windows, erasable optical-storage devices, photocatalysts, biological elements [2], electrodes for solar cells, lithium-ion batteries [3, 4], supercapacitors [5], x-ray screen phosphors, fireproofing fabrics, humidity and temperature sensors and in gas-sensing elements [6, 7]. They exhibit further important properties and functionalities. Recent works present a general review of nanostructured WO_x , their properties, methods of synthesis, and a description of how they can be used in unique ways for different applications. This oxide film shows both reversible wettability conversion between superhydrophobic and super-hydrophilic states.

12.1.1 Amorphous and Crystalline Tungsten Oxide Based Nanomaterials

Tungsten oxide based nanomaterials prepared by various techniques, such as thermal evaporation, chemical vapor deposition, sputtering, and sol–gel methods, possess different micro- and nanostructural properties. As in every chemical and physical deposition method, heat treatment plays an important role, because structural, optical and electrical properties will change by annealing. Therefore, it is important to characterize WO_3 films as a function of both the fabrication and the annealing conditions. WO_3 film grown at substrate temperature which is less than 300°C is amorphous and transparent. Crystallization of the films can be observed at a temperature of 350°C or higher [8].

Presently, scientists and engineers are paying much attention to nanostructured oxides with zero-dimensional quantum dots, one-dimensional nanowires and nanorods, and two-dimensional nanosheets and nanodisks, due to their novel physical and chemical properties which are different from their corresponding bulks and amorphous tungsten oxide films [9].

Tungsten trioxide (WO_3) has a perovskite structure (with no A-site cations) [10] which is referred to any material with the same type of crystal structure with general chemical formula of ABX_3 , where ‘A’ and ‘B’ are two cations of very different sizes, and X is an anion that bonds to both. As shown in Fig. 12.1 the tungsten ions occupy the corners of a primitive unit cell and oxygen ions bisect the unit cell edges [11].

Investigations by Ramana et al. on thin film of WO_3 deposited by pulsed-laser deposition (PLD) technique, represent the following crystal structure sequences depending on their temperature increment, which is also true for bulk tungsten oxides: amorphous ($<300^\circ\text{C}$) $\rightarrow$ monoclinic (350°C) $\rightarrow$ orthorhombic (500°C) $\rightarrow$ hexagonal (740°C) $\rightarrow$ tetragonal [8].

Fig. 12.1 Unit cell for the perovskite lattice and the octahedral symmetry

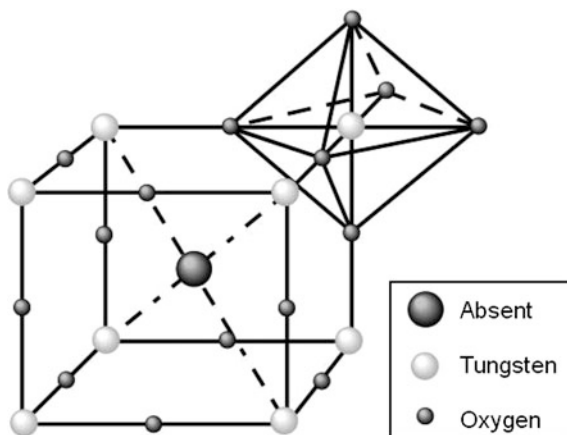
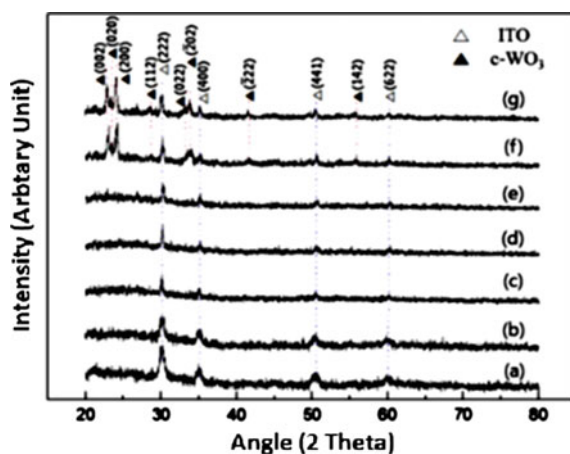


Fig. 12.2 X-ray diffraction patterns of the PLD WO_{3-x} thin films: **a** ITO; **b** Room Temperature; **c** annealed at 100 °C; **d** annealed at 200 °C; **e** annealed at 300 °C; **f** annealed at 400 °C; **g** annealed at 500 °C



The physical properties of slightly oxygen-deficient analogues tungsten trioxide, WO_{3-x} , possess several crystallographic phases [12]. WO_{3-x} films have been extensively studied for their chromism, which exhibit different chromic properties for different levels of oxygen deficiency: $x > 0.5$ films are metallic and conductive, $0.3 < x < 0.5$ films are blue and conductive, and $x < 0.3$ films are transparent and resistive; these findings are independent of the film preparation method [13].

Figure 12.2 demonstrates that amorphous WO_{3-x} thin films deposited by PLD remain in an amorphous phase up to 300 °C. The thin films form a crystalline structure by post annealing at 400 and 500 °C [14].

Reactive RF magnetron sputtering can be applied for depositing a slightly reduced amorphous WO_{3-x} comes up with the same result for crystallization of the

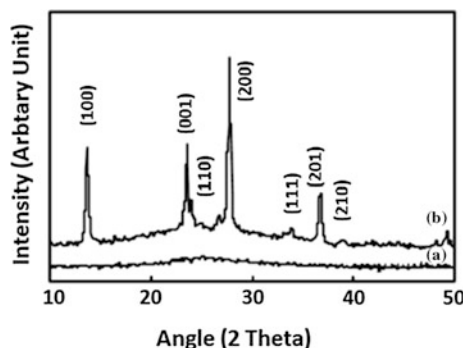
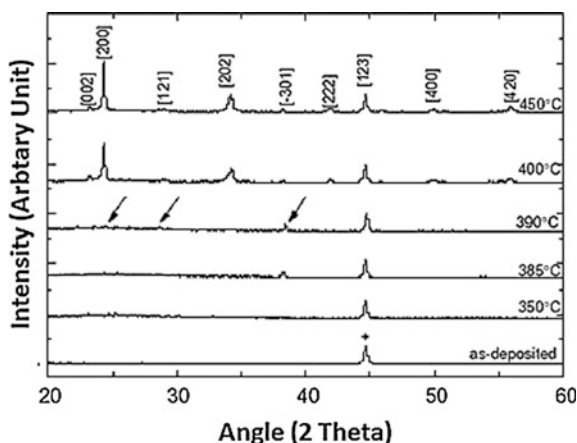


Fig. 12.3 XRD patterns for as-deposited and annealed WO_{3-x} films coated by reactive RF magnetron sputtering, **a** Room Temperature; **b** annealed at 400 °C

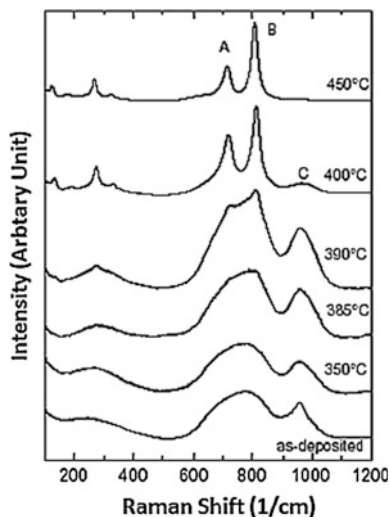
Fig. 12.4 XRD patterns of thermal evaporated tungsten oxide thin films annealed at different temperatures. (+) indicates the peaks from the substrate



film by annealing at 400 °C in vacuum [15]. Figure 12.3 shows the XRD patterns for the as-prepared and annealed tungsten oxide films. No XRD peaks were observed for the as-prepared films. Annealed films have a hexagonal single phase.

Thermally evaporated tungsten oxide films were obtained from WO_3 powder. As-deposited amorphous WO_{3-x} films were subjected to different annealing temperatures to obtain various crystalline phases. They were annealed at progressively increasing temperatures ranging from 350 to 450 °C in oxygen ambient. Figure 12.4 demonstrates that amorphous WO_{3-x} films remain in an amorphous phase up to 385 °C and begin to crystallize at 390 °C and then are completely crystallized at 450 °C. The peaks in the XRD pattern can be attributed to a monoclinic crystalline WO_3 structure.

Fig. 12.5 Raman spectra of thermal evaporated tungsten oxides thin films annealed at different temperatures



Raman spectra of the thermal evaporated WO_{3-x} films are shown in Fig. 12.5. The spectrum of an as-deposited film (a) shows a broad peak at 770 cm^{-1} , due to the $\text{W}^{6+}\text{-O}$ bonds. There is also a relatively sharp peak at 950 cm^{-1} , which has been assigned to the $\text{W}^{6+}=\text{O}$ stretching mode of terminal oxygen atoms. In addition, there is a weaker, broad peak at 220 cm^{-1} , which has been appeared from the presence of W^{4+} states. The peak at 770 cm^{-1} gets sharper and one new peak at 719 cm^{-1} develops at 385°C . At 390°C new peaks at 719 and 807 cm^{-1} develop. At 400°C annealing temperature, these peaks get sharper and the 950 cm^{-1} (C) component decreases. The two strongest peaks appear at 719 (A) and 807 (B) cm^{-1} in the Raman spectrum of crystalline WO_3 . Also two new peaks at 127 and 330 cm^{-1} develop in crystalline films. Overall, the amorphous WO_3 film begins to crystallize at 390°C and is completely crystallized at 450°C . A peak at 950 cm^{-1} , which is a key characteristic of the amorphous phase, remains up to 400°C indicating the coexistence of both crystalline and amorphous phases [16].

12.2 Tungsten Oxide Based Nanomaterials

Nanoscaled materials produced based on tungsten oxide have various physical and chemical properties according to their crystallography, confinement in one-dimension or two-dimensions, oxygen deficiency and consequently the band gap. In the following we will take a look at properties of tungsten oxide in thin film and nanofiber forms.

12.2.1 *Tungsten Oxide Based Thin Films and Mesoporous Thin Films*

Surfactants are able to modify and control the properties of electrode surfaces. The influence of surface active agents on the kinetics of electron transfer reactions at electrodes have been investigated in the past few decades [17]. Since the discovery of surfactant templated mesoporous materials, the templating method has been extended to a range of transition metal oxides [16]. Mesoporous tungsten oxide thin films prepared by nonionic surfactants as templates, show high-rate ion-insertion performance when used as electrochromic layers.

Different types of polymers were employed as a template to generate the mesoporous structure. These polymers are poly(ethylene glycol) (PEG), polyethylene glycol-ran-propylene glycol and polystyrene-co-allyl-alcohol. Here the application of PEG is given as an example. To avoid the varying degrees of crystallinity as a result of thermal treatment, a UV illumination method has been employed to remove the polymer surfactant. This room temperature approach uses ozone, generated during UV illumination in air to oxidize the organic compounds.

The electrochromic and optical properties of the mesoporous films are described and compared to standard sol-gel WO_3 films. Results are also presented on the samples prepared by thermal treatment. It has been demonstrated that the UV illumination treatment is a superior method to remove templates, leading to more detailed investigation of the effect of mesoporosity on the electrokinetics of ion insertion into WO_3 films. These mesoporous materials exhibit superior high rate ion-insertion performance when used as electrochromic layers, which is attributed to the high surface area of mesoporous WO_3 .

The introduction of mesoporosity into WO_3 layers has been predicted to improve their functional film properties due to the porous nature of these materials and better penetration of electrolytes into the inorganic framework. The improved electrochromic (EC) properties of mesoporous tungsten oxide films produced via these templates are described and compared to standard sol-gel films of tungsten oxides.

SEM was used to study the surface structure of all deposited layers shown in Figs. 12.6 and 12.7. The images indicate that polymer generates a remarkable

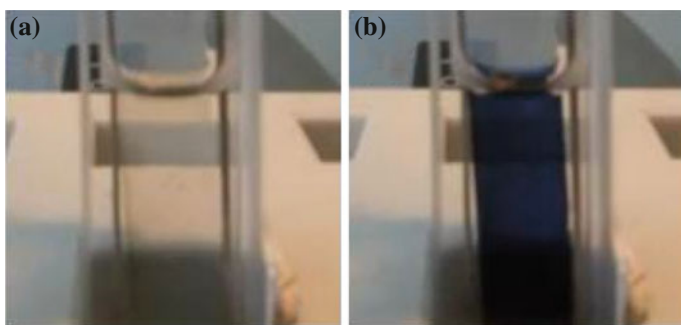


Fig. 12.6 SEM image of **a** as-deposited tungsten oxide film; **b** calcined tungsten oxide film at 400 °C

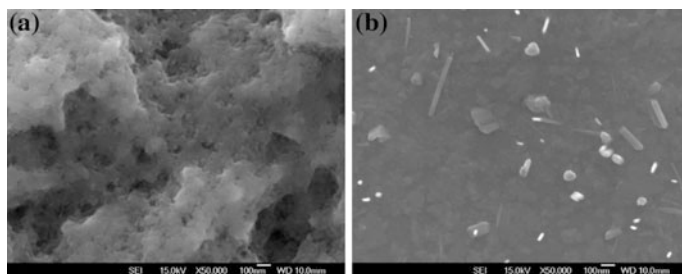


Fig. 12.7 SEM image of **a** as-deposited polymer-added tungsten oxide film; **b** calcined polymer-added tungsten oxide film at 400 °C

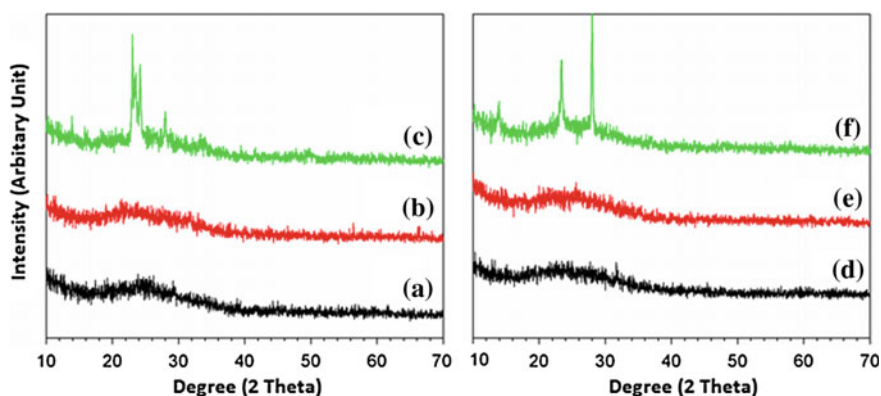
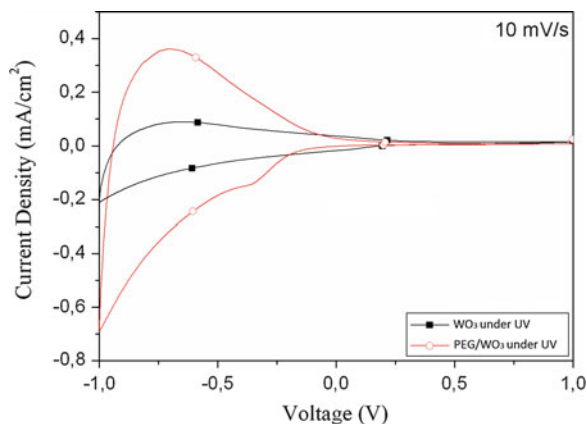


Fig. 12.8 XRD patterns of WO_{3-x} thin films: **a** As-deposited tungsten oxide film; **b** Tungsten oxide film treated by UV illumination; **c** annealed tungsten oxide film at 400 °C; **d** As-deposited polymer-added tungsten oxide film; **e** Polymer-added tungsten oxide film treated by UV illumination; **f** annealed polymer-added tungsten oxide film at 400 °C

change on the structural properties of tungsten oxide films. As shown in SEM results, while as-deposited tungsten oxide film has a porous structure, the surface of the as-deposited polymer-added tungsten oxide films is uniform and flat. Heat treatment causes a remarkable difference in the films' images.

Characterizations of tungsten oxide films were carried out with XRD and FTIR measurements. Standard and mesoporous tungsten oxide films remain almost amorphous up to 400 °C and crystallize at 400 °C. As-deposited and UV illuminated films are amorphous. Polymer-added tungsten oxide films have changed the crystal phases depending on the polymer matrix. Using polymer as a surfactant in the tungsten oxide films may retard crystallization and created different crystal phases. Calcined standard tungsten oxide films have triclinic crystalline structure. Following spectral deconvolution analysis of both films, their different crystal structures are presented in Fig. 12.8. PEG polymers are in linear and rigid rod form;

Fig. 12.9 Cyclic Voltammetry (CV) of tungsten oxide film treated by UV illumination (WO_3 under UV) and polymer-added tungsten oxide film treated by UV illumination (PEG/ WO_3 under UV)



this may positively affect the speed and the rate of crystallization of the structures in which PEG polymers are added.

All as-deposited films were then exposed to electrochemical and electrochromic investigations using LiClO_4 in propylene carbonate as an electrolyte. The CV of tungsten oxide films are given in Fig. 12.9. The results demonstrated that the polymer causes modifications of the electrochemical and structural properties of the films.

It is observed that better CV results (charge insertion/extraction) were obtained for mesoporous films as shown in the Fig. 12.9. Supposably using PEG polymer as a surfactant increases the active surface area. These mesoporous films exhibit superior high-rate ion-insertion performance in lithium liquid electrolyte as expected.

12.2.2 Tungsten Oxide Based Nanofibers and Nanowires

One-dimensional (1D) nanostructures of tungsten trioxide include nanowires [18], nanotubes [19], nanofibers [7], nanorods [20] and nanoribbons [21] which offer the best morphology for most of the electro-functional devices [22].

Tungsten oxide nanomaterials can be produced via various techniques including electrochemical etching [23], chemical vapor deposition (CVD) [24], hydrothermal reaction [25] and etc.

Among these methods and techniques, electrospinning is one of the most widely used processes for the production of nanofibers known since the 1930s. This technique employs electrostatic forces for stretching the viscoelastic fluid [26].

Electrospinning offers a relatively simple and versatile method for generating fibular mesostructures. This process involves the production of continuous one-dimensional nanofibers under electrostatic force of the charges on the surface of a liquid droplet in an electric field as strong as several $\text{kV} \cdot \text{cm}^{-1}$.

Here, method of producing nanofibers and nanowires of tungsten oxide via electrospinning has been applied using two different precursors.

WO₃ fibers can be produced by electrospinning technique using tungsten hexachloride (WCl₆) and tungsten (W) metallic powder as precursors. Surface and structural characterizations were performed by SEM, XRD and FTIR to demonstrate the crystallinity and structure of the fibers.

The electrospinning technique requires a viscose polymeric solution. This organic/inorganic blend contains precursors and polymer ingredients and their solvents. The sol-gel W metal powder and WCl₆ precursor mixed with a 10 % polyvinylpyrrolidone (PVP) solution in ethanol were loaded into a plastic syringe after stirring for minimum 5 hours. PVP has been preferred because it is a suitable capping reagent for various metal nanoparticles with low environmental and health risks.

12.2.2.1 Tungsten Oxide Fibers with Metallic Tungsten Precursor

To prepare the organic/inorganic blend for the electrospinning process, 10 ml of cold 30 % H₂O₂ solution was added to 0.7 g metallic tungsten powder. After stirring for about 4 hours, the powder dissolved completely and a faint yellowish solution (tungstic acid) was obtained. Then 10 ml ethanol and 1.0 g PVP were added into the above solution, followed by magnetic stirring for 2 hours. The solution was then loaded into a plastic syringe connected to a 1 cm short needle made of copper by a plastic pipe. The needle was connected to a high-voltage supply that is capable of generating DC voltages up to 35 kV. In this experiment a voltage of 15 kV was applied for electrospinning conducted in the air at room temperature. Annealing electrospun fibers in air at 500 °C for 2 h completely evaporates the solvents and polymer to attain calcined WO₃ nanofibers.

Microstructural properties of tungsten oxide fibers were investigated by FE-SEM. Homogenous nanofibers of tungsten oxide were prepared using metallic tungsten precursor as seen in the SEM images presented in Fig. 12.10. The diameter of the nanofibers are in the range of 112–184 nm. SEM images of calcined nanofibers of tungsten oxide are shown in Fig. 12.11. The thickness of the fibers are observed to be 35 % reduced that is between 42 and 152 nm.

FTIR analysis of as-prepared electrospun tungsten oxide nanofibers by a metallic tungsten precursor is given in Fig. 12.12. Inorganic compounds have vibrational bands mainly below 1200 cm⁻¹. The 816 cm⁻¹ band is assigned to the out of stretching vibrations of W–O–W mode, when hydrogen is located at a coplanar square of oxygen atoms. A relatively weak band at 1073 cm⁻¹, which is assigned to the plane deformational (bending) W–OH mode, was found. The peak at 1424 cm⁻¹ is assigned to ν(OH) and δ(OH) in OH, W–O group. In the frequency range of 400–1100 cm⁻¹, the peak in 977 cm⁻¹ (W=O terminal modes of surface grains) and W–O–W bridging mode 902 cm⁻¹ exist exhibiting the stretching vibrations of W–O bounds [27]. In addition, the sharp peak in 1290 cm⁻¹ is assigned to C–N stretches of the aromatic amines group. Organic compounds usually have vibrational bands

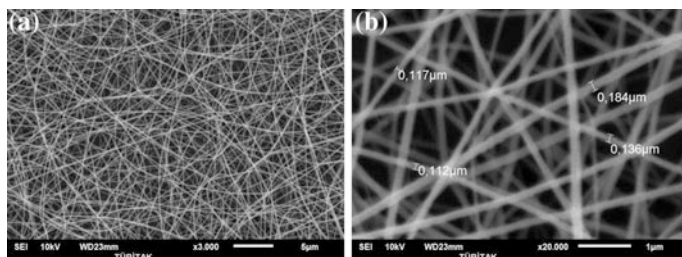


Fig. 12.10 SEM images of electrospun tungsten oxide nanofibers by W/PVP, **a** 5 μm scale, **b** 1 μm scale

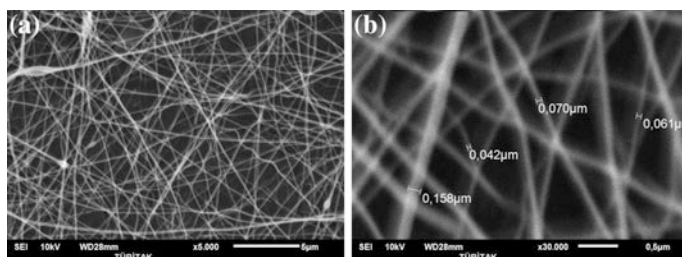
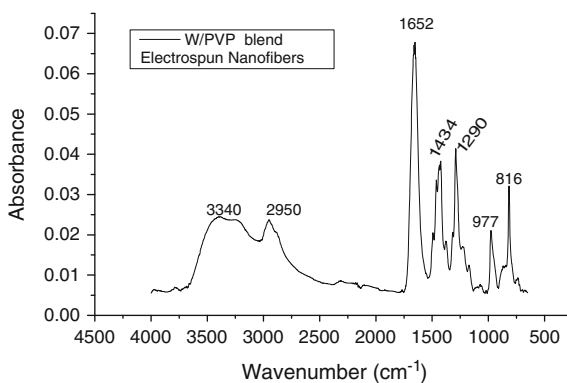


Fig. 12.11 SEM images of calcined tungsten oxide nanofibers by W/PVP, **a** 5 μm scale, **b** 0.5 μm scale

Fig. 12.12 FTIR of electrospun tungsten oxide nanofibers of W/PVP blend



mainly after 1500 cm⁻¹. The sharp peak around 1652 cm⁻¹ is assigned to C=O stretch of α, β-unsaturated aldehydes the ketone functional group. The medium peak at 2950 cm⁻¹ reveals the C-H stretch from an alkane group and finally the sharp and broad shoulder around 3340 cm⁻¹ is assigned to O-H stretch, from H-bonded alcohol the phenol functional group.

12.2.2.2 Tungsten Oxide Fibers with Tungsten Hexachloride Precursor

Tungsten hexachloride ($\text{WCl}_6 \geq 99.9\%$ trace metals basis) powder was dissolved in ethanol ($\text{CH}_3\text{CH}_2\text{OH}$), then mixed with polyvinylpyrrolidone (PVP) where N, N-dimethylformamide (DMF) was used as its solvent. The electrospinning process was held at room temperature using 20 kV DC voltage while the distance between the spinning nozzle and the collection plate was 15 cm, feeding rate was 6 $\mu\text{l}/\text{min}$ and substrate was a glass on aluminum foil.

SEM images of resultant electrospun nanofibers of the blend are presented in Figs. 12.13 and 12.14. The diameters of the composite nanofibers are distributed in the range of 190 and 350 nm with an average diameter of 270 nm.

To eliminate polymeric compounds from the film structure and obtain an inorganic surface the nanofiber mat was calcined. After calcination at 500 $^\circ\text{C}$ for 1 h, homogenous thin fibers in the range of 57 and 110 nm were obtained with an average diameter of 85 nm as shown in Fig. 12.14. After the calcination process, the average diameter of the nanofibers was reduced by nearly 70 %, proving that organic compounds have been removed from the film structure and tungsten oxide is the dominant remaining material in the surface. After calcination at 500 $^\circ\text{C}$ for 1 h, we obtained the thinner fibers as shown in SEM images.

FTIR analysis of as-prepared electrospun nanofibers of tungsten oxide using WCl_6/PVP blend is given in Fig. 12.15. The 679 cm^{-1} band is assigned to the out of plane deformation W–O–W mode, when hydrogen is located at a coplanar square of oxygen atoms, the 816 cm^{-1} peak is assigned to the out of plane stretching vibrations of the W–O–W mode. A relatively weak band at 1062 cm^{-1} , which is assigned to the plane deformational (bending) W–OH mode, was found. This is due to the creation of weakly bonded W–OH groups is formed in the as-deposited film. In the frequency range from 400 to 1100 cm^{-1} , a shoulder is observed around 979 cm^{-1} (W=O terminal modes of surface grains) and a W–O–W bridging mode at 902 cm^{-1} [27]. Organic compounds mostly have vibrational bands after 1500 cm^{-1} . The sharp peak around 1690 cm^{-1} is assigned to C=O stretch of α , β -unsaturated aldehydes, the ketone functional group. The medium peak at 2960 cm^{-1} reveals the C–H stretch from the alkane group and finally the sharp and broad shoulder around

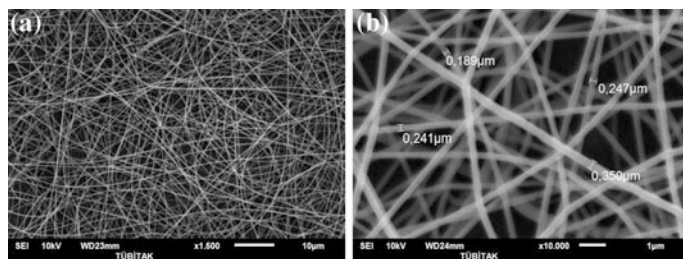


Fig. 12.13 SEM images of electrospun nanofibers by WCl_6/PVP , **a** 10 μm scale, **b** 1 μm scale

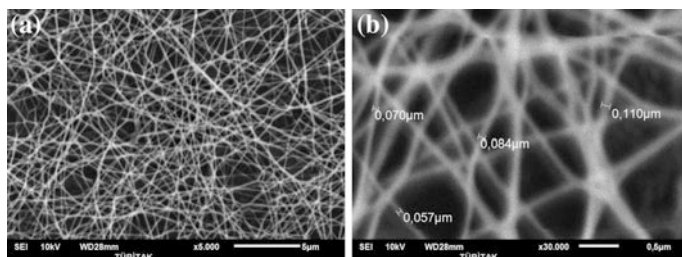
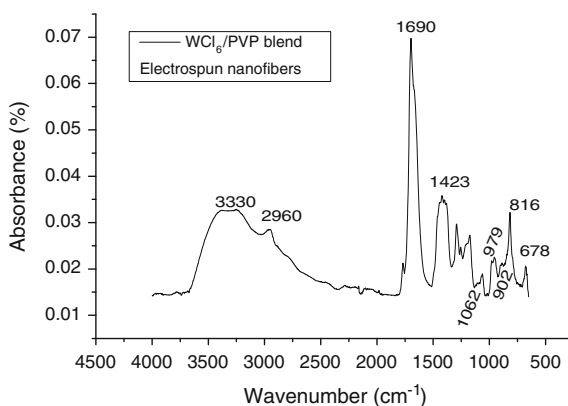


Fig. 12.14 SEM images of calcined nanofibers of tungsten oxide by WCl_6 /PVP blend

Fig. 12.15 FTIR of electrospun tungsten oxide nanofibers by WCl_6 /PVP blend



3330 cm^{-1} is assigned to O-H stretch, from H-bonded alcohol, the phenol functional group. In addition, the medium peak in 1423 is related to the C-H bend of one of the alkane groups.

To eliminate polymeric compounds from the film structure, and obtain an inorganic surface the nanofiber mat was calcined. After calcination at $500\text{ }^{\circ}\text{C}$ for 1 h, homogenous thin fibers in the range of 57 and 110 nm were obtained with an average diameter of 85 nm. After the calcination process, the average diameter of the nanofibers was reduced by nearly 70 %, proving that organic compounds have been removed from the film structure and tungsten oxide is the dominant remaining material in the surface.

Table 12.1 compares the diameter variations of WO_3 fibers obtained by two different precursors explained in the latter section. The average values indicate that as-prepared nanofibers of tungsten oxide polymeric hybrid are thinner for metallic tungsten precursor (148 vs. 270 nm), but for calcined tungsten oxide nanofibers, tungsten hexachloride precursor yields in thinner fibers and nanowires (83 vs. 100 nm).

Table 12.1 Diameter range of the nanofibers in different steps

Diameter	As-prepared nanofibers by WCl_6 precursor (nm)	Calcined nanofibers by WCl_6 precursor (nm)	As-prepared nanofibers by W metallic precursor (nm)	Calcined nanofibers by W metallic precursor (nm)
Min	189	57	112	42
Max	350	110	184	158
Average	270	83	148	100

12.3 Chromogenic Properties and Applications of Tungsten Oxide

Optically active thin film coatings can change their optical properties as a function of external stimuli. These substances, recently named ‘chromogenics’, include both inorganic and organic materials. Among various chromogenic property of films, well known ones are electrochromic (EC), thermochromic, photochromic and gaschromic. Materials with controllable light absorbance, transmittance or reflectance possess great technical relevance, because of their high potential applications. In the near future, optically active films may be utilized to regulate the throughput radiation energy for windows in the buildings and cars, to maintain comfortable lighting and temperature, in sunglasses, as an optically active filter, or in systems with variable reflectance, as automobile rear-view mirrors, in sensors, in detectors, in displays, as road signs, and so forth.

Tungsten oxide is the most investigated chromogenic material, especially used for the preparation of EC devices due to its high stability and fully reversible coloration ability.

12.3.1 Electrochromic Properties of Tungsten Oxide Films

The optical, electrochemical and structural properties of the deposited films were examined in both liquid and solid nafion electrolytes. Herein, electrochromic devices were fabricated by tungsten oxide active electrochromic layer prepared by tungsten chloride precursor and nafion electrolytes. The use of a polymer electrolyte, in contrast to commonly used liquid electrolytes, allows integration of the electrode into fully solid state devices with the following configuration: ITO/ WO_3 /Nafion/ITO.

The colored and bleached state in nafion solution was represented in Fig. 12.16. The potentiostatic measurements of produced films, were performed with the generation of an I-t curve. The film was colored at -1 V and bleached at $+1$ V, with each step being 60 s.

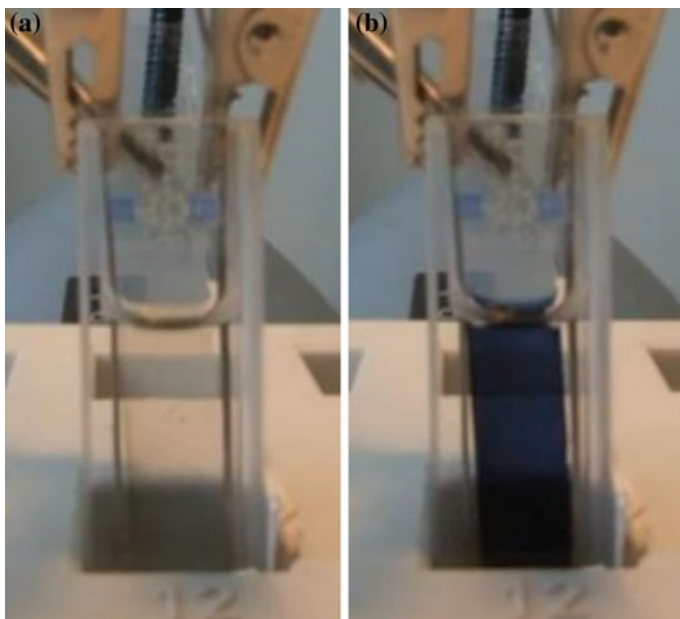
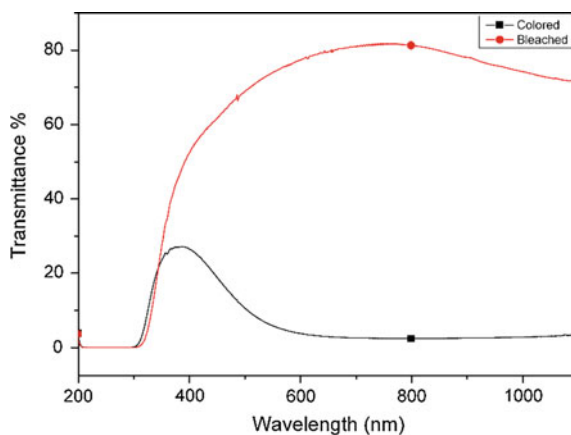


Fig. 12.16 Tungsten oxide electrodeposited film in nafion solution electrolytes **a** bleached state and **b** colored state

The in situ transmittance versus wavelength curves for tungsten oxide electrodeposited film by the WCl_6 precursor in nafion solution is shown in Fig. 12.17. The coloration and bleaching currents decay quickly for tungsten oxide electrodeposited film using the WCl_6 precursor. Additionally, in situ optical transmittance spectra of tungsten oxide electrodeposited film by the WCl_6 precursor was taken, the transmittance spectra for tungsten oxide electrodeposited film by the

Fig. 12.17 In-situ optical transmittance spectra of tungsten oxide electrodeposited film by the WCl_6 precursor in nafion liquid electrolyte



WCl₆ precursor exhibit high optical modulation. Transmittance values of the colored tungsten oxide film is 5.9 but bleached tungsten oxide has a transmittance of 74.0 at 550 nm. Transmittance values of the colored and bleached tungsten oxide films were obtained as 2.70 and 80.0 respectively at 680 nm.

12.3.2 Coloration Phenomena in WO₃

A variety of excitation processes such as thermal, optical, electrical, ionizing radiation, heating in vacuum, inert and reducing atmosphere, can be used for coloration of a material such as tungsten oxide. The spectrum of coloration is essentially similar in all cases. Coloration is structure-sensitive, while the absorption peak for amorphous WO₃ films is around 1.2 eV, for crystalline WO₃ films the absorption peak is around 0.7 eV.

Coloration occurs most efficiently in highly disordered amorphous WO_{3-x} films. Only sub-stoichiometric oxidized tungsten oxide crystalline films (WO₃) show coloration (Film formation occurs at high temperature under a reducing atmosphere such as vacuum, Ar, H₂, etc.). The photochromic effect (PC) occurs when films are irradiated with photon energies within the band gap (max at 3.6 eV) [28]. Ambient conditions can also effect optical and electrical coloration. Optical coloration efficiency increases in the presence of H₂O. Although there is no electrical coloration in vacuum, thermal, optical, electron and ion beam induced coloration is observed.

12.4 Conclusions

After four decades of extensive R&D on electrochromic (EC) properties of metal oxides, WO₃ still remains the leading EC-material. Major advances have been made in the fabrication and characterization (structural, optical, and electrical) of EC-materials and devices based on metal oxides. In this review, well-known preparation techniques of nanostructured tungsten oxide are briefly introduced, their similarities and differences according to the nanostructure of tungsten oxide produced are discussed. Two important factors, crystallinity of tungsten oxide film and the surface morphology play an important role in chromogenic properties of tungsten oxide nanomaterials.

References

1. X. Chi, Tungsten trioxide nanotubes with high sensitive and selective properties to acetone. *Sensors and Actuators B: Chemical* **194**, 33–37 (2014)
2. S. Novak, Effect of ingested tungsten oxide (WO_x) nanofibers on digestive gland tissue of porcellio scaber. *Am. Chem. Soc.* **47**, 11284–11292 (2013)

3. J. Lee, Simple fabrication of flexible electrodes with high metal-oxide content electrospun reduced tungsten oxidecarbon nanofibers for lithium ion battery applications. *Nanoscale* **6**, 10147–10155 (2014)
4. W. Li, Nanostructured WO₃ thin film as a new anode material for lithium-ion batteries. *Appl. Surf. Sci.* **256**, 2447–2452 (2010)
5. S. Cong, Y. Tian, Q. Li, Z. Zhao, F. Geng, Single-crystalline tungsten oxide quantum dots for fast pseudocapacitor and electrochromic applications. *Adv. Mater.* **26**, 4260–4267 (2014)
6. Y. Qin, Nanowire (nanorod) arrays-constructed tungsten oxide hierarchical structure and its unique NO₂-sensing performances. *J. Alloys Compd.* **615**, 616–623 (2014)
7. T.A. Nguyen, S. Park, J.B. Kim, T.K. Kim, G.H. Seong, J. Choo, Y.S. Kim, Polycrystalline tungsten oxide nanofibers for gas-sensing applications. *Sens. Actuators B: Chem.* **160**(1), 549–554 (2011)
8. C.V. Ramana, S. Utsunomiya, R.C. Ewing, C.M. Julien, U. Becker, Structural stability and phase transitions in WO₃ thin films. *J Phys. Chem. B* **110**, 10430–10435 (2006)
9. X. Chang, S. Shibin, Z. Li, X. Xu, Y. Qiu, Assembly of tungsten oxide nanobundles and their electrochromic properties. Elsevier, *Appl. Surf. Sci.* **257**, 5726–5730 (2011)
10. J. Sungpanich, Fabrication of WO₃ nanofibers by high voltage electrospinning. *Mater. Lett.* **65**, 3000–3004 (2011)
11. k Bange, Colouration of tungsten oxide films: a model for optically active coatings. Elsevier, *Solar Energy Mater. Solar Cells* **58**, 1–131 (1999)
12. E.K.H. Salje, S. Rehmann, F. Pobell, D. Morris, K.S. Knight, T. Herrmannsdorfer, M.T. Dove, Crystal structure and paramagnetic behaviour of e-WO_{3-x}. *J. Phys.: Condens. Matter* **9**, 6563–6577 (1997)
13. R. Chatten, A.V. Chadwick, A. Rougier, P.J.D. Lindan, The Oxygen Vacancy in Crystal Phases of WO₃. *J Phys. Chem. B* **109**, 3146–3156 (2005)
14. M.H. Kim, H.W. Choi, K.H. Kim, Properties of WO_{3-x} electrochromic thin film prepared by reactive sputtering with various post annealing temperatures. *IOP Jpn. Soc. Appl. Phys.* **52**, 11 (2013)
15. H. Miyazaki, Electrical and optical properties of amorphous and crystalline nondoped WO_{3-x} films prepared by reactive RF sputtering. *J. Appl. Phys.* **47**(9), 7295–7297 (2008)
16. E. Ozkan, S.H. Lee, P. Liu, C.E. Tracy, F.Z. Tepehan, J.R. Pitts, S.K. Deb, Electrochromic and optical properties of mesoporous tungsten oxide films. *J. Solid State Ionics* **149**, 139–146 (2002)
17. R. Vittal, H. Gomathi, K.J. Kim, Beneficial role of surfactants in electrochemistry and in the modification of electrodes. *Adv. Colloid Interface Sci.* **119**(1), 55–68 (2006)
18. W. Cheng, Y. Ju, P. Payamyar, D. Primc, J. Rao, C. Willa, M. Niederberger, Large-area alignment of tungsten oxide nanowires over flat and patterned substrates for room-temperature gas sensing. *Angew. Chem. Int. Ed.* **54**(1), 340–344 (2015)
19. X. Chi, C. Liu, L. Liu, Y. Li, Z. Wang, X. Bo, C. Su, Tungsten trioxide nanotubes with high sensitive and selective properties to acetone. *Sens. Actuators B: Chem.* **194**, 33–37 (2014)
20. P.V. Tong, N.D. Hoa, N.V. Duy, N.V. Hieu, Micro-wheels composed of self-assembled tungsten oxide nanorods for highly sensitive detection of low level toxic chlorine gas. *RSC Advances* **5**(32), 25204–25207 (2015)
21. M. Ahmadi, S. Sahoo, R. Younesi, A.P. Gaur, R.S. Katiyar, M.J. Guinel, WO₃ nano-ribbons: their phase transformation from tungstite (WO₃·H₂O) to tungsten oxide (WO₃). *J. Mater. Sci.* **49**(17), 5899–5909 (2014)
22. S. Salmaoui, F. Sediri, N. Gharbi, Characterization of h-WO₃ nanorods synthesized by hydrothermal process. *Polyhedron* **29**(7), 1771–1775 (2010)
23. G. Gu, B. Zheng, W.Q. Han, S. Roth, J. Liu, Tungsten oxide nanowires on tungsten substrates. *Nano Lett.* **2**(8), 849–851 (2002)
24. T.D. Manning, I.P. Parkin, M.E. Pemble, D. Sheel, D. Vernardou, Intelligent window coatings: atmospheric pressure chemical vapor deposition of tungsten-doped vanadium dioxide. *Chem. Mater.* **16**(4), 744–749 (2004)

25. H. Hassani, E. Marzbanrad, C. Zamani, B. Raissi, Effect of hydrothermal duration on synthesis of WO₃ nanorods. *J. Mater. Sci.: Mater. Electron.* **22**(9), 1264–1268 (2011)
26. R. Sahay, V. Thavasi, S. Ramakrishna, Design modifications in electrospinning setup for advanced applications. *J. Nanomater.* **17** (2011)
27. J. Díaz-Reyes, V. Dorantes-García, A. Pérez-Benítez, J.A. Balderas-López, Obtaining of films of tungsten trioxide (WO₃) by resistive heating of a tungsten filament. *Superficies Vacío* **21**(2), 12–17 (2008)
28. S.K. Deb, Opportunities and challenges in science and technology of WO₃ for electrochromic and related applications. *Sol. Energy Mater. Sol. Cells* **92**(2), 245–258 (2008)

Chapter 13

Electron Accumulation in InN Thin Films and Nanowires

L. Colakerol Arslan and K.E. Smith

Abstract An overview on the electron accumulation layer on InN thin film and nanowire surfaces is provided. The interactions between the valence and conduction bands due to the narrow band gap and high electron density at the surface of these materials have a big influence on the electronic structure and the device performance of these materials. We first review the current understanding on the electron accumulation on InN thin films, pointing out the role of defects and dislocations on the unintentional n-type conductivity. Then we carry out detailed investigation on tuning the surface charge properties of InN nanowires depending on the growth process.

13.1 Introduction

The narrow band gap semiconductors such as InAs, InSb, and InN are subject of intense study because of their combined electrical and optical functionalities. Their distinctive material characteristics such as high electron mobility, small effective mass and strong spin-orbit coupling provide excellent charge transport properties for applications in ultrahigh-speed nanophotonic devices. Among these materials, InN has attracted huge attention recently because of its near-surface properties. The low dissociation temperature of InN and the extremely high equilibrium vapor pressure of nitrogen prevented the preparation of high quality single crystals, leaving the optical, electrical and physical properties of this material difficult to

L. Colakerol Arslan (✉)

Department of Physics, Gebze Technical University, 41400 Kocaeli, Turkey

e-mail: lcolakerol@gtu.edu.tr

K.E. Smith

Department of Physics, Boston University, 02215 Boston, MA, USA

e-mail: ksmith@bu.edu

K.E. Smith

School of Chemical Sciences, The MacDiarmid Institute for Advanced Materials and Nanotechnology, The University of Auckland, 1142 Auckland, New Zealand

determine. Early InN samples, prepared by reactive RF sputtering of indium targets with nitrogen [1], have a band gap value of 1.9 eV due to the presence of oxygen in the films during growth [2]. Recently, the energy of the fundamental band gap has undergone a remarkable revision from the previously accepted value of approximately 1.9 eV to the much lower value of approximately 0.65 eV as a consequence of the development of epitaxial growth methods for high quality crystalline thin films of InN [3–5]. The band gap of the $\text{In}_x\text{Ga}_{1-x}\text{N}$ ternary alloy system now spans the near infrared (0.65 eV, $x = 1$) to the UV (3.4 eV, $x = 0$), enabling the entire optical window to be encompassed by a single material system [6]. This makes $\text{In}_x\text{Ga}_{1-x}\text{N}$ potentially suitable for applications such as high-brightness white light emitting diodes and high-efficiency solar cells. In addition, mobilities over $3000 \text{ cm}^2/\text{Vs}$ have been measured at room temperature [7], leading to high performance transistors operating at high frequencies up to the several THz range. The use of InN based optoelectronic devices offers the potential of an environmental-friendly red emitter with no toxic element, which may replace GaAs based devices. In order to fully realize the potential of InN, a broader investigation of the surface and interface properties was performed, since these will have a major influence on the design of low dimensional devices.

InN is unusual in that there is strong experimental evidence for the existence of an intrinsic electron accumulation layer near the surface of the thin films, where the density of electrons is over an order of magnitude larger than in the bulk. The existence of a large electron accumulation layer at all surface planes and the possibility of spin-splitting may also play an important role in electron spin transport and spintronics and extends InN's possible applications [8]. This electron accumulation could also potentially be used for gas sensor applications and for the next generation of THz radiation generators and sensors [9]. On the other hand, the high electron density makes it difficult to achieve p-type doping in InN. Therefore, understanding of the nature of this electron accumulation and surface electronic properties is of a great importance in the design and the analysis of high performance nitride-based electronic devices.

Adding up these entire qualities one can conclude that InN represents one of the most versatile semiconductor material systems. However, due to the lack of homoepitaxial substrates for nitride epitaxy the crystalline quality of planar films is not perfect. Despite high defect densities, nitride devices are extraordinarily performing. Recently, InN nanostructures in the forms of nanowires (NWs) have also received great attention due to the high surface-to-volume ratio and enhanced electron accumulation at the surface. NWs, which can be grown as single nano-crystals, show fewer structural defects than planar films and therefore, are expected to further improve the device quality. Especially, when the high surface electron density is combined with high mobility, InN NWs become attractive materials for the production of high electron mobility transistors operating at high frequencies. In addition, InN NWs can also be grown on several different substrates, including silicon, therefore the possibility of combining these highly lattice mismatched materials provides easy integration of these nanowires with the established Si microelectronics technology.

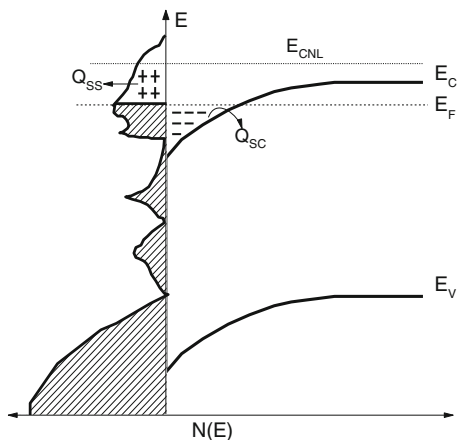
In this chapter, we have focused on the study of near surface electron accumulation in InN thin films and NWs. We begin by exploring the various mechanisms behind the origin of electron accumulation at InN surfaces. The accumulation layer is discussed in terms of the bulk Fermi level (E_F) lying below the pinned surface E_F , with a confining potential formed normal to surface due to the downward band bending facilitated by donor type surface states or impurities. Then the effect of these charge accumulation on the electronic properties of these materials was investigated. The effect of growth properties and reduced dimensions on the E_F pinning mechanism on InN nanowire surfaces is also discussed. Mg-doped InN NWs and the possibility to achieve p-type conductivity in InN NWs are presented.

13.2 Fermi Level Pinning

The pinning of the Fermi level (E_F) in semiconductors is observed for a wide class of physical phenomena, including interface formation, surface processes, chemical impurity doping, and high-energy radiation effects. For free surfaces of semiconductors, E_F pinning is driven by sufficiently high density of surface states of both donor and acceptor character [10, 11]. These surface states arise from the termination (or reconstruction) of the bulk crystal at the surface. They can also be caused by impurities or defects. If there surface states exist in the gap, they pin the Fermi level upward or downward depending on their position relative to charge neutrality level (CNL) [12]. The charge neutrality at the surface is satisfied when the total charge due to surface states, Q_{ss} , is compensated by an equal but opposite charge near the surface region called space charge, Q_{sc} . In the presence of ionized surface states at semiconductor free surfaces, charge neutrality is ensured by the bending of electronic bands due to the Fermi level being pinned at the surface. If the Fermi level is located above the CNL, the surface states are negatively charged and a depletion layer is created on the surface. If the electrons are filled to a point below the CNL, the surface states are positively charged and this results in an accumulation layer near the surface. The CNL allows a unique correspondence to be made between the band bending at the semiconductor surface and the population of surface states. For an intrinsic semiconductor, also called an undoped semiconductor, with homogeneous surfaces, the Fermi level should coincide with the CNL of the surface states.

Narrow band gap III–V semiconductors usually show an accumulation layer due to pinning of the Fermi level above the conduction band edge. This electron accumulation in these materials can be explained by the extremely low conduction band minimum at the Γ point compared with at other points in k-space. Therefore, the CNL is located above the CBM, leading to positive surface charge, Q_{ss} , as seen in Fig. 13.1. To compensate this surface charge, a negative space charge builds up on the surface and Fermi level is pinned below the CNL, resulting in an electron accumulation layer. The observed electron accumulation at the surface of these materials is due to the presence of positively charged donor-type surface states.

Fig. 13.1 Illustration of a surface accumulation layer. E_F is located below the charge neutrality level, the surface is positively charged forming an accumulation layer



Similar change of surface potential and the corresponding charge accumulation layer formation region occurs near surfaces of InN NWs. In this case, the nanowire can be modeled by two distinct regions along the lateral dimension, including the space charge region near the surface and the neutral region in the bulk.

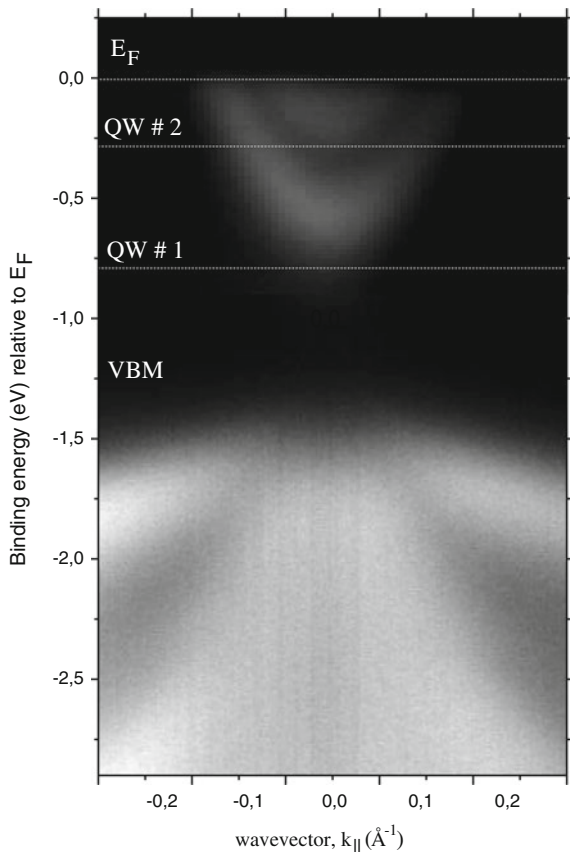
13.3 Surface Electron Accumulation in InN Thin Films

There have been several reports regarding surface electron accumulation on InN thin films. Evidence for the accumulation layer on InN thin films comes from high resolution electron energy loss spectroscopy (HREELS) studies of clean InN surfaces [13, 14], which revealed a conduction band plasmon whose energy increased as the incident electron energy was reduced, indicating a higher electron density within 80 Å of the surface than in the bulk. Further evidence comes from sheet carrier density measurements as a function of film thickness, capacitance-voltage measurements, a photoemission study of defective InN coated with Ti [15, 16]. Recent tunneling spectroscopy experiments add further credence to the existence of this layer, with features associated with electronic sub-bands being observed, indicating that the electrons are in quantized states [17].

The most direct evidence of electron accumulation layer on InN is obtained from the high resolution angle resolved photoelectron spectroscopy (ARPES) spectra showing the energy region between E_F and the top of the valence band, as shown in Fig. 13.2. ARPES is one of the most direct methods to determine the energy dispersion of electronic states in solids. This technique measures the energy and the angular distribution of electrons emitted from the surface of a material when it is illuminated with monochromatic photons of sufficient energy. ARPES reveals two discrete states, centered around the Γ -point. The states were characterized by varying the photon energy, and thus probing the band dispersion along k_z (surface normal)

Fig. 13.2 ARPES

photocurrent intensity map of states within 3 eV of E_F . $h\nu = 69$ eV and sample temperature was 177 K. Sample was annealed to 300 °C in UHV for 30 min. The false color intensity reflects the photocurrent, with lighter intensity indicating higher current



and angle of emission. The negligible dispersion along k_z and narrow range of emission angles over which these features are observed, indicating that the emitting states are highly localized in k space. The measured separation between the VBM and E_F is 1.2 eV for this surface, which is higher than the band gap of InN, confirming that E_F is above the conduction band minimum (CBM). The energy difference between E_F and the CBM for this surface is 0.65 eV. Therefore, these states are the quantum well states in the accumulation layer in InN due to the quantization of electrons trapped in the conduction band minimum. These states, although derived from the conduction band, arise from the existence of a potential well perpendicular to the film surface. Downward band bending forms a one-dimensional (1D) potential well and the resulting two dimensional (2D) electron gas is quantized along the surface normal. This band bending is the result of donor-type surface states or N vacancies pinning E_F above the CBM [15, 18]. Similar behavior is also observed from other In containing III–V compounds such as InAs ($E_F = 0.36$ eV) [19], InSb ($E_F = 0.18$ eV) [20], and InP ($E_F = 1.35$ eV) [21].

The existence of 2D electron accumulation layer due to downward band bending and the small value of the energy gap in InN lead to strong interaction of quantized electron states and valence band states. Thus, the dispersion of the top of the valence band in InN within the quantum well potential associated with the electron accumulation layer is affected. Although InN is well known to be a direct gap semiconductor in the bulk, intrinsic quantum well potential produces complicated hole dispersion curves and a minimum in the dispersion of the top of the valence band is measured at the Γ -point, as can be seen in Fig. 13.2. Although two dispersing heavy and light hole valence bands locate around the top of the valence band, the intensity of the heavy hole valence band vanishes near zone center, resulting in a camel-back dispersion of the top of the valence band [22].

There is a significant body of experimental evidence that InN is a direct gap semiconductor, including a recent study of the bulk electronic structure using resonant x-ray emission spectroscopy and state-of-the-art hybrid density functional theory [23]. Clear intermixing between the heavy-hole and light-hole valence bands in the intrinsic quantum well potential is observed associated with the near-surface electron accumulation layer; results in an inverted band structure, with the valence band maximum lying away from the Brillouin zone center. Such structures have been previously deduced from earlier transport and tunneling experiments from engineered single $\text{HgTe}/\text{Hg}_{1-x}\text{Cd}_x\text{Te}$ [24] and GaAs/AlAs quantum wells [25]. Theoretical modeling of electronic band structure of InN demonstrates that InN films under biaxial strain presents similar inverted valence band structure [26]. Tensile strain occurs during the early stages of growth on the InN films grown on GaN (0001) buffer layer due to elastic deformation of the crystal to fill the gaps between InN islands during coalescence [27, 28]. Chen et al. showed that InN films grown on GaN buffer layer form InN islands although InN films grown on InN buffer layer after nitridation does not form island structures [29]. Since the sample showing an inverted band structure was grown on InN buffer layer, don't have such strain to cause valence band mixing. Therefore intermixing between the heavy-hole and light-hole valence bands in the intrinsic quantum well potential is observed associated with the near-surface electron accumulation layer; results in an inverted band structure, with the valence band maximum lying away from the Brillouin zone center.

The origin of the surface states responsible for the Fermi level pinning mechanism at InN surfaces is associated with donor-like impurities or crystal defects. Due to nitrogen vacancies, dislocations, and incorporation of oxygen or hydrogen in the bulk during crystal growth; even high quality InN layers grown by plasma-assisted molecular beam epitaxy still show bulk free electron density of 10^{17} cm^{-3} [30–32]. According to models derived on the basis of single field electrical Hall effect measurements, the positively charged N vacancies associated with dislocations are considered to be the origin of unintentional n-type conductivity [33]. Moreover, surface reconstruction, surface adsorbates and formation of In overlayers or droplets are also identified as factors that cause donor type surface states [34, 35]. Models derived on the basis of single field electrical Hall effect measurements favor positively charged N vacancies (V_N^+) as the major origin of n-type conductivity.

In fact, ARPES spectra obtained from InN that is treated with two cycles of 500 eV Ar^+ ion bombardment and annealing in UHV to 300 °C show that modifying the InN surface leads to an increase in charge density in the conduction band [36]. As can be seen in Fig. 13.3a, the subband energy levels are localized at 0.511 and 0.797 eV below E_F after sputtering and annealing, while they are located at 0.233 and 0.662 below E_F for clean InN surface. The increase in the energy levels is directly related to increase in the charge density at the bottom of the conduction band. It is likely that the bombardment of InN surface with inert gas ion leads to N vacancies since this method may leave binary or ternary nitrides in a non-stoichiometric state [37, 38].

A series of momentum distribution curves (MDCs) extracted from Fig. 13.3a is presented in Fig. 13.3b to emphasize the dispersion of the quantized states in the conduction band. Dispersion of the bands is fitted assuming both parabolic (dashed gray) and non-parabolic bands (solid black). The non-parabolic curve is fitted to the outer band which is more resolved and has larger non-parabolicity effect because of its higher energy and larger occupation probability. It can be seen clearly that the dispersion of the states appreciably deviate from the parabolic dispersions. Rather, it is linear away from the subband minima and the non-parabolic dispersion follows the relation given by Kane's two band k.p model for narrow band gap semiconductors [39]. Note that non-parabolic single-band tight binding approach has also been applied to model surface electron accumulation at InN surfaces [40]. In narrow

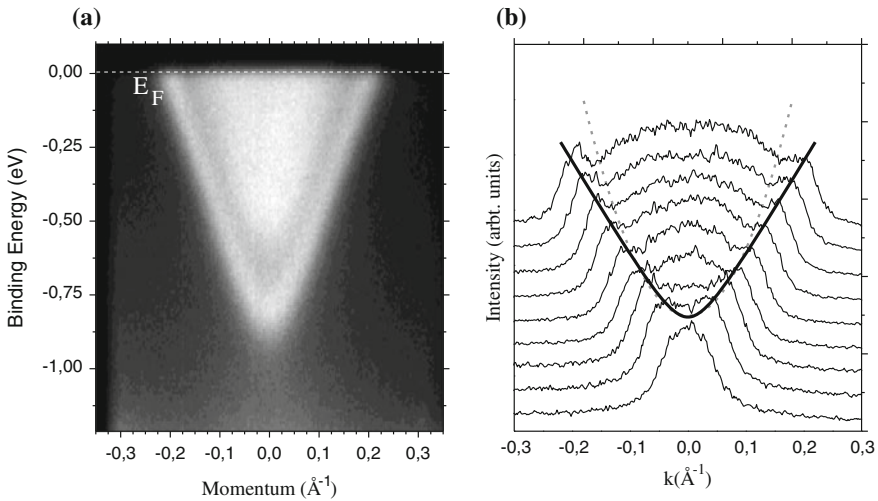


Fig. 13.3 **a** ARPES photocurrent intensity map of states within 1.2 eV of E_F . $h\nu = 70$ eV, and sample temperature was 60 K. Sample was prepared by two cycles of 500 eV Ar^+ ion bombardment and annealing in UHV to 300 °C. The momentum direction is along ΓSM in the surface plane. **b** Momentum distribution curve derived dispersion of the conduction band subbands extracted from the spectra on the left. The dispersion of the states are fitted by parabolic (dashed gray) and non-parabolic 2D (solid black) based on Kane's model [39] with effective masses respectively $0.16m_0$ and $0.1m_0$

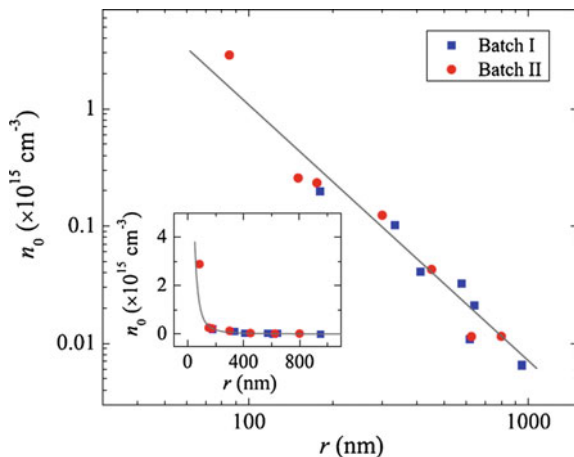


Fig. 13.4 The derived free electron concentration n_0 as a function of the nanowire radius r in a logarithmic scale. The line is a linear fit that gives a slope $k = -2.3 \pm 2$. The *inset* shows the plot in a linear scale, and the curve is given by $n_0 \propto 1/r^{2.3}$. Reprinted with permission from [50]. Copyright 2013, AIP Publishing LLC

band gap semiconductors, strong k.p interaction between the conduction band and the valence bands results in strong non-parabolicity in the conduction band [41]. Similar non-parabolic dispersion of InN conduction band is also observed by other groups by measuring the relation between the effective mass and carrier concentration [42].

Much debate is ongoing about the origin of the Fermi level pinning whether it is an intrinsic property or it occurs under certain reconstruction conditions. There are recent reports on the absence of surface Fermi level pinning on non-polar surfaces, which were cleaved under ultra-high vacuum conditions [43]. Theoretical calculations also suggest that surface charge accumulation would not be expected on non-polar surfaces [44]. However, the universal nature of the electron accumulation is found at both polar and nonpolar InN surfaces under ambient conditions due to contamination or adatom adsorption. Various growth conditions or surface treatment methods may affect the density and the microscopic nature of charge accumulation. However, all InN surfaces exhibit a tendency toward electron accumulation, because the bulk Fermi level is located below CNL level. This is confirmed by King et al. who showed that neither the film polarity nor the growth parameters have any effect on the Fermi level pinning mechanism [45]. This is in contrast to the interpretation of recent Hall effect measurements, which suggested different surface sheet densities for In- and N-polarity samples and the samples grown at N-rich and In-rich conditions. However, theoretical and experimental investigations show that the variation in the sheet density of the surface electron accumulation with buffer layer, polarity, or growth conditions is related with the dislocation densities at the interface [46–48].

Although this intrinsic band bending and electron accumulation may affect the electronic properties near the surface region, they have relatively low influence on the bulk electronic properties in semiconductor thin films. In semiconductor nanowires, on the other hand, the dimensions of the NW are often on the order of the Debye screening length, so charge accumulation layer has a larger impact on the electronic properties due to high surface-to-volume ratio.

13.4 Surface Charge Accumulation on InN NWs

InN nanowires are highly interesting as they can exhibit considerable confinement effects which provide them significantly different properties than InN thin films. Although it is quite challenging to grow InN thin films with high crystalline quality due to the lack of homoepitaxial substrates, NWs can be form on various substrates with lower defect densities than planar films. Numerous studies are focused on the investigation of InN NW structure devices with improved quality and the possibility of integration with silicon technology [49–51]. Moreover, surface space charge layer on polar and non-polar surfaces of InN NWs renders them ideal candidates for use in nanoscale sensor devices and infrared light-emitting diodes and lasers [52].

The thermodynamically stable phase of InN NWs is hexagonal wurtzite structure with two possible orientations depending on the growth method. The MBE grown nanowires are oriented along the [0001] polar direction, with their sidewalls being nonpolar *m*-planes [49]; while the ones deposited by chemical vapor deposition using the vapor-liquid-solid (VLS) growth mechanism are oriented along the [11–20] direction, and the sidewalls are formed by the polar {0001} facets [53]. The strong ionicity of the metal-nitrogen bond and absence of inversion symmetry leads to strong macroscopic polarization along the [0001] direction in the III-nitrides at zero strain [54, 55]. If an external electric field is applied to the III–V nitride lattice, crystal lattice will induce piezoelectric polarization to accommodate the stress. The overall polarization strength of crystal affects the free electron motion inside the crystal. Due to extraordinary surface-to-volume ratio and absence of interface affects, polarity of the surface has a dramatic effect on surface Fermi level pinning and therefore optical and electrical properties of InN nanowires, although the surface Fermi level pinning was shown to be the same for *a*-plane and for both polarities of *c*-plane InN.

The lateral surfaces of most of the InN NWs exhibit high density of accumulated electrons due to the presence of positively charged donor-type defects. An accumulation layer is formed due to Fermi level pinning on the sidewalls of NW surface, leaving a region with free-electron concentration in the middle of the wire. Extend of this region depends on the thickness of the wire. Several efforts have been made to characterize and to control E_F pinning and surface charge properties of InN nanowires. To prove the presence of a surface accumulation layer, electrical measurements were performed on InN nanowires and the resistance was displayed as a function of wire radius. It is expected that the conductance should scale with

the cross sectional area of the wire (πr_{nw}^2) as the current flows through the bulk of a wire, while the conductance should scale with πr_{nw} as the current flows through the surface of a wire. The data presented by Richter et al. showed an approximately linear relation between the conductance and diameter for wires with a diameter below 100 nm, suggesting that major contribution to conduction through InN nanowires is coming from the surface accumulation layer [56]. A decreasing resistivity with decreasing wire diameter is also obtained by others [57]. Another supporting evidence for the existence of surface electron accumulation on InN NW surface comes from the analysis of the slope of the conductance as a function of temperature [56]. The resistance of InN NW decreases linearly with decreasing temperature. This metallic like behavior indicates that the main conductance contribution arises from the surface electron accumulation layer. The pinning of the Fermi level 0.6–0.7 eV above the conduction band minimum at the surface of the nanowires is determined by polymer electrolyte gating and three-dimensional electrostatic modeling of charge distribution [58].

The existence of an electron accumulation layer on InN surface is also proven by photoluminescence (PL) spectroscopy which has been a key technique to study the optical processes in semiconductors with particular sensitivity to point defects. The shape and peak position of the PL spectra of InN NWs grown under In rich conditions with electron densities in the range of 1×10^{18} to $6 \times 10^{18} \text{ cm}^{-3}$ exhibit a weak dependence on the laser intensity which can be characterized by the presence of a weak band filling effect due to photogenerated charge carriers [59]. The reason for the weak excitation power dependence is a large broadening of the PL peak. Moreover, a high energy tail exists in the PL spectrum and it extends to higher energies with increasing temperature. The authors have correlated the observed features in the PL spectra with fluctuations of the Fermi level along the wire, relative to the conduction band edge and modelled based on an accumulation layer at the nanowire surface.

Raman spectroscopy is also used as an effective method to investigate surface properties of semiconductors as it is a nondestructive and no contact characterization tool. In most cases, Raman spectra taken perpendicular to c-axis of InN wires display a strong uncoupled $E_1(\text{LO})$ mode near the frequency of the $A_1(\text{LO})$ phonon in addition to bulk phonon modes of $E_2^H(\text{LO})$ and $A_1(\text{LO})$. The physical origin of this LO mode is thought to be the scattering of photons associated with the charge density fluctuations. The change in the $E_1(\text{LO})$ frequency with excitation wavelength indicates the changes in the free carrier concentration through the wire thickness due to the different probing depths. Furthermore, enhancement in the relative intensity of $E_1(\text{LO})$ mode is observed for the samples with morphologies benefiting the scattering through the lateral walls [60]. Therefore this mode is attributed to free electron accumulation on NW surfaces. A similar LO mode is also observed in Raman spectrum of a single InN NW independently of the scattering geometry, indicating the presence of high electron concentration at the wire surfaces [61].

Detailed investigation on the origin of an electron accumulation layer at InN NW surface is performed by using an electrical nanoprobe technique in a scanning

electron microscope environment [50]. It is observed that the derived free electron density increases as $n \sim r^{-2.3}$ shown in Fig. 13.4. The existence of high carrier concentration in small-radii NWs is explained by the presence of surface states. Considering the fact that the diameter of nanowires grown at low temperatures is smaller than that of the ones grown at higher temperatures, this is an indication of that nanowires with smaller radii have relatively higher carrier concentration due to surface defects. Recent study on the surface charge properties of InN nanowires has also revealed that the surface defects are mainly responsible for large carrier concentration and corresponding electron accumulation on InN surfaces [62]. In fact, the surface charge concentration can be tuned by varying the nanowire morphology through In seeding mediated growth approach [62].

InN NWs grown with the use of an in situ deposited In seeding layer under nitrogen rich conditions have carrier concentrations of $\sim 1 \times 10^{15-16} \text{ cm}^{-3}$. Moreover, the optical excitation power and temperature has a strong influence on their PL [63, 64]. Such behavior is a clear indication of low carrier density and absence or a negligible level of surface electron accumulation. The authors recorded an emission peak at 0.64 eV (at 300 K) due to band-to-band recombination. According to ARPES results, the Fermi level position of undoped InN NWs is located 0.5 eV above the VBM, which would imply that the near-surface E_F for InN NWs lies well below the CBM, as shown in Fig. 13.5a. In contrast, the PL spectra

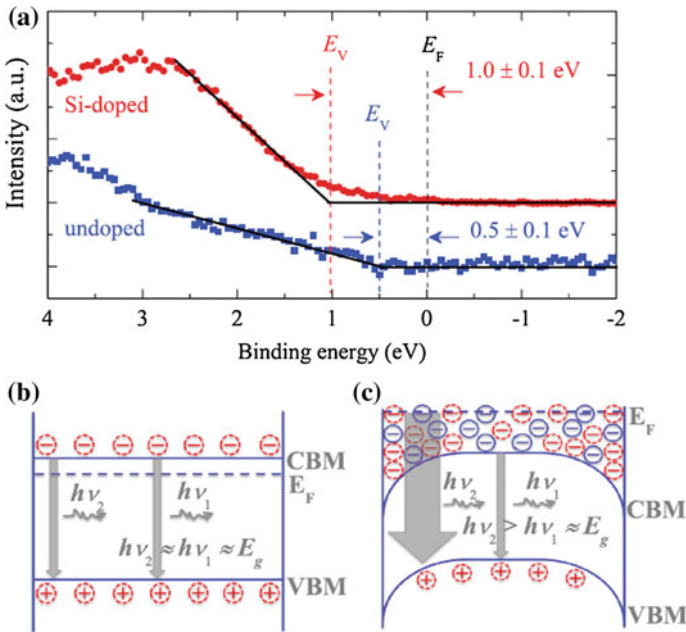


Fig. 13.5 **a** X-ray photoelectron spectroscopy of intrinsic and moderately Si-doped InN nanowires. **b** and **c** show the carrier distributions and recombination processes in intrinsic and InN: Si nanowires, respectively. The blue circles represent residual free charge carriers, and the dotted red circles represent photogenerated charge carriers. Reprinted with permission from [63]. Copyright (2012) by the American Physical Society

of Si doped InN NWs show two major peaks at 0.65 and ~ 0.75 eV. While the low energy emission is associated with band-to-band carrier recombination in the relatively low-doped bulk region, the high energy peak is associated with the emission of Mahan exciton in the high-doped nanowire near-surface region. Such behavior is confirmed by the pinning of the $E_F \sim 0.35$ eV above the CBM. A schematic band diagrams to illustrate the carrier recombination processes are shown in Fig. 13.5b, c for intrinsic and Si-doped InN NWs, respectively. Photogenerated electron-hole pairs are distributed nearly uniformly across the cross section of undoped InN NWs. For InN:Si nanowires, however, a surface charge density results in a downward band bending. Then photoexcited electrons will migrate to the surface along the potential slope produced by band-bending, leading PL emissions from bulk states and from surface trap states.

Werner et al. Showed that the surfaces of InN NWs are covered by a thin layer of In_2O_3 according to the results of XPS. The shell conductivity was ascribed to the accumulation layer located at the interface between InN and In_2O_3 layers formed during exposure to air [57].

13.5 Control of Surface Electron Accumulation in InN Nanowires

The band bending can be detrimental for technological applications because it precludes external control of the electronic properties and the realization of p-type conductivity in InN although it is essential for the fabrication of p-n junctions and building the basis of all electronic devices. Growth process and the quality of NWs have a major influence on the band bending mechanism.

The 2D electron accumulation level on the surfaces of intrinsically doped InN nanowires can be controlled by varying the epitaxial growth process. In this section, we will discuss the effects of growth parameters on the 2D electron accumulation layer on InN NW surfaces. Similar to InN thin films, early InN NWs prepared by solvo-thermal method, catalytic chemical vapor deposition, and vapor-liquid-solid were poorly crystalline and the band gap value was about 1.9 eV [65–67]. Recent improvements in its molecular beam epitaxial growth process have resulted in superior quality InN nanowire structures with a band gap of ~ 0.7 eV on various substrates becoming available [62, 68, 69]. However, most of the MBE grown InN NWs exhibit uncontrolled tapered morphologies and poor crystal quality because of the low decomposition temperature of InN, high vapor pressure of nitrogen over InN and the lack of lattice matched substrates. Mi and coworkers synthesized InN NWs directly on silicon substrate with the utilization of in-situ deposited In seeding layer [69]. The use of In seeding layer promotes the nucleation of InN and provides the growth of non-tapered InN NWs with a well-defined hexagonal structure along [0001] direction. Figure 13.6 shows the SEM images of InN nanowires grown at different temperatures. Optimum substrate temperature of 480 °C results in NWs

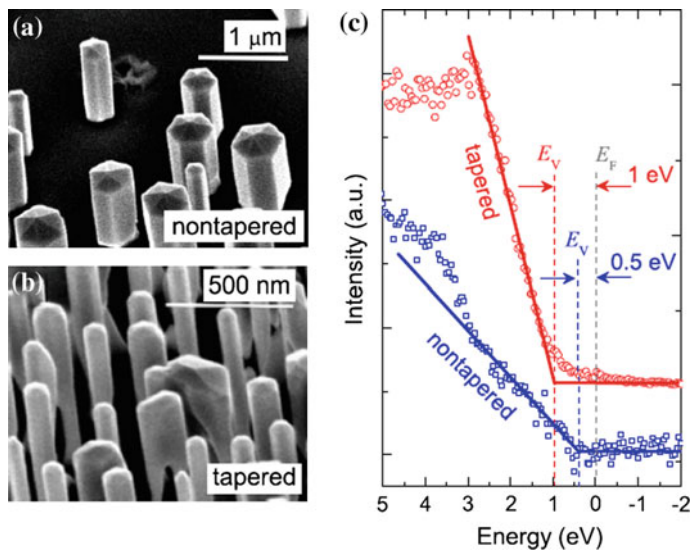


Fig. 13.6 The role of surface defects on InN nanowire quality and the near-surface Fermi-level. **a** and **b** SEM images of nontapered and tapered InN nanowires, respectively. **c** The XPS spectra for nontapered and tapered InN nanowires measured from the sidewalls. Reprinted with permission from [70]. Copyright (2014) by John Wiley and Sons, Inc.

with a uniform diameter as well as smooth hexagonal facets. Lower growth temperatures result in a strong tapering of the NW. The InN NWs synthesized at optimal substrate temperature have a rather low carrier concentration of low-to-mid 10^{13} cm^{-3} and high electron mobility of $12,000 \text{ cm}^2/\text{Vs}$ [50]. The valence band spectrum taken from the tapered NWs sidewalls shows that E_F is located about $\sim 0.15 \text{ eV}$ below the CBM (see Fig. 13.6c) The E_F of non-tapered InN nanowires, on the other hand, is positioned $\sim 0.5 \text{ eV}$ above the VBM, which implies that the surface Fermi level is not pinned above the conduction band. This was a major step toward realization of p-type doping.

Another promising evidence for p-type conductivity in InN nanowires comes from transport measurements. Due to the large surface to volume ratio in NWs, the surface has an even higher impact on electrical transport properties. The electrical properties (I-V) of as fabricated single InN nanowire is characterized a power-law dependence of the type $I \propto V^\beta$, with good ohmic behavior ($\beta \sim 1$) at lower biases, higher order response at larger biases, $\beta > 2$, associated with the trap-filled space charge limited conductivity due to the presence of charge traps at high bias. It is well known that when a dc voltage is applied to insulators or nearly insulating materials, space charge is formed. Therefore, such non-linear behavior implies that InN NWs have extremely low background concentration, i.e., the E_F of those non-tapered InN nanowires is located below the conduction band edge.

Following the observation of the lack of surface electron accumulation on cleaved InN surfaces and theoretical calculations on the surface states of InN which

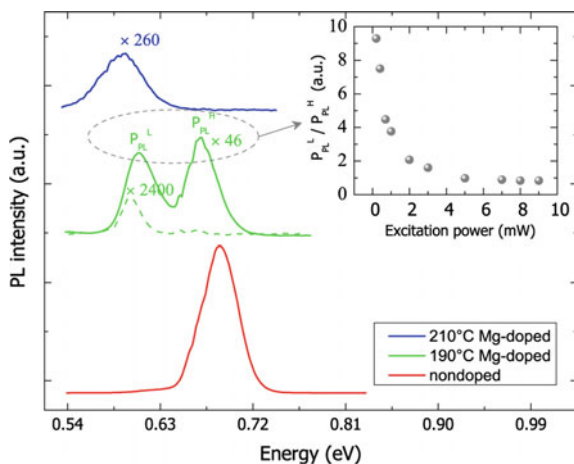


Fig. 13.7 The PL spectra of Mg-doped InN nanowires measured at 7 K under a 9 mW optical excitation. The PL spectrum of nondoped InN nanowires measured at 20 K under the same power is also shown for a comparison. The spectra were normalized by the PL peak intensity of the nondoped InN nanowires and were shifted for display purpose. The *dotted green curve* shows the PL spectrum measured under an excitation of 200 μ W from the Mg-doped sample with a Mg cell temperature of 190 $^{\circ}$ C; and the *inset* shows the PL peak intensity ratio of P_{PL}^L over P_{PL}^H as a function of the excitation power. Reprinted with permission from [78]. Copyright (2013) American Chemical Society

indicate that the Fermi-level can be unpinned on nonpolar planes, there has been intense study in the production of p-type doping in InN [70–76]. Since Mg has been proven to be an effective p-type dopant in GaN [77], Mg is also considered to be the most suitable candidate for obtaining p-type InN. However most studies did not confirm the presence of hole carriers in Mg doped InN because the surface potential originated by the depletion layer prevents the electric field from penetrating into the bulk of the film transport measurements are mostly dominated by surface charge layer and could not be used to determine bulk carrier density. Nevertheless, some evidence of the p-type conductivity in Mg doped InN NWs have been provided by PL and ARPES measurements [78]. Figure 13.7 shows the PL spectra of Mg-doped InN nanowires exhibiting non-tapered surface morphology measured at 7 K along with the PL spectrum of nondoped InN nanowire measured at 20 K. While the PL spectrum of the nondoped InN NWs is dominated by a single peak at 0.69 eV, PL spectrum of Mg doped InN NWs contains two emission peaks at energies $E_{PL}^L \sim 0.61$ eV and $E_{PL}^H \sim 0.67$ eV. With the increase of Mg doping, only PL^L peak can be observed and it is redshifted due to band gap renormalization. Enhanced defect incorporation leads to disappearance of high energy PL^H peak. To further investigate the origin of these two emission peaks, excitation power dependence of PL spectra was examined, as shown in the inset of Fig. 13.7. At low excitation power, low energy peak is dominant. Since holes at the Mg-acceptor

energy level and electrons in conduction band level participate in radiation recombination, this lower energy peak is ascribed to Mg-dopant associated acceptor energy level. The intensity of high energy peak, on the other hand, increases with the excitation power indicating that it is originating from valence to conduction band transitions in InN NWs.

ARPES measurements were performed at a grazing incidence on the samples with various Mg doping level to determine the position of the Fermi level at the surface as a function of bulk carrier concentration [78]. No states associated with the surface electron accumulation layer observed around the leading edges of the valence band photoemission spectra for high Mg-doped InN NWs. For moderate Mg concentrations ($T_{\text{Mg}} < 220^\circ$), the Fermi level is located ~ 0.4 eV above the VBM independent of the doping level. The surface Fermi level positions for high Mg concentrations ($T_{\text{MG}} > 230^\circ$) monotonically decrease with increasing Mg density and reaches to the level of 0.1 eV above the VBM at highest concentration, indicating the p-type doping of InN NWs. Because NW growth temperature is much higher than the Mg effusion cell temperature, it is quite challenging to incorporate Mg atoms around the surface region. However, at elevated Mg doping levels, enhanced acceptor incorporation in the near-surface region can balance to a certain extent the large surface desorption of Mg atoms during the epitaxial growth process.

13.6 Conclusions

This chapter is a review of experimental progress that has been achieved in understanding and controlling the surface electron accumulation in InN thin films and NWs. The existence of the intrinsic electron accumulation layer is explained in terms of a particularly low Γ -point conduction band minimum in wurtzite InN. As a result, the surface Fermi level is pinned high in the conduction band due to excess charge at the surface created by impurities, dislocations or donor-type surface states. Although nearly all InN thin films exhibit a strong charge accumulation, the degree of charge accumulation in InN NWs depends on the growth mechanism and specific surface conditions. In fact, the absence of an accumulation layer on highly crystalline InN NWs raises the question if the Fermi level pinning is an intrinsic property of InN or if it only occurs due to growth related defects or surface reconstruction.

Acknowledgments Thank Asst. Prof. L.F.J. Piper and Prof. T.D. Moustakas for their contributions to this work.

References

1. A.G. Bhuiyan, A. Hashimoto, A. Yamamoto, J. Appl. Phys. **94**, 2779–2808 (2003)
2. T.L. Tansley, C.P. Foley, J. Appl. Phys. **59**, 3241–3244 (1986)
3. V.Y. Davydov, A.A. Klochikhin, R.P. Seisyan, V.V. Emtsev, S.V. Ivanov, F. Bechstedt, J. Furthmuller, H. Harima, V. Mudryi, J. Aderhold, O. Semchinova, J. Graul, Phys. Status Solidi B **229**, R1–R3 (2002)
4. J. Wu, W. Walukiewicz, K.M. Yu, J.W. Ager, E.E. Haller, H. Lu, W.J. Schaff, Y. Saito, Y. Nanishi, Appl. Phys. Lett. **80**, 3967–3969 (2002)
5. T. Matsuoka, H. Okamoto, M. Nakao, H. Harima, E. Kurimoto, Appl. Phys. Lett. **81**, 1246–1248 (2002)
6. J. Wu, W. Walukiewicz, K.M. Yu, J.W. Ager, E.E. Haller, H. Lu, W.J. Schaff, Appl. Phys. Lett. **80**, 4741–4743 (2002)
7. X.Q. Wang, S.T. Liu, N. Ma, L. Feng, G. Chen, F.J. Xu, N. Tang, S. Huang, K.J. Chen, S.Q. Zhou, B. Shen, Appl. Phys. Express **5**, 015502 (2012)
8. C.M. Yin, H.T. Yuan, X.Q. Wang, S.T. Liu, S. Zhang, N. Tang, F.J. Xu, Z.Y. Chen, H. Shimotani, Y. Iwasa, Y.H. Chen, W.K. Ge, B. Shen, Nano Lett. **13**, 2024–2029 (2013)
9. H. Ahn, Y.J. Yeh, Y.L. Hong, S. Gwo, Appl. Phys. Lett. **95**, 232104 (2009)
10. H. Lüth, *Surfaces and Interfaces of Solids* (Springer, Berlin, 1993)
11. P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors* (Springer-Verlag, Berlin, 1996)
12. R.T. Tung, Mat Sci Eng R **35**, 1–138 (2001)
13. I. Mahboob, T.D. Veal, C.F. McConville, H. Lu, W.J. Schaff, Phys. Rev. Lett. **92**, 036804 (2004)
14. I. Mahboob, T.D. Veal, L.F.J. Piper, C.F. McConville, H. Lu, W.J. Schaff, J. Furthmuller, F. Bechstedt, Phys. Rev. B **69**, 201307(R) (2004)
15. H. Lu, W.J. Schaff, L.F. Eastman, C.E. Stutz, Appl. Phys. Lett. **82**, 1736–1738 (2003)
16. K.A. Rickert, A.B. Ellis, F.J. Himpsel, H. Lu, W. Schaff, J.M. Redwing, F. Dwikusuma, T.F. Kuech, Appl. Phys. Lett. **82**, 3254–3256 (2003)
17. T.D. Veal, L.F.J. Piper, M.R. Phillips, M.H. Zareie, H. Lu, W.J. Schaff, C.F. McConville, Phys. Status Solidi A **203**, 85–92 (2006)
18. C. Stampfl, C.G. Van de Walle, D. Vogel, P. Krüger, J Pollmann Phys. Rev. B **61**, 7846–7849 (2000)
19. L.Ö. Olsson, C.B.M. Andersson, M.C. Håkansson, J. Kanski, L. Ilver, U.O. Karlsson, Phys. Rev. Lett. **76**, 3626–3629 (1996)
20. V.Y. Aristov, V.M. Zhilin, C. Grupp, A. Taleb-Ibrahimi, H.J. Kim, P.S. Mangat, P. Soukiassian, G. Le Lay, Appl. Surf. Sci. **166**, 263–267 (2000)
21. G. Abstreiter, R. Huber, G. Trankle, B. Vinter, Solid State Commun. **47**, 651–654 (1983)
22. L. Colakerol, L.F.J. Piper, A. Fedorov, T.C. Chen, T.D. Moustakas, K.E. Smith, EPL (Europhysics Letters) **83**, 47003 (2008)
23. L.F.J. Piper, L. Colakerol, T. Learmonth, P.A. Glans, K.E. Smith, F. Fuchs, J. Furthmueller, F. Bechstedt, T.C. Chen, T.D. Moustakas, J.H. Guo, Phys. Rev. B **76**, 245204 (2007)
24. K. Ortner, X.C. Zhang, A. Pfeuffer-Jeschke, C.R. Becker, G. Landwehr, L.W. Molenkamp, Phys. Rev. B **66**, 075322 (2002)
25. R.K. Hayden, D.K. Maude, L. Eaves, E.C. Valadares, M. Henini, F.W. Sheard, O.H. Hughes, J.C. Portal, L. Cury, Phys. Rev. Lett. **66**, 1749–1752 (1991)
26. J. Bhattacharyya, S. Ghosh, Physica status solidia (a) **204**, 439–446 (2007)
27. A. Delimitis, P. Komninou, G.P. Dimitrakopoulos, T. Kehagias, J. Kioseoglou, T. Karakostas, G. Nouet, Appl. Phys. Lett. **90**, 061920 (2007)
28. T. Bottcher, S. Einfeldt, S. Figge, R. Chierchia, H. Heinke, D. Hommel, J.S. Speck, Appl. Phys. Lett. **78**, 1976–1978 (2001)
29. T.C.P. Chen, C. Thomidis, J. Abell, W. Li, T.D. Moustakas, J. Cryst. Growth **288**, 254–260 (2006)

30. V. Darakchieva, K. Lorenz, N.P. Barradas, E. Alves, B. Monemar, M. Schubert, N. Franco, C.L. Hsiao, L.C. Chen, W.J. Schaff, L.W. Tu, T. Yamaguchi, Y. Nanishi, *Appl. Phys. Lett.* **96**, 081907 (2010)
31. E. Kalesaki, J. Kioseoglou, L. Lympierakis, P. Komninou, T. Karakostas, *Appl. Phys. Lett.* **98**, 072103 (2011)
32. N. Miller, E.E. Haller, G. Koblmuller, C. Gallinat, J.S. Speck, W.J. Schaff, M.E. Hawkrige, K.M. Yu, J.W. Ager, *Phys. Rev. B* **84**, 075315 (2011)
33. V. Cimalla, V. Lebedev, F.M. Morales, R. Goldhahn, O. Ambacher, *Appl. Phys. Lett.* **89**, 172109 (2006)
34. T. Nagata, G. Koblmuller, O. Bierwagen, C.S. Gallinat, J.S. Speck, *Appl. Phys. Lett.* **95**, 132104 (2009)
35. D. Segev, C.G. Van de Walle, *Europhys. Lett.* **76**, 305–311 (2006)
36. L. Colakerol, T.D. Veal, H.K. Jeong, L. Plucinski, A. DeMasi, T. Learmonth, P.A. Glans, S.C. Wang, Y.F. Zhang, L.F.J. Piper, P.H. Jefferson, A. Fedorov, T.C. Chen, T.D. Moustakas, C.F. McConville, K.E. Smith, *Phys. Rev. Lett.* **97**, 237601 (2006)
37. Y.H. Lai, C.T. Yeh, J.M. Hwang, H.L. Hwang, C.T. Chen, W.H. Hung, *J. Phys. Chem. B* **105**, 10029–10036 (2001)
38. S. Krischok, V. Yanev, O. Balykov, M. Himmerlich, J.A. Schaefer, R. Kosiba, G. Ecke, I. Cimalla, V. Cimalla, O. Ambacher, H. Lu, W.J. Schaff, L.F. Eastman, *Surf. Sci.* **566**, 849–855 (2004)
39. E.O. Kane, *J. Phys. Chem. Solids* **1**, 249–261 (1957)
40. K.H. Bevan, M.S. Hossain, *Semicond. Sci. Tech.* **27**, 105029 (2012)
41. B.R. Nag, Springer, Berlin, (1980)
42. W. Walukiewicz, J.W. Ager, K.M. Yu, Z. Liliental-Weber, J. Wu, S.X. Li, R.E. Jones, J.D. Denlinger, *J. Phys. D Appl. Phys.* **39**, R83–R99 (2006)
43. C.L. Wu, H.M. Lee, C.T. Kuo, C.H. Chen, S. Gwo, *Phys. Rev. Lett.* **101**, 106803 (2008)
44. C.G. Van de Walle, D. Segev, *J. Appl. Phys.* **101**, 081704 (2007)
45. P.D.C. King, T.D. Veal, C.S. Gallinat, G. Koblmuller, L.R. Bailey, J.S. Speck, C.F. McConville, *J. Appl. Phys.* **104**, 103703 (2008)
46. Y. Ishitani, M. Fujiwara, X. Wang, S.B. Che, A. Yoshikawa, *Appl Phys Lett*, 92 (2008)
47. G.F. Brown, J.W. Ager, W. Walukiewicz, W.J. Schaff, J. Wu, *Appl. Phys. Lett.* **93**, 262105 (2008)
48. V. Lebedev, V. Cimalla, T. Baumann, O. Ambacher, F.M. Morales, J.G. Lozano, D. Gonzalez, *J. Appl. Phys.* **100**, 094903 (2006)
49. T. Stoica, R. Meijers, R. Calarco, T. Richter, H. Luth, *J. Cryst. Growth* **290**, 241–247 (2006)
50. S. Zhao, O. Salehzadeh, S. Alagha, K.L. Kavanagh, S.P. Watkins, Z. Mi, *Appl. Phys. Lett.* **102**, 073102 (2013)
51. B.H. Le, S.R. Zhao, N.H. Tran, Z.T. Mi, *Appl. Phys. Lett.* **105**, 231124 (2014)
52. E. Calleja, J. Grandal, M.A. Sanchez-Garcia, M. Niebelschutz, V. Cimalla, O. Ambacher, *Appl. Phys. Lett.* **90**, 262110 (2007)
53. R.S. Chen, T.H. Yang, H.Y. Chen, L.C. Chen, K.H. Chen, Y.J. Yang, C.H. Su, C.R. Lin, *Nanotechnology* **22**, 425702 (2011)
54. O. Ambacher, *J. Phys. D Appl. Phys.* **31**, 2653–2710 (1998)
55. F. Bernardini, V. Fiorentini, D. Vanderbilt, *Phys. Rev. B* **56**, 10024–10027 (1997)
56. T. Richter, H. Luth, T. Schapers, R. Meijers, K. Jeganathan, S.E. Hernandez, R. Calarco, M. Marso, *Nanotechnology* **20**, 405206 (2009)
57. F. Werner, F. Limbach, M. Carsten, C. Denker, J. Malindretos, A. Rizzi, *Nano Lett.* **9**, 1567–1571 (2009)
58. D.R. Khanal, W. Walukiewicz, J. Grandal, E. Calleja, J. Wu, *Appl. Phys. Lett.* **95**, 173114 (2009)
59. T. Stoica, R.J. Meijers, R. Calarco, T. Richter, E. Sutter, H. Luth, *Nano Lett.* **6**, 1541–1547 (2006)
60. S. Lazic, E. Gallardo, J.M. Calleja, F. Agullo-Rueda, J. Grandal, M.A. Sanchez-Garcia, E. Calleja, E. Luna, A. Trampert, *Phys. Rev. B* **76**, 205319 (2007)

61. E.O. Schafer-Nolte, T. Stoica, T. Gotschke, F.A. Limbach, E. Sutter, P. Sutter, D. Grutzmacher, R. Calarco, *Nanotechnology* **21** (2010)
62. S. Zhao, S. Fatholouloumi, K.H. Bevan, D.P. Liu, M.G. Kibria, Q. Li, G.T. Wang, H. Guo, Z. Mi, *Nano Lett.* **12**, 2877–2882 (2012)
63. S. Zhao, Z. Mi, M.G. Kibria, Q. Li, G.T. Wang, *Phys. Rev. B* **85**, 245313 (2012)
64. S.R. Zhao, Z.T. Mi, B.H. Le, *Gallium Nitride Materials and Devices IX* **8986**, 898617 (2014)
65. Y.J. Bai, Z.G. Liu, X.G. Xu, D.L. Cui, X.P. Hao, X. Feng, Q.L. Wang, *J. Cryst. Growth* **241**, 189–192 (2002)
66. Z.H. Lan, W.M. Wang, C.L. Sun, S.C. Shi, C.W. Hsu, T.T. Chen, K.H. Chen, C.C. Chen, Y.F. Chen, L.C. Chen, *J. Cryst. Growth* **269**, 87–94 (2004)
67. J. Zhang, L. Zhang, X.S. Peng, X.F. Wang, *J. Mater. Chem.* **12**, 802–804 (2002)
68. X.Q. Wang, S.B. Che, Y. Ishitani, A. Yoshikawa, *Physica Status Solidi C—Curr. Topics Solid State Phys.* **3**(6), 1561–1565 (2006)
69. Y.L. Chang, F. Li, A. Fatehi, Z.T. Mi, *Nanotechnology* **20**, 345203 (2009)
70. J.H. Chai, T.H. Myers, Y.W. Song, R.J. Reeves, W.M. Linhart, R.J.H. Morris, T.D. Veal, M.G. Dowsett, C.F. McConville, S.M. Durbin, *J. Vac. Sci. Technol. B* **30**, 02b124 (2012)
71. J.H. Chai, Y.W. Song, R.J. Reeves, S.M. Durbin, *Phys. Status Solidi A* **209**, 95–99 (2012)
72. M.A. Mayer, S. Choi, O. Bierwagen, H.M. Smith, E.E. Haller, J.S. Speck, W. Walukiewicz, *J. Appl. Phys.* **110**, 123707 (2011)
73. I. Wilke, *Ultrafast Phenom. Semicond. Nanostruct. Mater.* **XIV**, 76001 (2010)
74. J.W. Ager, N. Miller, R.E. Jones, K.M. Yu, J. Wu, W.J. Schaff, W. Walukiewicz, *Physica Status Solidi B—Basic Solid State Phys.* **245**, 873–877 (2008)
75. P.D.C. King, T.D. Veal, P.H. Jefferson, C.F. McConville, H. Lu, W.J. Schaff, *Phys. Rev. B* **75**, 115312 (2007)
76. J.W. Ager, K.M. Yu, R.E. Jones, D.M. Yamaguchi, S.X. Li, W. Walukiewicz, E.E. Haller, H. Lu, W.J. Schaff, *AIP Conf. Proc.* **893**, 343–344 (2007)
77. K.M. Tracy, W.J. Mecouch, R.F. Davis, R.J. Nemanich, *J. Appl. Phys.* **94**, 3163–3172 (2003)
78. S. Zhao, B.H. Le, D.P. Liu, X.D. Liu, M.G. Kibria, T. Szkopek, H. Guo, Z. Mi, *Nano Lett.* **13**, 5509–5513 (2013)

Chapter 14

Optical and Structural Properties of Quantum Dots

M.H. Yükselici, A. Aşıkoğlu Bozkurt, Ç. Allahverdi, Z. Nassar, D. Bulut, B. Can Ömür, M.K. Torun and A.T. İnce

Abstract We report (i) thickness dependent evolution of structural disorder, strain on crystalline planes and grain size in chemical bath deposited (CBD) CdS thin films studied through a combinative evaluation of the results of optical absorption, Raman spectroscopies, X-Ray diffraction (XRD) and Scanning Electron Microscopy (SEM). (ii) refer briefly to $\text{CdSe}_x\text{S}_{1-x}$ nanocrystals in liquid and (iii) address quantum size effect in $\text{CdSe}_x\text{S}_{1-x}$ quantum dots embedded in glass studied through steady state photoluminescence spectroscopy. The asymptotic optical absorption edge is red shifted while the long wavelength tail narrows with increasing thickness which is proportional to deposition time. We employ effective mass theory under quantum size effect to estimate average grain size from the energetic position of asymptotic optical absorption edge. The long wavelength tail optical absorption is due presumably to the micro-electric field induced by crystalline defects. The transmission probability through the potential energy barrier created by micro-electric field is calculated with the help of WKB (Wentzel, Kramers, Brillouin) approximation. We conclude that as the deposition time increases from 10 to 150 min, the average grain radius changes by 2 nm, Urbach energy and the electric micro-field decrease from 600 to 400 meV and 2240–820 kV/mm respectively. The shift in XRD pattern shows that the compressive strain decreases with growth. The Raman LO_1 vibrational mode display an increase up to 22 min of deposition time and then a decrease.

M.H. Yükselici (✉) · A. Aşıkoğlu Bozkurt · Z. Nassar · B. Can Ömür
Faculty of Science and Letters, Department of Physics, Yıldız Technical University,
Davutpaşa Campus, Esenler, Davutpaşa Mah., Davutpaşa Caddesi, 34220 İstanbul, Turkey
e-mail: mhyukselici@gmail.com

Ç. Allahverdi
Electrical-Electronics Engineering Department, Faculty of Engineering,
Toros University, 45 Evler Campus/Bahçelievler Mahallesi,
16. Street, No: 1/7 Yenisehir, 33140 Mersin, Turkey

D. Bulut · M.K. Torun · A.T. İnce
Department of Physics, İnönüMah. Kayışdağı Cad., Yeditepe University, 26
AğustosYerleşimi, 34755 Ataşehir, İstanbul, Turkey

14.1 Introduction

Cd chalcogenide compounds of II–VI semiconductor nanocrystals (quantum dots) are of special interest due to their technological applications especially in solar cells, opto-electronics, biotechnology, medicine and light emitting diodes. When the size of bulk semiconductor is reduced to nanometer scales, the crystal gains new optical and electronic properties. Energy levels are quantized as in atoms and the energy difference between levels widens with the inverse of the size squared. The efficiency for the luminescence emission of the levels increases. The techniques for obtaining II–VI nanocrystals with good luminescent characteristics and uniform size distribution include colloidal synthesis of nanocrystals from precursors at the presence of capping agents or stabilizers in order to achieve the desired size-dependent optical properties and solid phase precipitation of the crystallites in glass matrix through nucleation followed by diffusion limited growth and then Ostwald ripening among others. II–VI thin films deposited by various techniques such as solution growth and thermal evaporation possess optical and structural properties which might be explained with the help of quantum size effect. In this chapter we firstly review our research on (i) $\text{CdSe}_x\text{S}_{1-x}$ nanocrystals which have been extensively studied model system and (ii) the most proper nanocrystal of group II–VI semiconductors which take an advantage of strong quantum size effect that is, CdTe of band gap energy of 1.5 eV (850 nm) and thirdly report (iii) thickness dependent evolution of structural disorder, strain on crystalline planes and grain size in chemical bath deposited CdS thin films. The asymptotic optical absorption edge is red shifted while the long wavelength tail narrows with increasing thickness which is proportional to deposition time. We employ effective mass theory under quantum size effect to estimate average grain size from the energetic position of asymptotic optical absorption edge. The long wavelength tail optical absorption is due presumably to the micro-electric field induced by crystalline defects. The transmission probability through the potential energy barrier created by micro-electric field is calculated with the help of WKB approximation. We conclude that as the deposition time increases from 10 to 150 min, the average grain radius changes by 2 nm, Urbach energy and the electric micro-field decrease from 600 to 400 meV and 2240 to 820 kV/mm respectively.

14.2 $\text{CdSe}_x\text{S}_{1-x}$ Nanocrystals

Synthesis of binary and ternary II–VI group semiconductor nanocrystal systems which have convenient size, size distribution, shape and composition in liquid media is highly important with respect to recent technological applications. For instance, the II–VI group semiconductor nanocrystals which have sufficient quantum efficiency are used in active layers of light emission diodes (LEDs) in order to improve color purity and color variety of the diodes [1]. Semiconductor nanocrystals can be

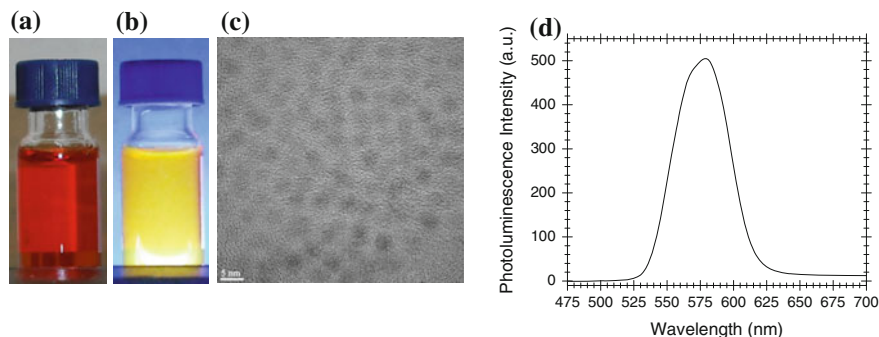


Fig. 14.1 CdSe nanocrystals dispersed in toluene. Their appearance under daylight (a) and under ~ 366 nm UV light (b). HRTEM image (c) and photoluminescence spectrum (d) of these nanocrystals. TEM scale of 5 nm is given at the bottom left corner in (c)

formed and grown in aqueous and non-aqueous liquids with methods of microemulsion [2], chemical reduction [3], cold and hot injection [4] etc. As shown in Fig. 14.1a, CdSe nanocrystals growth in non-coordinating octadecene with hot injection method and then dispersed into toluene are shown. These nanocrystals whose surfaces are modified with stearic acid ($C_{17}H_{35}COOH$) molecules can be kept in toluene longer than one year. Their appearance under ~ 366 nm UV light is shown in Fig. 14.1b. It can be deduced that these nanocrystals are nearly spherical and have diameter of ~ 3 nm from their high resolution transmission electron (HRTEM) image given at Fig. 14.1c. Measured photoluminescence spectrum of these synthesized nanocrystals is shown in Fig. 14.1d.

14.3 Investigation of Raman Spectroscopy for CdTe Thin Film

Four CdTe thin film samples are produced by evaporating CdTe compound onto glass substrates by using physical vapor deposition (PVD) technique. The samples are investigated structurally and optically by Raman scattering. Relative to the bulk phonon frequency, longitudinal acoustics (LA) phonon frequency (~ 100 cm^{-1}) is redshifted for all samples. This shift may be produced as a result of the effect of finite particle size and/or strain in the thin film.

14.3.1 Experimental Details

The data for Raman scattering were measured at room temperature by using a spectrometer employing Rayleigh line rejection filter. The samples were excited by 3 mW laser having a wave length of 532 nm and a spot size of 1 μm .

14.3.2 Modelling

First order Raman spectrum can be fitted to the confinement model discussed in [5]:

$$I(\omega) = \int_{q=0}^{q=\frac{2\pi}{a}} \frac{|C(0, q)|^2 d^3 q}{[\omega - \omega(q)]^2 + (\Gamma_0/2)^2} \quad (3.1)$$

$$|C(0, q)|^2 = \exp(-q^2 d^2 / 2\alpha) \quad (3.2)$$

where d is the diameter of the grain, a is the lattice parameter, q is the phonon wave vector, α is a value determines how rapidly the wave function of the phonon decays as the boundary is approached, Γ_0 is the natural line width of the optical phonon in bulk, $\omega(q)$ is the dispersion curve:

$$\omega(q) = \omega_0 - \Delta\omega \sin^2(qa/4) \quad (3.3)$$

ω_0 is the phonon frequency and $\Delta\omega$ is the width of the dispersion curve.

14.3.3 Discussion

First order Raman spectra for the CdTe thin film samples are plotted in Fig. 14.2 according to the model in (3.1). Figure 14.2 shows lines centered at phonon frequencies closed to that for the L-critical-point LA phonon which has a frequency of $3.25 \text{ THz} = 108.333 \text{ cm}^{-1}$ [6]. In this work, we observe that phonon frequency red shifts as grain size decreases. This decrease in phonon frequency can be attributed to the effect of compressive strain and/or phonon confinement due to finite grain size.

Fig. 14.2 Intensity against Raman shift for CdTe thin films. Line position is $\sim 100 \text{ cm}^{-1}$ which is close to L-critical-point LA phonon

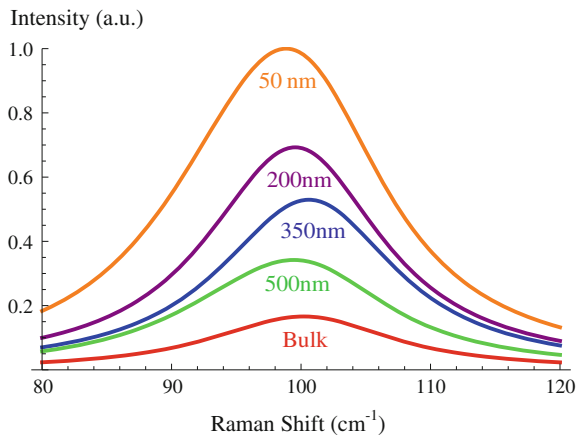


Table 14.1 Line position, line width and grain size calculated by Raman spectroscopy for CdTe thin film

Thickness (nm)		Line position (cm^{-1})	Line width (cm^{-1})	R_{ave} (nm)
50		98.9	18.4	1.380
200		99.6	16.4	1.602
350		100.6	16.4	1.750
500		99.4	18.0	1.984
Bulk	Measured	100.2	16.8	
	[2]	108.333		

Referring to Table 14.1, the grain size calculated by Raman spectroscopy decreases to a value less than Bohr radius of CdTe bulk (7.3 nm [7]) as the film thickness decreases and hence the phonon confinement becomes stronger with decreasing the thickness causing red shift in phonon frequency.

14.3.4 Importance of the Subject

In general, the compounds of group II–VI can be considered as materials of valuable practical importance, the fact that they exhibit many considerable solid-state phenomena. As a special case of these compounds, CdTe is a material which can be considered as a completely representative of the group II–VI for its distinctive physical properties. Moreover, CdTe is a promising compound since it is used in electronic and optical manufactures. Studies about CdTe are scarce especially those related to Raman spectroscopy. Some articles have investigated the structural and optical properties of CdTe using XRD and/or optical transmission spectroscopy but they have not used Raman spectroscopy, the references [8–15] are given as examples for some of these articles. On the other hand, there are limited number of articles which are related to Raman spectroscopy analysis for CdTe [16–18]. These articles didn't use the confinement model described in the present work, which has the advantage of that it can compute the fitting parameters like grain size and line width, for analysing the spectrum; in [19] the data was fitted to Lorentzian profile and the spectrum in [20] was studied using a different model described in [21], while [16–18] did not use the fitting at all and hence they could not estimate the grain size and line width.

14.3.5 Section Summary

PVD technique was used for depositing CdTe on glass plates. Four samples of CdTe thin films of different thicknesses were obtained. The samples are investigated by Raman spectroscopy to study the effect of grain size on phonon

confinement. The confinement model evaluated grain sizes ranging from about 1.4 to 2.0 nm. The importance of this work lies in the fact that there are few articles that concern to analysis of Raman spectra for CdTe using the above model.

14.4 Steady State Photoluminescence Spectroscopy

The experimental data, results and discussions presented in this section are mainly cited from and/or based on [22, 23]. In previous studies of $\text{CdS}_x\text{Se}_{1-x}$ which is a member of group II–IV quantum dots, it is revealed that photoluminescence (PL) spectra generally has a three peak structure; one of them is close to the absorption (ABS) edge due to direct electron hole recombination, the other one is attributed to recombination from another energy levels within the band gap and the last low energy peak is related to deep trap levels. The first peak mentioned above is assigned to volume contribution and the other peaks are presumed to be surface states [24, 25]. Hache et al. studied with emphasis on evolution of intensity of the excitation beam in PL spectra. It is employed that at low intensity there is only one peak which is related to deep trap levels. As the intensity increases, higher energy peak (L_1) due to direct electron hole pair recombination remains stationary, whereas lower energy peak (L_2) that leads to recombination through the new unidentified states becomes visible. At higher intensities, both of the peaks show up and begin to get sharper with respect to decrease in deep trap level luminescence. In addition to this, they also figured out that recombination from another states has more efficiency than direct electron hole pair recombination. G. Mei et al., who studied on variously sized quantum dots embedded in glass matrix, pointed out that the PL peak near the absorption edge is related to bound-exciton recombination and indicates the impurity or surface states of $\text{CdS}_x\text{Se}_{1-x}$ quantum dots. As the particle size gets smaller, more surface states occur in PL spectrum. In other words, since the surface to volume ratio of quantum dot spheres increases, the PL spectrum is dominated by surface states of decreasing particle sizes [26, 22]. Rodden et al. scrutinized the effect of annealing at different temperatures and doping concentrations for $\text{CdS}_x\text{Se}_{1-x}$ samples [27]. In PL spectra a shift is observed through the lower energies by increasing temperatures or doping concentrations. Increasing temperature results in an increase of crystallite size. As the crystallite size becomes larger, the effective energy gap of the $\text{CdS}_x\text{Se}_{1-x}$ sample shifts towards lower energies, indicating an existence of three-dimensional quantum confinement of electrons and holes. A reduce in doping concentrations causes a growth of smaller particles, hence a shift towards the higher energy is observed in PL spectra. In more recent works, nanocrystals in a range of 2.3–4.2 nm radii are analyzed with a quantized state effective mass theory by Allahverdi et al. [23]. They figured out the PL spectra with the help of the model Ravindran et al. and also concentrate on deep trap peak and amount of its shift from the first exciton band. The peaks which are located at ~ 1.56 and ~ 1.6 eV remains stationary as the heat treatment progresses. However, deep trap peak in PL and the first exciton peak in ABS shift towards

lower energies (38 meV for deep trap peak) due to quantum confinement effect. The shift can also arise from the variation in S content that may increase/decrease with the heat treatment [26]. Since the shift due to S content is about ~ 20 meV, it is presumed that the redshift of 38 meV in deep trap peak represents the quantum confinement effect. With the help of the model Ravindran et al. presented, Allahverdi et al. observed that a decrease in energy occurs between the deep trap and first exciton peak with increasing average particle radius. It can be inferred from the ABS and PL spectra, the narrower shift, the larger are the particles. This indicates that deep trap levels have less sensitivity to variations in average quantum dot radius than that of direct recombination levels in $\text{Cd}_{0.08}\text{Se}_{0.92}$ quantum dots.

We have focussed on trap sites of variously sized samples of $\text{CdS}_x\text{Se}_{1-x}$ QDs in glass. As-received RG695 Schott borosilicate glass doped with $\text{CdS}_{0.08}\text{Se}_{0.92}$ was melted at 1000 °C for 15 min in order to dissolve the particles and after quenching to room temperature, the glass was annealed at 450 °C for 5 h to decrease the stress in it. Finally the samples were exposed to heat treatment gradually from 600 to 675 °C for different periods of times as listed in Table 14.2 [22].

In order to identify the trap sites of gradually annealed $\text{CdS}_x\text{Se}_{1-x}$ samples, we present the steady state PL spectra (light colored curve) with the Gaussian model lineshape (dark colored curve) on it in Fig. 14.3 [22].

The modelling function was employed below, was also described in [28]. Hence, the calculated PL intensity has the following form

$$I_{PL}(h\nu) = \frac{A}{\sqrt{2\pi}\sigma_R} \exp\left[-\frac{(h\nu - E_{trap})^2}{2\Gamma_{trap}^2}\right] + \sum_R \left\{ \frac{1}{\sqrt{2\pi}\sigma_R} \exp\left[-\frac{(R - R_0)^2}{2\sigma_R^2}\right] \frac{B}{\sqrt{2\pi}\Gamma} \exp\left[-\frac{(h\nu - E_R)^2}{2\Gamma^2}\right] \right\} \quad (4.1)$$

Table 14.2 Heat treated $\text{CdS}_x\text{Se}_{1-x}$ samples for different temperatures and times

Sample number	Annealing temperature (°C)	Time (h)	R_0^* (nm)	R_0^{**} (nm)	σ^* (nm)
1	600	1/2	—	—	—
2	625	1	1.7 ± 0.38	2.1	1.2 ± 0.3
3	625	2	1.8 ± 0.36	2.2	1.3 ± 0.3
4	625	4	2.9 ± 0.10	2.5	0.32 ± 0.01
5	625	8	3.3 ± 0.73	2.7	0.22 ± 0.05
6	650	2	3.4 ± 0.17	3.1	0.55 ± 0.03
7	650	4	4.3 ± 0.87	3.4	0.70 ± 0.14
8	650	8	4.8 ± 0.17	3.5	0.70 ± 0.03
9	675	2	7.2 ± 0.52	4.5	1.3 ± 0.1
10	675	4	8.5 ± 1.3	4.7	1.3 ± 0.2

Average radius (R_0^*) and size dispersion (σ^*) calculated by the employed model of PL spectra. Average radius (R_0^{**}) obtained from the energetic position of the first exciton peak is pointed out for comparison

where A and B = the constants, R_0 = radius of the spherical quantum dot, σ_R = standard deviation of the radius, $\Gamma = 2$ % width of homogenous broadening, $\Gamma_{trap} = 10$ –20 % width of trap band broadening.

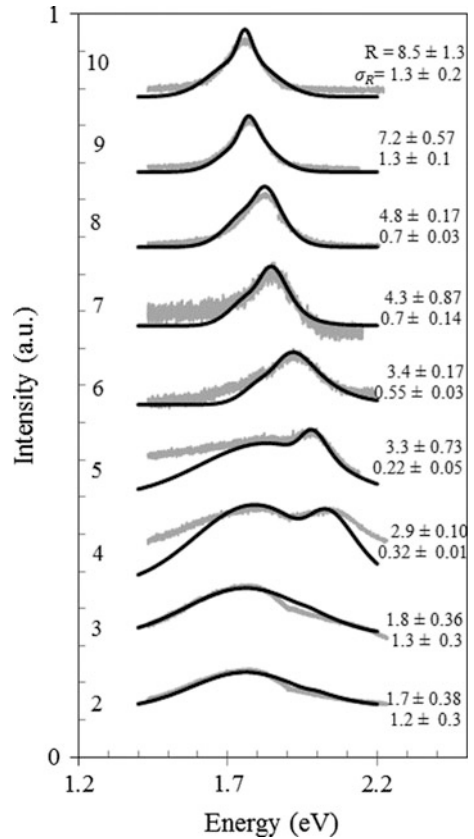
We consider that $\Gamma = (\Delta E/E) \times 100$ where ΔE indicates the standard deviation homogenous at an energy of E . The size dependent first exciton transition energy is characterized by following relation [29].

$$E_R \text{ (eV)} = E_g \text{ (eV)} - \frac{0.14}{R \text{ (nm)}} + \frac{0.356}{\mu [R \text{ (nm)}]^2} \quad (4.2)$$

The first term leads to the bulk band gap of $\text{CdS}_{0.08}\text{Se}_{0.92}$ of $E_g = 1.727$ eV while the second term is responsible for the Coulomb potential energy between an electron and a hole. μ represents $0.106m_0$ that m_0 indicates electron rest mass.

As seen in Fig. 14.3, in the first periods of times and temperatures the dark colored curves 2 and 3 exhibits one peak at ~ 1.8 eV. We presume that this peak is attributed to volume traps and the model figures out that it is almost stationary. When the heat treatment progresses, another peak begins to resolve at ~ 2.1 eV and shifts

Fig. 14.3 Light colored lines indicate the steady state PL spectra obtained with 488 nm cw Ar^+ laser of the intensity of ~ 70 nW/mm². Dark colored lines refer to simulated PL spectra. Average nanocrystal radius (R_0) and size dispersion (σ_0) are the best fit parameters which are presented beside each pair of curve. Heat treatment temperatures and times numbered as (1) 600 $\frac{1}{2}$, (2) 625, 1; (3) 625, 2; (4) 625, 4; (5) 625, 8; (6) 650, 2; (7) 650, 4; (8) 650, 8; (9) 675, 2; (10) 675, 4, respectively. See Table 14.2 for details



through the lower energies for samples 4 and 5. During the gradual increase in heat treatment, this peak overlaps with the stationary peak which leads to the volume trap band from samples 5 to 10. According to the modelling, this peak is related to a size dependent band and identified as a surface assisted electron hole recombination band. It is redshifted from the first exciton peak by amount of 120 meV and the length of the shift decreases with respect to increasing radius of the QDs.

14.5 The Progression of Strain and Micro-electric Field Dependent Urbach Energy with Deposition Time in Chemical Bath Deposited CdS Thin Films

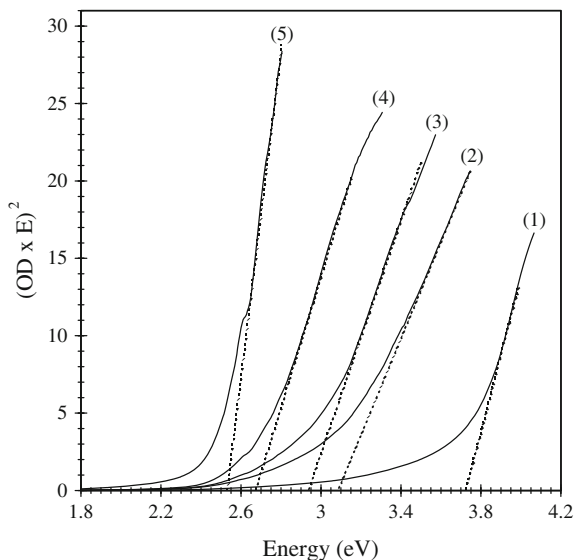
Thin film CdS semiconductors are of special interest due to their technological applications in solar cells. CdS can be used as a window material in hetero-junctions due to its wide band gap energy of ~ 2.5 eV [30–34]. Chemical bath deposition (CBD) is a low cost method to grow thin films over wide surfaces [35–40]. Most of studies on CBD CdS have gone into the optimization of experimental parameters such as bath temperature, pH value, concentration of reactants, deposition time and post-deposition heat treatment. On the other hand, recent Raman studies show that the interfacial strain and phonon confinement modify the phonon vibrational modes. A 0.35 % change in lattice constant for CdSe nanotubes produces a blue shift of 2.4 cm^{-1} in LO_1 mode in Raman spectrum [41]. CdSe nanocrystals of an average radius of 3 nm and a size distribution of 10 % or FWHM of 0.3 nm result in a red shift of 1.82 cm^{-1} in LO_1 mode relative to bulk structure of the same composition [42]. In this section we employ optical absorption spectroscopy (ABS), Scanning Electron Microscopy (SEM) and X-Ray Diffraction (XRD) spectroscopy for initial characterization of thin film CBD CdS structures. We analyze the optical absorption (ABS) data with the help of Wentzel, Kramers, Brillouin (WKB) approach to relate the long wavelength tail below the fundamental absorption edge to micro-electric fields induced by crystalline disorder. Finally we employ resonant Raman spectroscopy to assess the relative weight of results obtained by XRD data with the help of Grüneisen model and phonon confinement.

14.5.1 Experimental

14.5.1.1 Sample Preparation

CdS thin films were deposited on glass substrates by chemical bath deposition (CBD). Glass substrates were cleaned and dipped into a 60 °C reaction bath of pH = 11.6 prepared from 0.1 M of cadmium sulfate, 1 M of thiourea each separately

Fig. 14.4 $(OD \times hv)^2$ plotted against photon energy ($h\nu$) for chemical bath-deposited CdS thin films. Deposition times are numbered as follows: 1 $\rightarrow$ 10, 2 $\rightarrow$ 16, 3 $\rightarrow$ 22, 4 $\rightarrow$ 40, 5 $\rightarrow$ 150 min



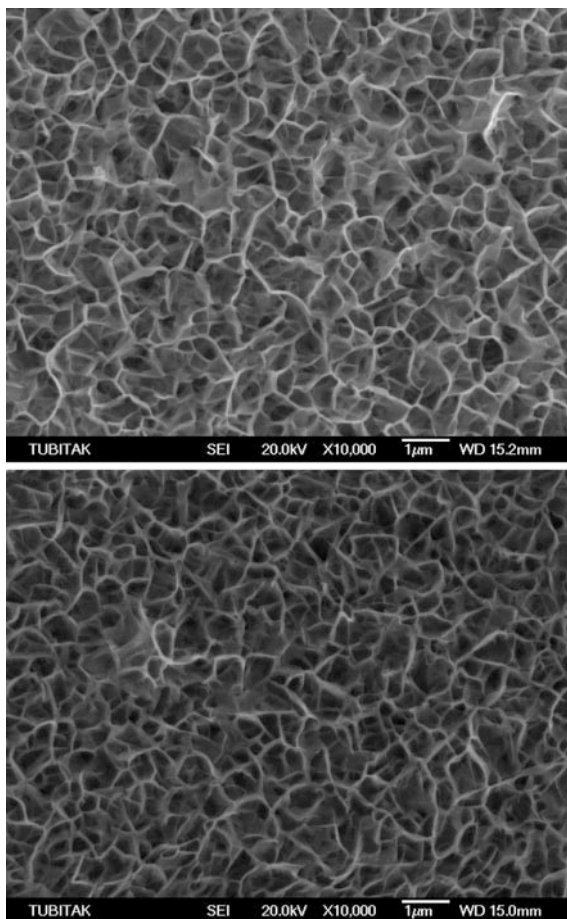
dissolved in 30 ml of distilled water and 5 % ammonium hydroxide according to the procedure reported in detail elsewhere [43]. A set of 5 glass slides deposited for 10, 16, 22, 40 and 150 min under stirring at 250 rpm is scrutinized in this work.

14.5.1.2 Optical Absorption

Optical density (OD) for each sample is calculated by optical transmission measurements. As shown in Fig. 14.4, the straight-line portions of the $(OD \times hv)^2 - hv$ curves are extrapolated to obtain the band gap energies (E_g) as intercept with the x axis according to the relation $(OD \times hv)^2 = (A^*d)^2(hv - E_g)$, where A^* is a constant and d film thickness [44]. As the deposition time increases, the asymptotic absorption edge with a tail structure shifts to lower energies. A solution to Schrödinger Equation for a particle in a spherical box under effective mass model yields a relation between the energetic position of first exciton peak and average particle's radius [45]. The blue shift is related to grain size by 1E_R (eV) = E_g (eV) - $0.1655/R$ (nm) + $0.376/\mu[R$ (nm)]² where the first term is bulk band gap of CdS crystal of $E_g = 2.402$ eV, the second term the attractive Coulomb potential energy and the third term the kinetic energy for a particle in a spherical box ($\mu = 0.1674m_0$, m_0 is the electron rest mass). The average grain radii R_{ave} (nm) for each film are calculated as 1.21, 1.71, 1.91, 2.60 and 3.59 nm for samples 1, 2, 3, 4 and 5 respectively.

¹Derived by using [16].

Fig. 14.5 Surface SEM micrographs for sample 2 (*top*) and sample 4 (*bottom*)



14.5.1.3 SEM and EDS

JEOL/35M-6335F scanning electron microscope (SEM) was employed for the surface morphological analysis. SEM surface micrographs are shown in Fig. 14.5 and cross-sectional micrographs in Fig. 14.6 for samples 2 and 4 respectively. SEM micrographs have an irregular and anfractuous surface like a web pattern. We assume that powdery and nonsticky film deposition give rise to an extremely rough surface. Figure 14.7 shows Energy Dispersive X-ray spectrum (EDS) of CdS film deposited for 40 min (sample 4). The most intense peak is due to Si of atomic % of 22.47 which is presumably from glass substrate. The second and third intense peaks might be due to Cd of atomic % of 11.89 and S of 11.55.

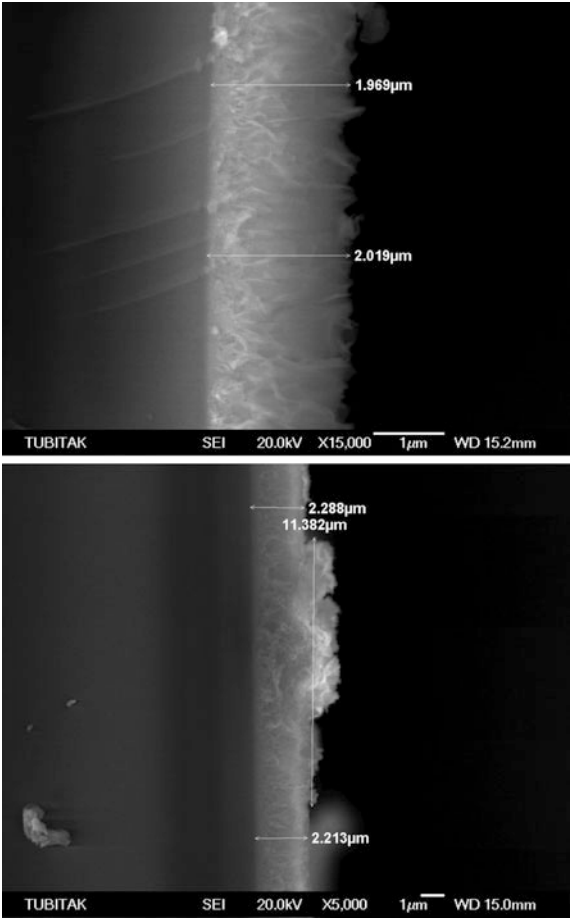


Fig. 14.6 Cross sectional SEM micrographs for sample 2 (*top*) and sample 4 (*bottom*)

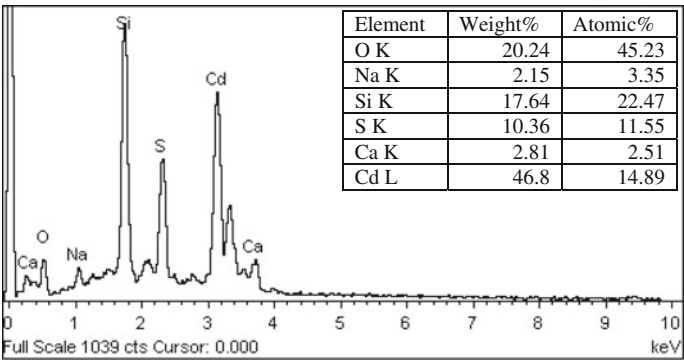
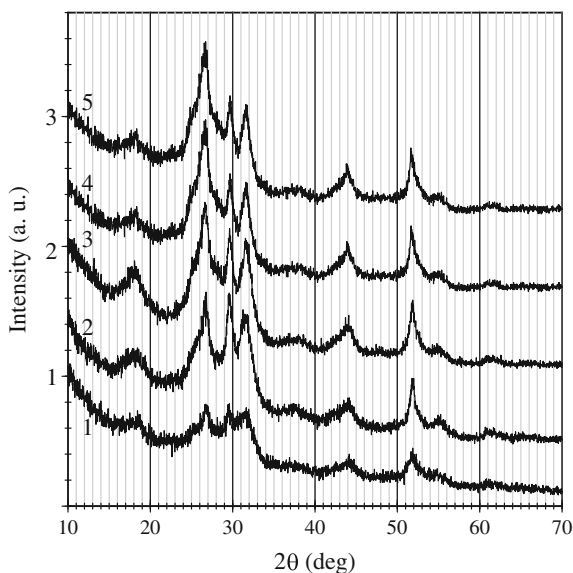


Fig. 14.7 Energy dispersive X-ray spectrum (EDS) of CdS thin film deposited for 40 min (sample 4)

14.5.1.4 XRD

XRD patterns recorded on Panalytical MPD diffractometer are given in Fig. 14.8. We fit each XRD spectrum with a Lorentzian line profile for each peak/shoulder/hump and a cubic background function and then find best fitting parameters for peak positions and full width at half maximum (FWHM). We presume that the peaks at $2\theta = 26.5^\circ$, 43.9° , 51.9° and 54.7° correspond to H[002]/C[111], H[110]/C[220], H[112]/C[311] and H[004]/C[222] planes of hexagonal (H) (ICDD PDF No: 03-065-8873) and/or cubic (C) (ICDD PDF No: 03-065-2887) planes of CdS crystal. The peaks at 18.3° and 31.6° are assigned to cubic planes of CdSO_4 structure (ICDD PDF No: 00-003-0453). The reported studies show that CBD CdS thin films grow on glass substrates in cubic zincblende, hexagonal wurtzite or mixed H and C structures depending on bath temperature and post-deposition heat-treatment. The broad structure at 65.6° is assigned to [322] cubic plane. A shoulder at 25.3° and a hump at 28.2° identified at the earlier stages of film growth are assigned to [110] and [101] hexagonal planes of CdS crystal respectively. We observe that the peak at 26.6° grows at the expense of the clear structures at between 25° and 35° for thin film deposited for 150 min. The peak structures assigned to H[100] and H[101] disappear for later stages of thin film growth. However the persistence of C[322] suggests a crystal growth which favors cubic phase with the peak at 26.5° assigned to C[111].

Fig. 14.8 XRD pattern for CBD CdS thin films of different thicknesses

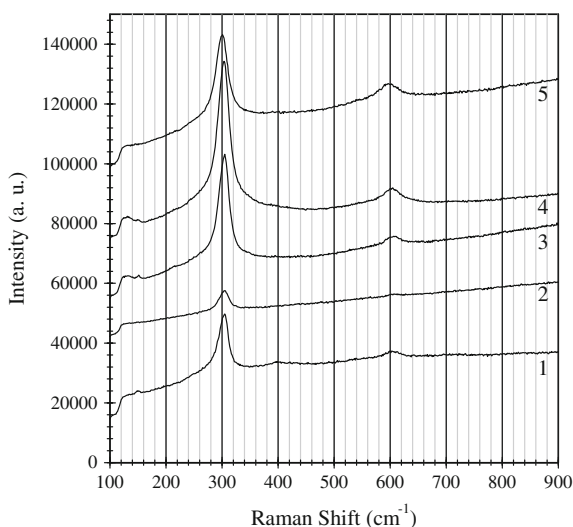


14.5.1.5 Raman Spectroscopy

Resonant Raman scattering measurements were carried out at room temperature on Renishaw 250 mm focal length in Via Reflex Spectrometer system using Rayleigh line rejection filter allowing Raman spectrum to 100 cm^{-1} from a laser line. An accuracy of $\pm 0.1\text{ cm}^{-1}$ was estimated for the spectral region studied. The Raman spectra are shown in Fig. 14.9. The Raman spectra are excited by a cw laser line at 532 nm and a total power of 3 mW within a spot size of $1\text{ }\mu\text{m}$ on the sample. Because the x-ray pattern in Fig. 14.8 displays quite amorphous background, the Raman spectra were collected from crystalline regions in the substrate. The spectra display a sharp peak at between $298\text{--}305\text{ cm}^{-1}$ attributed to zone center longitudinal optical (LO) mode of cubic phonons and hexagonal $A_1(\text{LO}) + E_1(\text{LO})$ phonons and overtones at $592\text{--}600\text{ cm}^{-1}$ [46]. We assume that the broad structure at the low frequency side of the most intense mode for the film deposited for 10 min is possibly originated from a cubic phonon mode based on a previous study by O. Z-Angel et al. [46] who reported that $E_1(\text{TO})$ mode of cubic structure at 240 cm^{-1} is shifted to 277 cm^{-1} as a result of cubic to hexagonal phase transformation. A Lorentzian fit which is centered at 278 cm^{-1} is consistent with their proposal. It is also apparent that the most intense mode is asymmetrically broadened toward the lower frequency which can be fitted by a Lorentzian profile centered at $\sim 297\text{ cm}^{-1}$ which might be due to disordered activated zone edge optical phonons as reported by previous researchers [47]. A broad peak at around 400 cm^{-1} which appears for the thinnest film is due presumably to Si–O vibrational modes from glass substrate (Fig. 14.9).

The shift in LO_1 peak position is related to grain radius by digitizing the calculated curve between LO_1 mode and grain radius at $295\text{ }^\circ\text{C}$ in [48]. We find a relation for the radius dependent vibrational frequency: $\omega(R)\text{ (cm}^{-1}\text{)} = 303.44 - 20.57/R^2\text{ (nm}^2\text{)}$.

Fig. 14.9 Raman spectra for CBD CdS thin films of different thicknesses



14.5.2 Modeling of the Urbach Tail

The defect levels below the band gap energy broaden due mainly to thermal vibrations and phonon-electron interactions. The long wavelength absorption tail is explained by the excitation of a low energy electron to conduction band through tunneling the potential energy incline created by electric micro fields [49]. The schematic energy band diagram is shown in Fig. 14.10 without field (a) and with field (b). The potential well of depth U_o represents a trap level over which an excited charge from the valence band tunnels through the classically forbidden region II as shown in (b) under electric field. The exponential tail below the fundamental absorption edge (E_g) related to the degree of structural disorder is determined by Urbach rule. We follow the procedure as described by Liboff [50] and derive transmission probability for an electron to tunnel through a potential barrier under electric field. Micro-field (F) of defects employing basic quantum mechanics [50]. In the WKB approximation the wave function for the charged particles is approximated by modified plane wave form:

$$\varphi(x) = Ae^{iS(x)} \quad (5.1)$$

where $S(x) = S_0(x) + \hbar S_1(x) + \frac{1}{2}\hbar^2 S_2(x) + \dots$ is the WKB expansion. After substituting (5.1) in Schrödinger equation and keeping only the first term in the expansion, we obtain the WKB solution in region III [50],

$$\varphi_{III}(x) = \frac{A}{\sqrt{k}} \exp \left[i \int_{x_2}^x k dx - \frac{\pi}{4} \right] \quad (5.2)$$

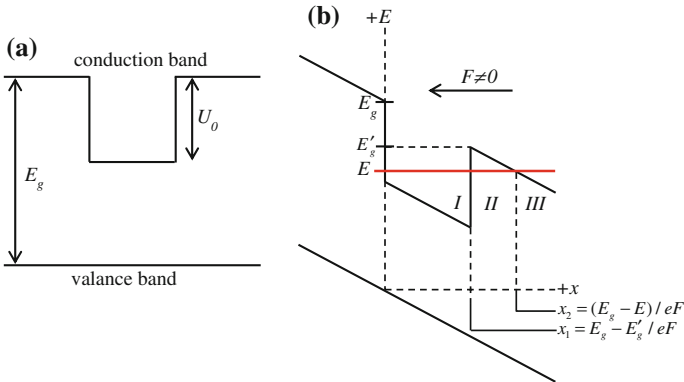


Fig. 14.10 The schematic energy band diagram with no field (a) with field (b) applied

which is related to φ_I over φ_{II} by connecting formulas,

$$\varphi_I(x) = \frac{C}{\sqrt{k}} \exp \left[i \int_{x_1}^x k' dx - \frac{\pi}{4} \right] \quad (5.3)$$

where $r = \exp \left[\int_{x_1}^{x_2} K dx \right]$ and k or $k' = \sqrt{2m(E - U)\hbar^2}$ and $K = \sqrt{2m(U - E)\hbar^2}$

are the magnitudes of wave vectors in regions I ($U = -eFx + E_g - U_0 < E$) or III ($U = -eFx + E_g < E$) and in region II ($U = -eFx + E_g > E$) respectively. The transmission probability (T) which is the ratio of transmitted (J_{trans}) to incident (J_{inc}) current densities in regions III and I respectively is calculated as $T = r^{-2}$ where $J = \frac{\hbar}{2mi} \left(\varphi^* \frac{d\varphi}{dx} - \varphi \frac{d\varphi^*}{dx} \right)$ [50]. A photon of energy $E = h\nu$ below the shifted band gap E'_g will be absorbed depending on the transmission probability and therefore we propose that $\alpha \approx T$ based on the earlier work by Dow and Redfield [51] and calculate the absorption coefficient by integrating the argument of r from $x_1 = (E'_g - E)/eF$ to $x_2 = (E_g - E)/eF$,

$$\alpha(E, F) = C_1 \exp \left[-C_2 \frac{(E'_g - E)^{3/2}}{F} \right] \quad (5.4)$$

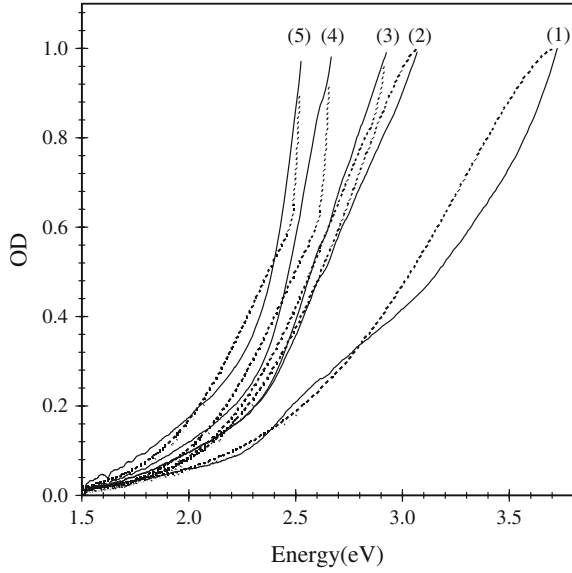
where C_1 is a constant, $C_2 = \frac{3}{4e} \sqrt{2\mu/\hbar^2}$ and E_u is Urbach energy. In Fig. 14.11, the calculated optical tail densities are plotted against photon energies for different thickness films along with experimental tail optical densities. As the electric micro-field increases, the absorption tail width related to Urbach energy increases. We model each optical absorption tail for different thickness films and find best average electric micro-field values. In Fig. 14.12, we plot electric micro-field as a function of deposition time. We see that as the deposition time increases, electric micro-field strength decreases. As thickness of the film of $\sim 2 \mu\text{m}$ at 16 min of deposition time increases to $\sim 2.5 \mu\text{m}$ at the end of 40 min, the average electric micro-field decreases. Average electric micro-field can be described by a power law decay function given by the expression: $F(\text{kV/mm}) = 3.4 \times [t(\text{min})]^{-0.32}$.

We relate electric micro-field to Urbach energy to explore the relation between disorder and internal electric field induced by crystalline disorder. The optical density is given by,

$$OD = OD_0 \exp \left[\frac{h\nu - E_g}{E_u} \right] \quad (5.5)$$

here OD_0 is a constant. The Urbach tail width represented by the Urbach energy (E_u) is calculated from the slope of least-squares linear fit to the straight portion of $\ln OD$ against the $h\nu$ curve [52]. The Urbach energy is plotted against the average

Fig. 14.11 Experimental and calculated (*dashed curve*) optical absorption tail spectra below band gap energy for different thickness thin films



electric micro-field in Fig. 14.13. The right vertical axis is the unitless structural disorder parameter (X) calculated by the equation [53],

$$E_u = \frac{E_p}{2\sigma_0} \left[X + \coth\left(\frac{E_p}{2k_{BT}}\right) \right] \quad (5.6)$$

where E_p of 25 meV is the phonon energy, σ_0 of 1.19 is a material-dependent parameter and X is a measure of structural disorder. The straight line through the data points is a linear fit given by $y = 0.16x + 254$ where $y = E_u$ (meV) or X (%) and $x = F$ (kV/mm).

14.5.3 The Progression of Strain with Deposition Time

The XRD diffraction lines broaden due to finite crystalline size, strain and instrumental error. The size broadening is given by Debye-Scherrer relation

$$\Delta(2\theta_{hkl})_{size} = \frac{K\lambda}{2R_{ave} \cos \theta_{hkl}} \quad (5.7)$$

where R_{ave} is the average radius of grains. The strain broadening and the strain coefficient are given by,

Fig. 14.12 Electric micro-field plotted against deposition time. The curve is a power law fit the data

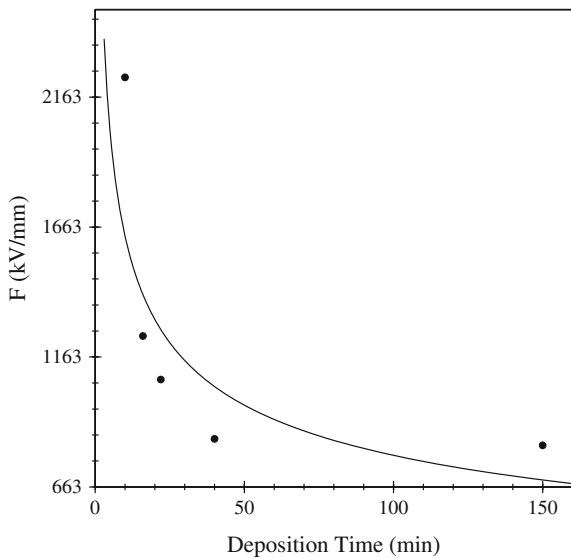
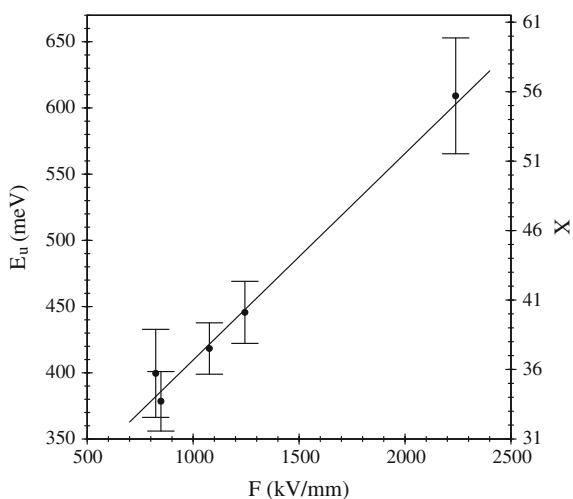


Fig. 14.13 Average electric micro-field against Urbach energy and disorder. The straight line is a linear fit to the data with slope of 0.16



$$\Delta(2\theta_{hkl})_{strain} = 2 \epsilon \tan \theta_{hkl} \quad (5.8a)$$

$$\epsilon = - \frac{d_{hkl} - d_{hkl}^{bulk}}{d_{hkl}^{bulk}} \quad (5.8b)$$

We approximate the angular dependent instrumental error function through a scan on the Si powder as $\Delta(2\theta_{hkl})_{inst}(\text{mrad}) = -9 \times \tan \theta_{hkl} + 4.5/\cos \theta_{hkl}$ which estimates that the maximum instrumental broadening could be only 20 % of the least FWHM.

Table 14.3 The summary of the results for different thickness CdS thin films

Deposition time (min)	Urbach energy (meV)	Grain radius (nm)		F (kV/mm)	Strain $\times 10^{-3}$		Raman shift (cm ⁻¹)	E _g (eV)	Thickness (μm) SEM
		ABS	Raman		XRD	Raman			
10	605	1.21	1.93	2239	9.55	2.24	297.90	3.72	–
16	445	1.71	3.47	1244	7.73	0.19	301.74	3.08	~2.0
22	417	1.91	3.88	1077	4.80	0.09	302.08	2.93	–
40	377	2.60	3.24	848	3.52	0.33	301.49	2.67	~2.5
150	396	3.59	1.96	823	3.48	3.13	298.11	2.53	–

We related peak position for C[111] for each sample to the interplanar distance d_{111} by employing Bragg Law for cubic structure with a bulk lattice constant of $a = 0.582$ nm. The progression of strain values calculated by employing XRD data and (5.8b) are listed in Table 14.3.

Raman spectroscopy was employed to assess the relative weigh of the result for the strain obtained by XRD. Vibrational modes in Raman spectra are fitted by a Lorentzian line profile for each vibrational mode and a cubic background function. The best fitting parameters for the LO_1 peak position of P_i are recorded to calculate the strain coefficient which is related to the vibrational frequency by [54],

$$\frac{\Delta a}{a} = \frac{\left(\frac{P_i - \Delta\omega_i^c}{LO_i^B}\right)^{-\frac{1}{\gamma}} - 1}{3} \tag{5.9}$$

where LO_i^B is the bulk frequency of LO phonon, $\Delta\omega_i^c$ is the red-shift due to phonon confinement and γ is the Grüneisen parameter of 1.51 for bulk CdS [55]. Previous studies show that the shift in Raman frequency due to phonon confinement is negligible and can be ignored compared to the shift due to strain. The strain values calculated by using LO_1 peak positions in Fig. 14.9 are tabulated in Table 14.3.

14.5.4 Section Conclusion

Different thickness CdS thin films were grown on glass by CBD. The red shift observed in asymptotic optical absorption edge is used to calculate the average grain radii from 1.2 to 3.6 nm with the help of effective mass model under quantum confinement of charged carriers. The optical absorption below the band gap energy is simulated by a simple quantum mechanical model which presumes the WKB solution to Schrödinger equation under a disorder induced micro-electric field. The simulation shows that the micro-electric field decreases with grain size. The exponential tail width proportional to Urbach energy varies nearly linearly with micro-field. The results extracted directly or through calculation from the optical absorption, XRD, SEM and Raman data are summarized in Table 14.3.

The red shift in Raman LO₁ vibrational modes and the shift in C(111) diffraction peak position in XRD spectra yield a compressive strain on the order of power of 3. We conclude that as the deposition time increases from 10 to 150 min, the average grain radius changes by 2 nm, Urbach energy and the electric micro-field presumably induced by crystalline disorder decrease approximately from 600 to 400 meV and 2240–820 kV/mm respectively. The shift in XRD pattern shows that the compressive strain decreases with growth from 9.55×10^{-3} to 3.48×10^{-3} . The Raman LO₁ vibrational mode displays an increase up to 22 min of deposition time and then a decrease. There might be two contributions to observed shift in LO phonon mode: strain and phonon confinement. Phonon confinement leads to a decrease in frequency because the zone center LO phonon frequency of 303.44 cm^{-1} for bulk CdS is at a maximum. The shift due to strain might be positive or negative. At the earlier stages of grain growth when the size is 1.2 nm in radius, we observe a decrease of about 5 cm^{-1} which is too high to be due to phonon confinement since Raman frequency is red shifted by only 1 cm^{-1} as the nanocrystalline radius change between 1 and 3 nm as reported in [56]. As shown in Table 14.3 the strain calculated by Grüneisen model decreases with deposition time up to 22 min and increases later on which might be due to a phase transition to cubic structure from a mixed phase of hexagonal and cubic structures at the earlier stages of film growth as is evident from XRD pattern.

14.6 In Situ Low Temperature Optical Absorption Spectroscopy

The low temperature optical transmission measurements were conducted on one CdSe doped glass sample by employing a custom-made liquid nitrogen optical chamber. A diagram of the sample geometry is shown in Fig. 14.14. The temperature was monitored with a thermo-couple positioned within 2 cm of the center of the crystal fixed on a sample holder in the chamber. The light from a 50 W tungsten lamp, which was powered by a constant current source, was collected by a lens, passed through the quartz window of the chamber and then the sample. The transmitted light was collected by another lens behind the quartz window at the

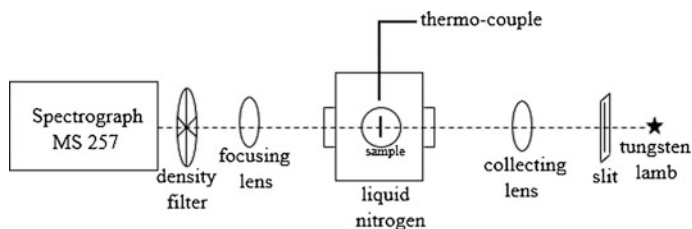
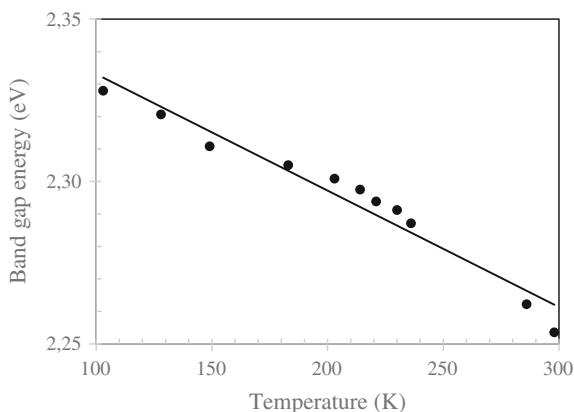


Fig. 14.14 The experimental setup for in situ low temperature optical absorption spectroscopy

Fig. 14.15 The energetic position of asymptotic optical absorption edge plotted against temperature. The line through the data is a linear fit with a slope of -4×10^{-4} eV/K



opposite side of the chamber and focused on the entrance slit of a $\frac{1}{4}$ m spectrograph. The transmitted spectra were measured as a function of temperature at between 104 and 253 K. The band gap energy is plotted as a function of temperature in Fig. 14.15. The line is a least-squares straight line fit whose slope is $\frac{dE}{dT} = -4 \times 10^{-4}$ eV/K.

Acknowledgements This work has been supported by Yıldız Technical University Scientific Research Project Coordination under project nos, 2012-01-01-KAP09 and 2012-01-01-KAP03 2011-01-01-DOP01. XRD and SEM were conducted at TÜBİTAK MRS Industrial services. We thank Dr. Barış Yağcı at Surface Science and Technology Center (KUYTAM), Koç University for conducting Raman measurement.

References

1. V. Wood, V. Bulović, Colloidal quantum dot light-emitting devices. *Nano Rev.* **1**(5202), 1–7 (2010)
2. L. Liu, Q. Peng, Y. Li, An effective oxidation route to blue emission CdSe quantum dots. *Inorg. Chem.* **47**(8), 3182–3187 (2008)
3. J. Yoon, W. Chae, S. Im, Y. Kim, Mild synthesis of ultra-small CdSe quantum dots in ethylenediamine solution. *Mater. Lett.* **59**, 1430–1433 (2005)
4. C. Mell Donegá, P. Liljeroth, D. Vanmaekelbergh, Physicochemical evaluation of the hot-injection method, a synthesis route for monodisperse nanocrystals. *Small* **12**, 1152–1116 (2005)
5. A.K. Arora, M. Rajalakshmi, T.R. Ravindran, Phonon confinement in nanostructured materials. *Encycl. Nanosci. Nanotechnol.* **8**, 499–512 (2004)
6. H. Matsuo, T. Soma, Lattice dynamics of tetrahedral compounds with the local Heine-Abarenkov model potential. *Phys. Stat. Sol. (b)* **124**, 37–44 (1984)
7. F. Wu, J.W. Lewis, D.S. Kliger, J.Z. Zhang, Unusual excitation intensity dependence of fluorescence of CdTe nanoparticles. *J. Chem. Phys.* **118**, 12 (2003)
8. M.H. Yükselici, Ç. Allahverdi, Size-dependent photo-induced shift of the first exciton band in CdTe quantum dots in glass prepared by a two-stage heat-treatment process. *J. Lumin.* **128**(3), 537 (2008)

9. V.M. Nikale, S.S. Shinde, C.H. Bhosale, K.Y. Rajpure, Physical properties of spray deposited CdTe thin films: PEC performance. *J. Semicond.* **32**, 033001–1 (2011)
10. H.R. Moutinho, F.S. Hasoon, F. Abulfotuh, L.L. Kazmerski, Investigation of polycrystalline CdTe thin films deposited by physical vapor deposition, close-spaced sublimation, and sputtering. *J. Vac. Sci. Technol., A* **13**, 2877 (1995)
11. Ç. Allahverdi, M.H. Yükselici, Temperature dependence of absorption band edge of CdTe nanocrystals in glass. *New J. Phys.* **10**, 103029 (2008)
12. G. Gordillo, J.M. Florez, L.C. Hernandez, Preparation and characterization of CdTe thin films deposited by CSS. *Sol. Energy Mater. Sol. Cells* **37**, 273–281 (1995)
13. M. H. Yükselici, Ç. Allahverdi, Solid-phase precipitation of CdTe nanoparticles in glass, *Phys. Stat. Sol. (b)* **236**(3), 649 (2003)
14. S. Lalitha, R. Sathyamoorthy, S. Senthilarasu, A. Subbarayan, K. Natarajan, Characterization of CdTe thin film—dependence of structural and optical properties on temperature and thickness. *Sol. Energy Mater. Sol. Cells* **82**, 187–199 (2004)
15. S. Wageh, A.A. Higazy, M.A. Algrade, optical properties and activation energy of a novel system of CdTe nanoparticles embedded in phosphate glass matrix. *J. Modern Phys.* 913–921 (2011)
16. P.T.C. Freire, M.A. Araujo Silva, V.C.S. Reynoso, A.R. Vaz, V. Lemos, Pressure Raman scattering of CdTe quantum dots. *Phys. Rev. B* **55**(11), 6743 (1997)
17. S.S. Islam, S. Rath, K.P. Jain, S.C. Abbi, Forbidden one-LO-phonon resonant Raman scattering and multiphonon scattering in pure CdTe crystals. *Phys. Rev. B* **46**(8), 4982 (1997)
18. V.S. Vinogradov, G. Karczewski, I.V. Kucherenko, N.N. Mel'nik, P. Fernandez, Raman spectra of structures with CdTe-, ZnTe-, and CdSe-based quantum dots and their relation to the fabrication technology. *Phys. Solid State* **50**(1), 164–167 (2008)
19. V. Dzhagan, I. Lokteva, C. Himcinschi, X. Jin, J. Kolny-Olesiak, D.R.T. Zahn, Phonon Raman spectra of colloidal CdTe nanocrystals: effect of size, non-stoichiometry and ligand exchange. *Nanoscale Res. Lett.* **6**, 79 (2011)
20. A.G. Rolo, M.I. Vasilevskiy, N.P. Gaponik, A.L. Rogach, M.J.M. Gomes, Confined optical vibrations in CdTe quantum dots and clusters. *Phys. Stat. Sol. (b)* **229**(1), 433–437 (2002)
21. M.I. Vasilevskiy, A.G. Rolo, M.J.M. Gomes, O.V. Vikhrova, C. Ricolleau, Impact of disorder on optical phonons confined in CdS nano-crystallites embedded in a SiO₂ matrix. *J. Phys.: Condens. Matter* **13**, 3491–3509 (2001)
22. B.C. Ömür, A. Aşıkoğlu, Ç. Allahverdi, M.H. Yükselici, CdS_xSe_{1-x} quantum dots studied through optical absorption, steady state photoluminescence and resonant Raman spectroscopy. *J. Mater. Sci.* **45**, 112–117 (2010)
23. Ç. Allahverdi, M.H. Yükselici, R. Turan, A. Seyhan, Photoluminescence spectroscopy in the study of growth of CdSe_xS_{1-x} nanocrystals in glass. *Semicond. Sci. Technol.* **19**, 1005 (2004)
24. N.F. Borrelli, D.W. Hall, H.J. Holland, D.W. Smith, Quantum confinement effects of semiconducting microcrystallites in glass. *J. Appl. Phys.* **61**, 5399 (1987)
25. F. Hache, M.C. Kellin, D. Ricard, C. Flytzanis, Photoluminescence study of Schott commercial and experimental CdS_{Se}-doped glasses: observation of surface states. *J. Opt. Soc. Am. B* **8**, 1802 (1991)
26. G. Mei, A photoluminescence study of CdSe_xS_{1-x} semiconductor quantum dots. *J. Phys.: Condens. Matter* **4**, 7521 (1992)
27. W.S.O. Rodden, C.N. Ironside, C.M.S. Torres, A study of the growth of CdSe_xS_{1-x} crystallites within a glass matrix. *Semicond. Sci. Technol.* **9**, 1839 (1994)
28. T.R. Ravindran, A.K. Arora, B. Balamurugan, B.R. Mehta, Inhomogeneous broadening in the photoluminescence spectrum of CdS nanoparticles. *Nanostruct. Mater.* **11**, 603–609 (1999)
29. F. Henneberger, J. Puls, C. Spiegelberg, A. Schülzgen, H. Rossman, V. Jungnickel, A.I. Ekimov, Optical and electro-optical properties of II-VI quantum dots. *Semicond. Sci. Technol.* **16**, A41 (1991)
30. A.D. Compaan, The status of and challenges in CdTe thin-film solar-cell technology. *Mater. Res. Soc. Symp. Proc.* **808**, 545–555 (2004)

31. J. Britt, C. Ferekides, Thin-film CdS/CdTe solar cell with 15.8 % efficiency. *Appl. Phys. Lett.* **62**, 2851–2852 (1993)
32. K.L. Chopra, P.D. Paulson, V. Dutta, Thin-film solar cells: an overview. *Prog. Photovolt. Res. Appl.* **12**, 69–92 (2004)
33. S. Chun, Y. Jung, J. Kim, D. Kim, The analysis of CdS thin film at the processes of manufacturing CdS/CdTe solar cells. *J. Cryst. Growth* **326**, 152–156 (2011)
34. C. Gretener, J. Perrenoud, L. Kranz, L. Kneer, R. Schmitt, S. Buecheler, A.N. Tiwari, CdTe/CdS thin film solar cells grown in substrate configuration. *Prog. Photovolt. Res. Appl.* (2012). doi:[10.1002/pip.2233](https://doi.org/10.1002/pip.2233)
35. C. Sahu, S.N. Sahu, Preparation of CdS semiconductor thin films by a solution growth technique. *Thin Solid Films* **235**, 17–19 (1993)
36. R. Ortega-Borges, D. Lincot, Mechanism of chemical bath deposition of cadmium sulfide thin films in the ammonia-thiourea system in situ kinetic study and modelization. *J. Electrochem. Soc.* **140**, 3464–3473 (1993)
37. J.M. Doña, J. Herrero, Chemical bath deposition of CdS thin films: an approach to the chemical mechanism through study of the film microstructure. *J. Electrochem. Soc.* **144**, 4081–4091 (1997)
38. C. Guill'en, M. A. Mart'inez, J. Herrero, Accurate control of thin film CdS growth process by adjusting the chemical bath deposition parameters. *Thin Solid Films* **335**, 37–42 (1998)
39. R.S. Mane, C.D. Lokhande, Chemical deposition method for metal chalcogenide thin films. *Mater. Chem. Phys.* **65**, 1–31 (2000)
40. J. Herrero, M.T. Gutiérrez, J.M. Doña, M.A. Martínez, A.M. Chaparro, R. Bayon, Photovoltaic windows by chemical bath deposition. *Thin Solid Films* **361–362**, 28–33 (2000)
41. R. Venugopal, P.-I. Lin, C.-C. Liu, Y.-T. Chen, Surface-enhanced Raman scattering and polarized photoluminescence from catalytically grown CdSe nanobelts and sheets. *J. Am. Chem. Soc.* **127**, 11262–11268 (2005)
42. E.S.F. Neto, N.O. Dantas, S.W. da Silva, P.C. Morais, M.A. Pereira da Silva, Confirming the lattice contraction in CdSe nanocrystals grown in a glass matrix by Raman scattering. *J. Raman Spectrosc.* **41**, 1302–1305 (2010)
43. A. Aşıkoğlu, M.H. Yükselici, Evolution of the energy band structure in chemical-bath-deposited CdS thin films studied by optical absorption spectroscopy. *Semicond. Sci. Technol.* **26**, 055012 (2011)
44. J.I. Pankove, *Optical Processes in Semiconductors* (Dover, New York, 1971)
45. F. Henneberger, J. Puls, A. Ch Spiegelberg, H. Rossman Schulzgen, V. Jungnickel, A.I. Ekimov, Optical and electro-optical properties of II-VI quantum dots. *Semicond. Sci. Technol.* **6**, A41–A50 (1991)
46. O. Zelya-Angel, Raman studies in CdS thin films in the evolution from cubic to hexagonal phase. *Solid State Commun.* **104**, 161–166 (1997)
47. S. Mishra, A. Ingale, U.N. Roy, A. Gupta, Study of annealing-induced changes in CdS thin films using X-ray diffraction and Raman spectroscopy. *Thin Solid Films* **516**, 91–98 (2007)
48. E.S. Freitas Neto, N.O. Dantas, S.W. da Silva, P.C. Morais, M.A. Pereira-da-Silva, A.J.D. Moreno, V. López-Richard, G.E. Marques, C. Trallero-Giner, Temperature-dependent Raman study of thermal parameters in CdS quantum dots. *Nanotechnology* **23**, 1–8 (2012)
49. H. Hartmut, S.W. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductors*, 4th edn. (World Scientific, Singapore, 2004)
50. R.L. Liboff, *Introductory Quantum Mechanics*, 2nd edn. (Addison-Wesley, San Francisco, 1980)
51. J.D. Dow, D. Redfield, toward a unified theory of Urbach's rule and exponential absorption edges. *Phys. Rev. B* **5**, 594–610 (1972)
52. M.V. Kurik, Urbach Rule. *Phys. Stat. Sol. (a)* **8**, 9–45 (1971)
53. G.D. Cody, T. Tiedje, B. Abeles, B. Brooks, Y. Goldstein, Disorder and the optical absorption edge of hydrogenated amorphous silicon. *J. Phys. Colloques* **C4**, 301–304 (1981)

54. V. Spagnolo, G. Scamarcio, M. Lugara, G.C. Righini, Raman scattering in $\text{CdTe}_{1-x}\text{Se}_x$ and $\text{CdS}_{1-x}\text{Se}_x$ nanocrystals embedded in glass. *Superlattices Microstruct.* **16**, 51–54 (1994)
55. Landolt-Borstein, *Numerical Data and Functional Relationships in Science and Technology*, vol. III/41b. (Springer, Berlin, 1999)
56. X.S. Zhao, J. Shroeder, P. Persans, T. Bilodeau, Resonant-Raman-scattering and photoluminescence studies in glass-composite and colloidal CdS. *Phys. Rev. B* **43**, 12580–12589 (1991)

Chapter 15

One-Dimensional Nano-structured Solar Cells

H. Karaağaç, E. Peksu, E.U. Arici and M. Saif Islam

Abstract The solar light harvesting has long been regarded as promising way to meet the increasing world's annual energy consumption as well as the solution to prevent the detrimental long-term effect of carbon-monoxide emission released by fossil fuel sources. Due to the high cost of today's conventional PV technology, however, it is not possible to compete with the energy supplied from fossil fuel sources. The use of one-dimensional nanostructures, including nanowires (NWs), nanorods (NRs), nanopillars (NPs) and nanotubes (NTs) in solar cells with different device architectures (e.g. axial, radial, and nanorod/nanowire array embedded in a thin film) provides peculiar and fascinating advantages over single-crystalline and thin film based solar cells in terms of power conversion efficiency and manufacturing cost due to their large surface/interface area, the ability to grow single-crystalline nanowires on inexpensive substrates without resorting to complex epitaxial routes, single-crystalline structure and light trapping function. In this chapter, we review the recent studies conducted on nanowire/nanorod arrays based solar cells with different device architectures for the realization of high-efficiency solar cells at an economically viable cost.

H. Karaağaç (✉) · E. Peksu

Department of Physics, Istanbul Technical University, 34469 Maslak, Istanbul, Turkey
e-mail: karaagacha@itu.edu.tr

E.U. Arici

Energy Institute, Istanbul Technical University, 34469 Maslak, Istanbul, Turkey
e-mail: aricibogner@itu.edu.tr

M. Saif Islam

Department of Electrical and Computer Engineering, University of California at Davis, Davis, CA 95616, USA
e-mail: sislam@ucdavis.edu

15.1 Introduction

Today's world annual energy consumption has been estimated to be around 15 TWYs and it has been predicted to be around 30 TWYs by 2050 [1]. To meet this huge energy requirement, harnessing energy from the sun has long been regarded as a promising solution due to its abundance and free from contaminations and carbon emissions [2]. The direct conversion of solar energy is realized through solar cells, which is based on photovoltaic (PV) effect [3]. Due to the high cost of today's conventional PV technology it is not possible to compete with the energy supplied from fossil fuel sources. However, in recent years, there has been an increasing interest in reducing the price, or cost per watt, of solar cells. In particular, the use of nanostructures, defined as structures having at least one dimension in the range of 1–100 nm, including nanowires (NWs), quantum wells (QWs), and quantum dots (QDs), in solar cells is identified as promising approach to generate inexpensive photovoltaic electricity [4]. The use of nanostructures in solar cells offers many advantages over their conventional planer counterparts owing to their unique properties such as implanting new physical mechanisms, tailoring the material properties and low fabrication cost. In particular, one-dimensional (1D) nanostructures including nanowires, nanorods (NRs), nanopillars (NPs) and nanotubes (NTs) provide peculiar and fascinating advantages over thin films in various opto-electronic device applications such as solar cells and sensors, arising from their large surface/interface area, single-crystalline structure and light trapping function [5].

The objective of this chapter is to review recent research activities on 1-D nanostructures based solar cells. First section of this chapter deals with the most widely preferred 1D nanostructures based solar cell architectures/models. Next section focuses on the most commonly employed growth techniques for the synthesis of 1D nanostructures. Third section includes well-known materials (Silicon, Zinc-Oxide, Titanium-dioxide, and carbon) based 1D nanostructures based solar cells. The final section concludes with prospective on 1D nanostructures and their solar cell application.

15.2 One-Dimensional Nanostructures Based Solar Cell Architectures

In the past years, a variety of concepts have been proposed for the realization of high-efficiency solar cells at an economically viable cost [6, 7]. Utilization of one-dimensional (1D) nanowire/nanorods array for the construction of three-dimensional (3D) structures is among the widely preferred approach to fabricate high-efficiency solar cells at lower cost. In recent decades, various photovoltaic (PV) device architectures based on nanowires/nanorods/nanopillars have been developed to construct 3D solar cells. In particular, the device configuration based

on nanowire/nanorods array with radial, axial, and nanowire/nanorods array embedded in thin film junctions are the most widely preferred types of PV devices [5, 8, 9], as depicted in Fig. 15.1. Each of these device architectures has its own advantages and drawbacks [5]. All types of these structures offer many potential benefits over the traditional planar thin-film or wafer-based solar cells in terms of cost, effective charge separation, strain relaxation and some electrical and optical properties [10]. Of the aforementioned solar cell architectures, core-shell (radial) configuration takes advantages of all these effects, such as high surface/volume ratio, reduced surface reflection and enhanced carrier collection efficiency. The core-shell configuration provides not only tuning the optical properties of the structure by adjusting the size and pitch of nanowire array but also the possibility to improve the carrier collection efficiency by varying the core-shell radii [5]. However, in this type device architecture a conformal coating requirement still remains a problem [11]. For the nanowire/nanorods array embedded in a thin film configuration, although there is an enhanced carrier collection efficiency, its optical properties are almost similar to thin film or wafer-based solar cell architecture owing to its planar top surface [11]. Of the 1D nanostructures based solar cell models, PV devices based on axial junction configuration have attracted less attention due to the limitation of the geometry in effective charge collections and the challenges in material synthesis [12]. It is, therefore, critical to fabricate sharp axial p-n junctions and well controlled doping level to achieve effective charge separation and charge transportation to counter electrodes with such a device configuration [12].

The past decade has seen increasingly rapid advances in the use of nanowires/nanorods for the realization of third-generation solar cells with either all-inorganic or inorganic-organic hybrid material combinations based on one of the aforementioned solar cell device architectures [13, 14]. In 2005, the first Si NWs based core-shell solar cell with a 14.5 % efficiency (theoretically) was proposed [15]. On the other hand, the efficiency of first Si NW solar cell having core-shell architecture with thin amorphous n-Si on p-Si NWs was only ~ 0.1 % [16]. The low efficiency of the constructed device was attributed to junction quality and impurities, which are assumed to be key parameters having direct effect on determining the efficiency of this type device structure. Another study based on Si NW core-shell solar cell was reported by Garnett and Yang using two-step procedures [9]. N-type Si NW array was derived from Si wafer using chemical etching, which

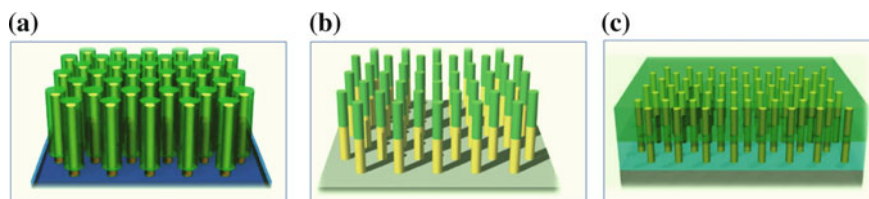


Fig. 15.1 One-dimensional nanostructures based solar cell architectures: **a** core-shell, **b** axial, **c** nanowire/nanorod array embedded in thin film junctions [5]

was sequentially decorated with a p-type Si layer using chemical vapor deposition. The solar cell parameters of the device exhibited a J_{sc} (short-current density) of 4.28 mA/cm^2 , V_{oc} (open-circuit voltage) of 0.29 V and FF (fill-factor) of 0.33 . Based on these parameters, the power conversion efficiency (PCE) was calculated to be 0.46% . The measured low PCE with this device geometry was attributed to the existence of severe interfacial recombinations and high series resistance, suggested to be overcome by surface passivation of Si NWs and reducing the bulk resistivity of p-Si cap layer.

There are also a number of studies in literature on nanowire array embedded in thin film type solar cells. For instance, Yang et al. demonstrated an all-oxide solar cell, in which n-type ZnO NW array covered by a p-type Cu_2O nanoparticles was utilized to realize environmentally benign stable solar cells [17]. It was revealed that the morphology, the thickness and the phase of Cu_2O nanoparticles were the key factors determining the performance of solar cell. It was also realized that when an immediate blocking layer introduced between the electrodes (having direct contact with NWs) and the absorber layer, there was a dramatic increase in the overall efficiency of the solar cell. For instance, a 10 nm TiO_2 blocking layer between the transparent electrode and Cu_2O resulted in 50 times higher efficiency (0.055%) with respect to the device configuration without the blocking layer. In another study, Fan et al. [18] demonstrated a solar cell consisting of n-CdS nanopillar array grown by CVD in a porous anodic alumina membrane (AAM) embedded in p-CdTe polycrystalline thin film by CVD so as to enhance light absorption and carrier collection efficiency. The performance of solar cell was tested under different illumination conditions, ranging from 17 to 100 mWcm^{-2} . For instance, under A.M.1.5 G illumination condition, the n-CdS/p-CdTe nanopillars solar cell exhibited a power conversion efficiency of $\sim 6 \%$, which is known to be higher than values that have been reported so far for the solar cell devices utilizing the same concepts such as the orthogonalizing the direction of light absorption and photo-generated carrier collection. Results of this investigation suggested that the power conversion efficiency achieved, $\sim 6 \%$, could be improved by applying several device-fabrication, contact issues, and nanopillars-geometry related optimization processes including the adjustment of the inter-pillar spacing of CdS pillar array, the transparency and resistances of contacts, coating nanopillars with a antireflection layer, and tuning the penetration depth of CdS nanopillars in the polycrystalline CdTe thin film layer. Additionally, in order to realize a flexible solar cell, CdS/CdTe nanopillars solar cell was constructed in a plastic layer (Polydimethylsiloxane, PDMS), which exhibited almost the same power conversion efficiency that obtained for the solar cells constructed on the rigid substrates. Furthermore, under different bending conditions device showed a little change in the cell performance. In a different study, single nanowire solar cells consisting of GaAs nanowires in radial p-n configuration were reported [8], which exhibited a 0.8% power conversion efficiency with $\text{FF} \sim 0.26$. The same group also investigated numerical simulations of current-voltage characteristics for III-V nanowire core-shell architecture in order to find the optimum nanowire geometry for the realization of high-efficiency solar cells [19]. Their analysis revealed that a $\sim 30 \%$

power conversion efficiency could be achieved using the core-shell device architecture by implanting effective surface passivation as well as designing the p-n junction with top contact geometry and minimal tip height. In another major study, Colombo et al. reported a 4.5 % power conversion efficiency for a radial p-i-n diode nanowire solar cell based on GaAs [20]. In addition to CdS, Si, ZnO and GaAs, core-shell solar cells based on GaN nanopillars were also reported as heterostructured solar cell with power conversion efficiency of ~ 2.7 % under AM 1.5 G illumination [21].

There have been relatively few recent studies on solar cells based on axial device architecture due to the challenges in synthesis of materials and the limitations in the charge collection. Sivakov et al. constructed an axial architecture based solar cell composed of multicrystalline p^+-n-n^+ -doped Si layer using electroless etching technique [22]. To form axial geometry, first p^+ , n , and n^+ amorphous Si films were deposited alternately by electron-beam evaporation technique, which was followed by crystallization using a laser beam, and then chemical etching to synthesize Si NWs. Results demonstrated that Si NWs exhibited a significant reduction in reflectance in the wavelength range of 300–1000 nm, when compared with its planar counterpart. Power conversion efficiency for the device was calculated to be 1.4 % under standard measurement conditions (A.M. 1.5 illumination). In another major study, GaAs nanowire array based solar cells with p-i-n axial junctions were reported by Yao et al. [23]. The results of investigation revealed that GaAs nanowire array was highly absorptive when synthesized with a low filling ratio. In addition to this, it was found that thicker nanowire arrays without surface passivation treatment were favorable owing to the high surface recombination velocity. A power conversion efficiency of 7.58 % was achieved for the solar cell with p-i-n axial junction consisting of a junction depth of 100 nm and nanowire diameter of 320 nm, indicating that GaAs nanowire array with axial geometry is among the promising candidates for the realization of light-efficiency and low-cost solar cells, which are regarded as next generation photovoltaics.

15.3 Synthesis of One-Dimensional Nanostructures

In this section, the most widely employed fabrication methods for the synthesis of nanostructures will be reviewed. In particular, the synthesis of one-dimensional nanostructures based on top-down and bottom-up methods will be introduced.

In general, top-down and bottom-up methods are regarded as two different approaches that embody all the fabrication methods for the synthesis of 1-D nanostructures. Fabrication methods based on top-down approach includes the reduction of the size of bulk material to the nanoscale by using combined lithographic steps and etching process. It is well known fact that this approach was first employed in the microelectronic industry. However, following the advances in thin film deposition and lithography techniques this approach has been used to synthesize one-dimensional nanostructures as well. Although a great success has been

achieved in the fabrication of 1-D microstructures by using top-down approach, the limit in the size of features (width, diameter) that can be formed imposed by lithographic patterning (e.g., electron beam) has constrained its wide use in the last half century [24]. However, this goal can be realized by bottom-up approach, regarded as a more promising way for the synthesis of 1-D nanostructures due to its precisely controlled growth dynamics. In other words, this approach mainly based on the synthesis of 1-D nanostructures from some basic blocks such as atoms, molecules and clusters in a precisely controlled fashion.

Generally, the methods for the fabrication of 1-D nanostructures are presented under two different routes. These are solution- and vapor-based techniques. Here, several commonly employed fabrication techniques for the synthesis of 1-D nanostructures will be presented.

15.3.1 Solution-Based Synthesis of 1-D Nanostructures

The fabrication of one-dimensional (1-D) nanostructures along one specific orientation requires the break of growth symmetry that can be accomplished in a wide range of solution-based routes.

Hydrothermal technique is one of the most commonly used solution-based techniques. Among several other solution-based techniques, hydrothermal is of a special interest due to its many advantages such as its low growth temperature, a variety of control parameters for tuning properties, and allowing the input of various elements into the crystal at relatively low temperatures [25]. The synthesis process using hydrothermal technique involves the use of a mixture of precursor and reagent to either regulate or template crystal growth like amine, which is followed by a series of chemical reactions under different pressures and temperatures. So far, a wide variety of 1-D nanostructures, such as ZnO NWs, TiO₂ NRs and MnO₂ NWs, have been synthesized using this technique [26–28]. Hydrothermal technique is particularly preferred to synthesize vertically well-aligned ZnO NWs that can be employed for a wide variety of opto-electronic devices [29–36]. For instance, recently, we demonstrated a 3D photodetectors based on ZnO NWs grown by hydrothermal technique (Fig. 15.2) [37].

Some of solution-based techniques are particularly convenient for certain types of materials. For instance, metal assisted chemical etching (MACE) is widely used to derive Si NWs from single-crystalline silicon (Si) wafers [38]. The fabrication of Si NWs can also be accomplished by using a number of techniques, such as vapor-liquid-solid (VLS) growth, electrochemical, or lithography-related etching routes [39]. However, most of these techniques are requiring complicated expensive equipments with demand of high vacuum and temperature. On the other hand, MACE is a simple and a low-cost technique that can be employed for the synthesis of vertically well-oriented Si NWs. It offers several advantages over the aforementioned techniques including its ability to tune the physical parameters (e.g., diameter, length, and density) by adjusting growth parameters (e.g., growth

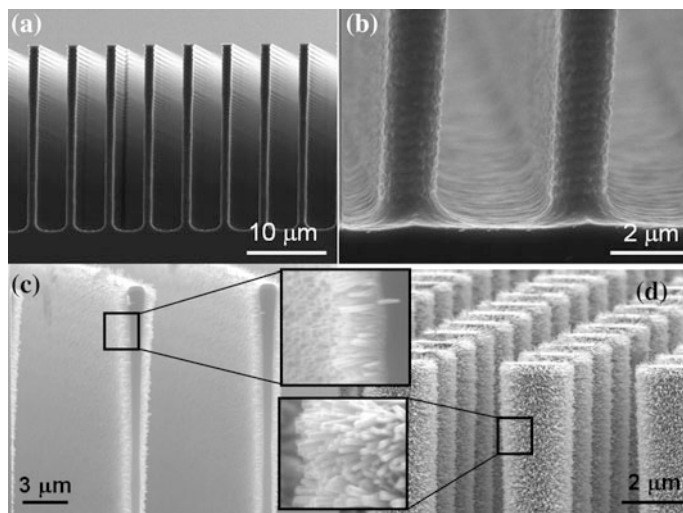


Fig. 15.2 SEM images of Si micro pillars/walls decorated with ZnO NWs grown by hydrothermal technique: **a** Si microwalls and **b** Si micropillars decorated with AZO seed layer (100 nm), ZnO NWs grown on **c** Si microwalls and **d** Si micropillars

temperature, precursor concentration and growth time), no requirement of expensive equipment and allowing mass producible fabrication of Si NWs. In a typical MACE process, a Si wafer first covered by a metal, which is subsequently immersed in a solution consisted of hydrofluoric (HF) acid and an oxidative agent [40]. Since the metal coated parts of Si wafer has a faster etching rate compared with uncoated parts, it leads to a selective etching as metal nanoparticles propagates through the Si wafer. Based on transferred metal (e.g., silver, gold, copper) pattern and the nature of used solution (AgNO_3/HF or $\text{H}_2\text{O}_2/\text{HF}$), Si nanostructures with various morphologies (e.g., cylinder, pillar, pyramid, wire, holes) can be realized [40]. In MACE process, a noble metal, such as silver (Ag), behaves as local cathode and injects holes into Si substrate in contact with metal layer, resulting in formation of SiO_2 . Since HF attacks SiO_2 , it can be easily etched away. In this way, Si nanostructures with a specific pattern and morphology can be synthesized, as shown in Fig. 15.3 [41].

One-dimensional nanostructures such as nanowires can also be synthesized via the sonochemical technique. Similar to hydrothermal technique, this route is based on a seed formation in a specific bulk solution. The energy required for the seed formation is supplied from ultrasonic waves as opposed to the case in hydrothermal approach, which is based on a temperature gradient driven process [42]. In a typical sonochemical process, seeds are formed by a phenomenon called cavitation. Regions where the pressure of liquid is below the vapor pressure are acting as the source of these cavitations, which are generated by ultrasonic waves (15–400 kHz).

In other words, ultrasonic waves triggers first the bubble formation, then col-lap-tion and finally energy release that can be used for the seeding. The localized

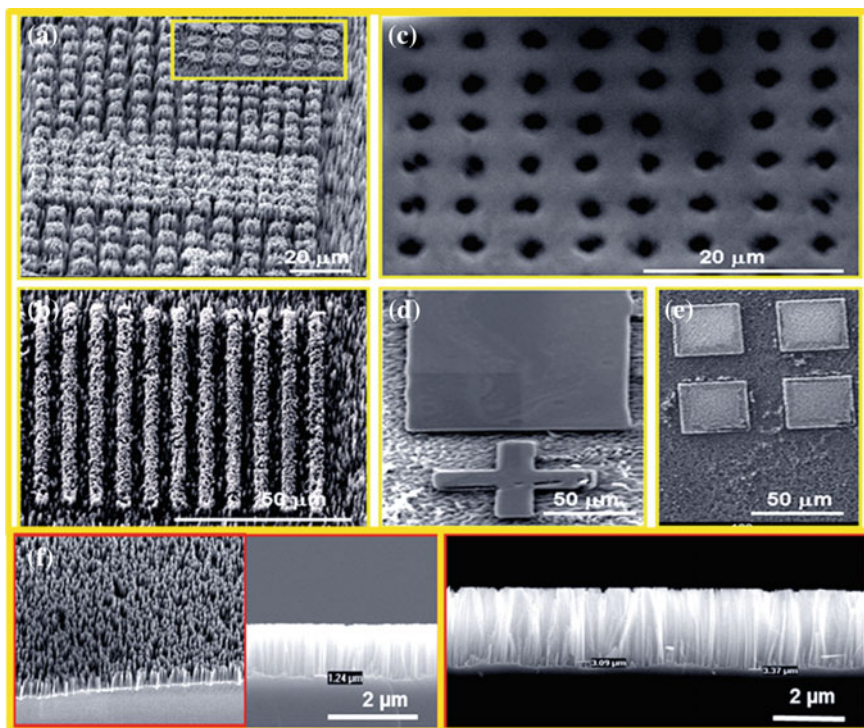


Fig. 15.3 One and two dimensional Si nano/micro structures synthesized using MACE process: **a** Si-pillars, **b** Si-microwalls, **c** Si-microholes with specific pattern, **d**, **e** different Si-micro geometrical features and **f** Si-nanowires

heat release during the collapse of bubbles is around 5000 K. It is this high local heat release that triggers the formation of seeds at extreme conditions. In recent years, there has been an increasing amount of literature on the growth of 1-D nanostructures by using sonochemical approach. For instance, in 2002, Gates and co-workers reported the synthesis of selenium (Se) nanowires using this technique [43]. To grow Se NWs first an aqueous solution composed of selenious acid and hydrazine was used to form Se colloids. After filtering these colloids through a membrane, they were dried and then dispersed in an alcoholic solution. Following these stages, solution was sonicated for about 30 s (42 kHz, 0.15 W/m²). The applied ultrasonic waves resulted in the formation of clusters of the Se colloids with irregular shapes, acting as seeds responsible for the growth of Se NWs based on solid-solution-solid mechanism. In another study, Kumar et al. reported the synthesis of magnetic nanorods using sonochemical process [44]. The sonochemical method was also employed for the growth of ZnO nanowires. We recently reported the synthesis of ZnO NWs in a single step, which took less than an hour growth cycle [45]. This technique allowed us to grow ZnO NWs on virtually any substrates (Si, Cr, polymer, cotton and human hair) as long as they withstand both water and

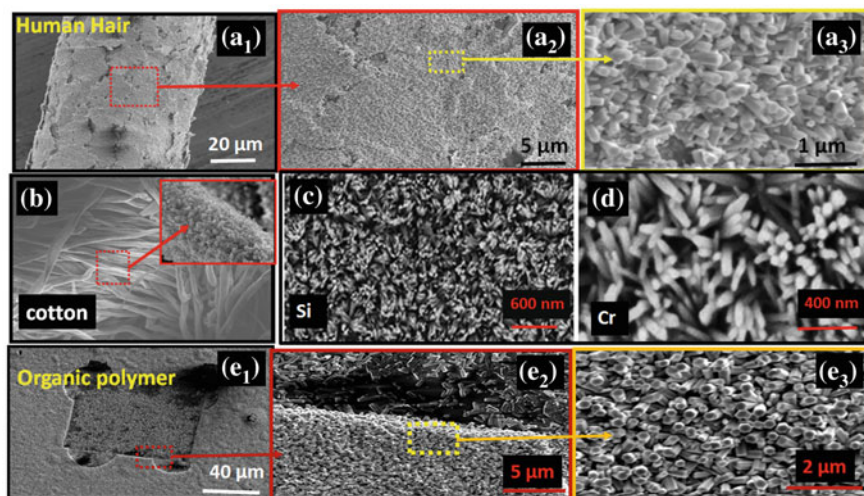


Fig. 15.4 SEM images of ZnO nanowires grown on **a1–a3** human hair, **b** cotton **c** Silicon (Si), **d** chromium (Cr) and **e1–e3** organic polymer (PMMA) substrates using sonochemical technique

alcohol, as shown in Fig. 15.4. Although sonochemical technique is relatively rapid and non-hazardous, it is very challenging to control the shape and size of nanowires (e.g., length and orientation).

15.3.2 Vapor Phase-Based Synthesis of 1-D Nanostructures

Vapor-phase based synthesis of one-dimensional (1-D) nanostructures are particularly preferred for a certain material systems [42]. The growth kinetics behind the vapor-phase based methods are generally catalogued into vapor-liquid-solid (VLS) and vapor-solid (VS) mechanisms.

15.3.2.1 Vapor-Liquid-Solid (VLS) Mechanism

The VLS growth mechanism involves selectively reaction of vapor-phase precursors on a substrate pre-coated with patterned catalyst layer. A selective growth is achieved through the catalyst particles, which form a eutectic liquid solution with nanowire material.

Following the dissolution of precursors from gas phase into the liquid drops, the supersaturation of solution results in precipitation of nuclei, which act as seed points for the growth of 1-D nanostructures [46]. During the growth process, both the diameter and location of 1-D nanostructures can be precisely controlled via the position and droplet size of catalysts; hence 1-D nanostructures with a specific

pattern and density can be produced [47]. A schematic representation for the growth of Si NWs by VLS mechanism is presented in Fig. 15.5.

The use of 1-D nanostructures by VLS mechanism was first reported by Wagner and co-workers to grow the unidirectional Si crystal growth [49]. There is a large volume of published studies on VLS mechanism based synthesis of 1-D nanostructures of a wide range of materials including elemental semiconductors (Ge and Si), III–V semiconductors (GaAs, GaN), II–VI semiconductors (ZnS, CdS and CdSe), and oxides (ZnO and SiO₂) [47, 50–54]. For instance, recently, we synthesized Si NWs via a VLS process by using Au thin film catalyst deposited on Si substrate for transfer printing from Si mother substrate to any carrier substrate, as shown Fig. 15.6 [55].

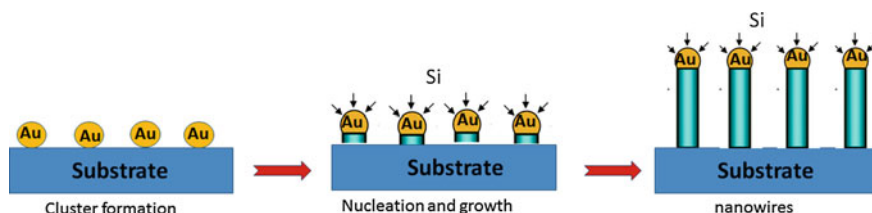


Fig. 15.5 VLS growth mechanism for Si nanowires using Au catalyst [48]

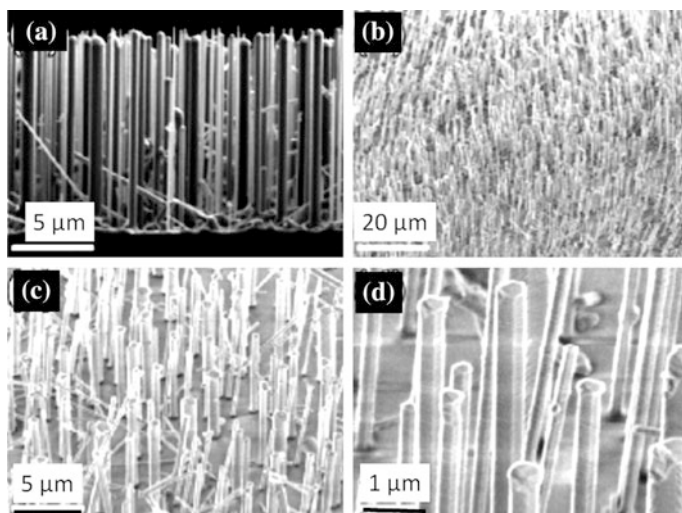


Fig. 15.6 SEM images of Si nanowires (NWs) grown by VLS process. **a** Si NWs grown on Si (111) wafer using Au metal catalyst, **b–d** high magnification SEM images of transferred VLS grown Si NWs from Si wafer and PMMA polymer (embedded in polymer)

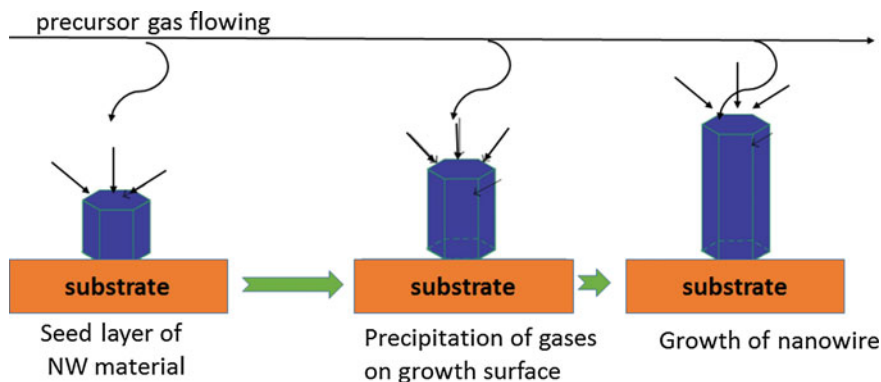


Fig. 15.7 A typical VS growth mechanism [57]

15.3.2.2 Vapor-Solid (VS) Mechanism

In contrast to VLS route, a catalyst is not employed in a typical vapor-solid (VS) route to synthesize one-dimensional nanostructures. In this process, vapor phase of 1-D nanostructure material is generated either by decomposition of precursors or thermal evaporation, which is subsequently transferred by carrier gases (e.g., N_2 or Ar) and condensed on a substrate kept a lower temperature [56], as shown in Fig. 15.7.

As long as one can control some growth parameters, particularly nucleation, it is possible to synthesize 1-D nanostructures by VS mechanism. During the growth cycle of 1-D nanostructures, it is generally believed that nucleation mechanism is driven by either certain defects (e.g., screw dislocation) or anisotropic growth. Although the growth mechanism is not well-understood, it is suggested that at the initial stage of nuclei formation there is a competition between certain crystallographic directions having different growth rates. In the end, one of these directions wins this race and hence promotes the formation of 1-D nanostructure in this favored direction [58]. By now, a considerable amount of literature has been published on VS route based synthesis of one-dimensional nanostructures including nanowires (Si_3N_4 , SiC), metal-oxides (oxides of Zn, In, Al, and Sn), nanobelts (ZnO, In_2O_3 , and SnO) [56, 59, 60] and carbon nanotubes [61–63].

15.4 Common Materials for 1-D Nano-Structured Solar Cells

In this section, we mainly introduce several commonly used materials such as silicon, zinc-oxide and titanium-dioxide that we have employed in our recent studies for the fabrication of one-dimensional nanostructures based solar cells.

In addition, we review the studies on the utilization of carbon nanotubes (CNTs) in dye sensitized solar cells as well as strategies devoted to the development of CNTs based more efficient photoanodes.

15.4.1 *Silicon*

The vast majority of today's solar cells is based on silicon (Si) material due to its various advantages including its abundance, stability, high reliability, and well established processing techniques. However, the requirement of high quality single crystalline Si for the power conversion efficiencies over 20 % increases the cost tremendously. To reduce the cost, recent research efforts are mainly focused on structural changes in devices to reach highly efficient, low cost solar cells. In this regard, Si nanowires (NWs) have received significant interest because of the introduction of new physical mechanisms such as quantum confinement, allowing for the tuning of material properties and overcoming the theoretical efficiency limits for conventional planar p-n junction solar cells. To date a considerable amount of literature has been published on Si nanowires/nanorods based solar cells. However, this subsection will mainly focus on our studies based on synthesis of Si NWs by electroless etching technique and solar cells based on their integration into p-AgGa_{0.5}In_{0.5}Se₂ (AGIS) chalcopyrite thin films. In addition, our study on the fabrication of a hybrid solar cell based on Si-NWs/PCBM will be introduced in detail.

15.4.1.1 **n-Si-NWs/p-AGIS Nanowires Embedded in a Thin Film Type Solar Cell**

Over the last two decades, a large a growing body of literature have been reported on developing high efficient low cost inorganic solar cells. A wide variety of materials are employed for achieving this goal, including poly-crystalline Si, amorphous Si, polymers, II–VI and I–III–VI₂ chalcopyrite (I–III–VI₂) compounds. In particular, chalcopyrite materials are strong candidates for the realization of highly-efficient solar cells based on the construction of p-n heterojunctions utilizing II–VI semiconductors due to their similar crystal structures, which is the key factor to reduce lattice mismatch in the device structure.

Copper (Cu–III–VI₂) and silver (Ag–III–VI₂) based chalcopyrites are well known members for their high absorption coefficient (10^5 cm^{-1}) and an optimum bandgap energy for the solar spectrum [64, 65]. So far, a number of studies have been reported on Ag based chalcopyrite thin film solar cells including n-Ag(In_{1-x}Ga_x)Se₂/CdS:ZnO(Al) [66], p-AgInSe₂/n-CdS [67], p-AgGaSe₂/n-CdS [68] and p-AgGa_{0.25}In_{0.75}Se₂/n-Zn_{0.35}Cd_{0.65}S [69] heterojunctions with power conversion efficiencies ranging from 4.5 to 7.8 %. In addition to incorporating new materials into solar cells, structural changes in devices may also serve the goal to reach high efficient solar cells at lower cost. Although there are many studies reported on one

of these two major research areas individually, so far, however, too little attention has been paid to combine these studies. In this sense, introducing Si NWs into chalcopyrite systems is very promising for fabricating high efficient inorganic solar cells.

In our recent study we reported the synthesis of Si NWs and solar cells based on their integration into p-AgGa_{0.5}In_{0.5}Se₂ (AGIS) chalcopyrite thin films. It was the first study reporting fabrication and characterization of Si NWs/p-AGIS heterojunction based thin film solar cells.

Si NWs were synthesized using metal assisted chemical etching (MACE) technique, details of which is given in earlier section. In order to synthesize Si NWs, an n-type Si wafer was immersed in a solution consisting of 4.6 M hydrofluoric acid (HF) and 0.002 M silver nitride (AgNO₃) [70]. The fabrication of Si NWs was accomplished based on the oxidation-reduction reactions took place in solution. After the synthesis of n-Si NWs, they were decorated with AGIS thin films (800 nm) to form p-n heterojunction based solar cells (see Fig. 15.8) [64]. Following this, indium (In) metallic top contacts with dot pattern were deposited by thermal evaporation. For the back contacts of the solar cell, Si wafer was coated with silver (Ag) by direct current (DC) sputtering.

For the synthesis of Si NWs, we have preferred the MACE technique applied to n-type (100) Si wafers (10 Ω cm resistivity). From the previous studies it is known that MACE route weakly dependent upon the orientation and doping type of the silicon wafer [71]. The growth parameters such as solution temperature and growth time were 60 °C and 5 min, respectively. The cross sectional SEM images of synthesized Si NWs are given in Fig. 15.9. From these images, it is apparent that there are vertically well oriented Si NWs on the Si substrate with an etching depth of $\sim$ 700 nm corresponding to the average NW length. Using MACE technique it is possible to tune the lengths and radii of NWs by adjusting growth parameters such as growth time, etchant concentration and the solution temperature [70].

Once the Si NWs were synthesized, AGIS thin films were deposited by electron-beam technique to realize Si NWs embedded in a thin film type solar cell. To determine the structural, chemical and optical properties of the AGIS thin films prior to the fabrication of p-n junction, AGIS films were deposited onto glass substrates as well. In addition, in order to understand the effect of post-annealing on

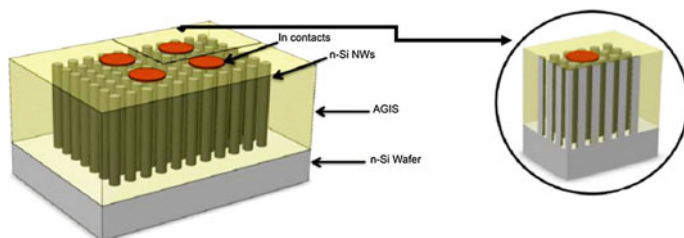


Fig. 15.8 Three-dimensional schematic views of the fabricated n-Si NWs/p-AGIS heterojunction solar cell

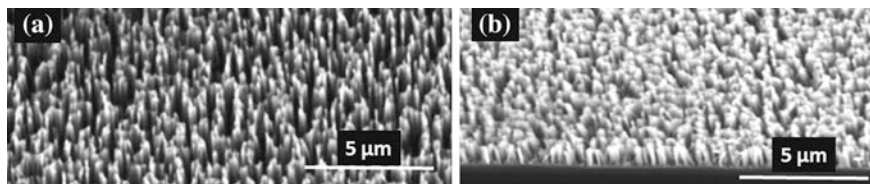
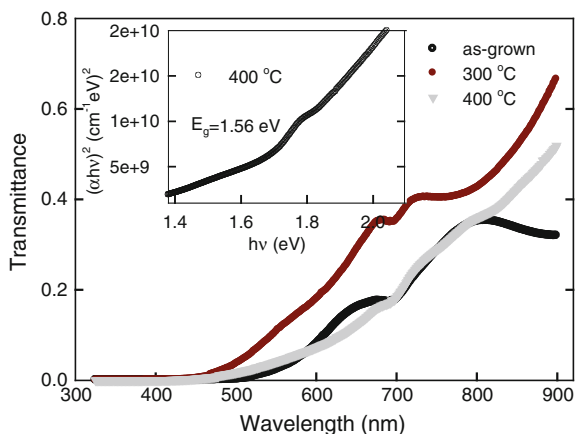


Fig. 15.9 SEM images of synthesized Si NWs using electroless etching techniques

Fig. 15.10 Transmission spectra recorded for as-grown and AGIS films annealed at different temperatures. *Inset* figure presents $(\alpha h\nu)^2$ versus $(h\nu)$ relation for AGIS film annealed at 400 °C



the properties of the films, AGIS thin films were subjected to annealing process in the temperature range of 300–400 °C for 30 min in N_2 gas atmosphere.

Transmission measurements were carried out in the wavelength range of 325–900 nm to determine the band gap of the deposited AGIS thin films. The spectrum obtained for both as-grown and annealed AGIS thin films are shown in Fig. 15.10. As can be seen from the transmission spectra, the average transmittance is around 40 % at 800 nm for all samples. Absorption coefficient (α) was evaluated via the transmission spectra. Through the relation between photon energy ($h\nu$) and absorption coefficient, $(\alpha h\nu) = A(h\nu - E_g)^n$, the nature and energy value of the band gap of AGIS films were determined from the plot of $(\alpha h\nu)^2$ versus $(h\nu)$, shown as an inset in Fig. 15.10 for an AGIS film annealed at 400 °C. The results obtained from the analysis show that there is a direct nature ($n = 1/2$) of the optical transition and the optical band gap value is around 1.56 eV, which is exactly matching the most abundant part of the solar spectrum (1.3–1.5 eV).

A nanowires embedded in a thin film type solar cell was constructed by embedding n-Si NWs in a p-AGIS poly-crystalline thin film. The cross-sectional view of the fabricated solar cell with $AgGa_{0.5}In_{0.5}Se_2/n\text{-Si}$ (NWs)/n-Si structure is presented in Fig. 15.11. The dark current (I)–voltage (V) characteristic of In/p-AGIS/n-Si NWs/n-Si/Ag solar cell was tested at room temperature to determine the formation of p-n junction between n-Si NWs and p-AGIS thin film. As shown in

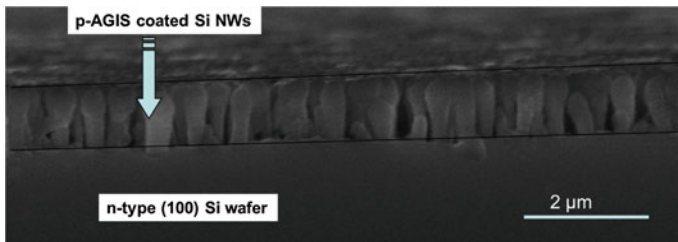


Fig. 15.11 SEM image of the fabricated n-Si NWs/p-AGIS structured solar cell (cross sectional view)

Fig. 15.12 Current (I)-voltage (V) characteristic of n-Si NWs/p-AGIS heterojunction measured in dark condition

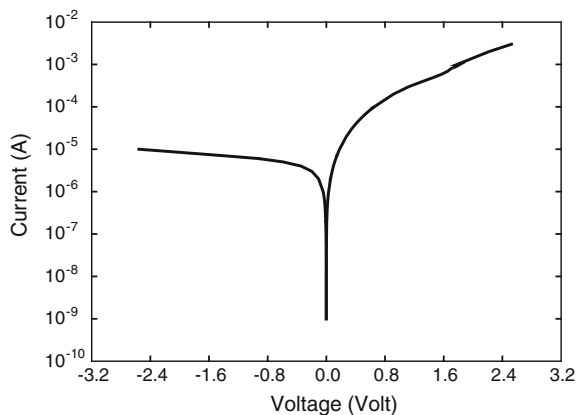


Fig. 15.12, the device shows the typical p-n junction characteristics with rectification ratio ($I_{\text{forward}}/I_{\text{reverse}}$) of 100 at 1 V, which verifies the formation of p-n heterojunction.

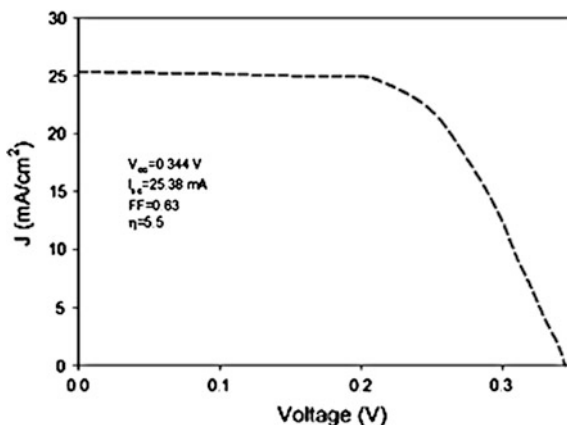
The performance of fabricated solar cell was tested by a solar simulator under AM 1.5 illumination (100 mW/cm^2). The I-V characteristic recorded under the replicated solar spectrum is presented in Fig. 15.13. The power conversion efficiency (η) of a typical solar cell can be calculated from

$$\eta = \frac{V_{oc} I_{sc} FF}{P_{in}} \quad (15.1)$$

where V_{oc} , I_{sc} , FF and P_{in} are open circuit voltage, short circuit current, fill factor and incident power, respectively.

The V_{oc} , J_{sc} , and FF of the solar cell were found to be 0.34 V, 25.38 mA/cm^2 and 63 %, respectively, corresponding to a power conversion efficiency of 5.50 %. Besides these, series (R_s) and shunt (R_{sh}) resistances, showing the quality of solar cell, were also calculated through the slope of the I-V curve near the J_{sc} and V_{oc} , which were calculated to be $32 \text{ k}\Omega$ and 15Ω , respectively.

Fig. 15.13 I–V characteristic of n-Si NWs/p-AGIS heterojunction solar cell recorded under the replicated solar spectrum



In reviewing the literature, no studies were found on a solar cell based on n-Si NWs embedded in p-AGIS thin films; therefore, it is not possible to compare our results with literature. Although constructed with different materials combination, a solar cell with a similar architecture, based on n-CdS nanopillar arrays embedded in a p-type polycrystalline CdTe thin film, was reported in literature [18]. A 6 % power conversion efficiency was achieved by embedding CdS nanopillars into CdTe thin film.

To date, a numerous studies have been reported in the literature on fabrication of Ag-based chalcopyrite thin film heterojunction planar solar cells constructed with a wide variety of materials including CdS: ZnO (Al), n-CdS and n-Zn_{0.35}Cd_{0.65}S thin films with power conversion efficiencies ranging from 4.5 to 7.8 % [66–69]. In our study, the efficiency achieved with Si NWs based solar cell lies in range of reported values. For instance, when compared with n-Ag(In_{1-x}Ga_x)Se₂/CdS:ZnO(Al) planar thin film heterojunction solar cells, we have measured an almost two times larger short circuit current density with the n-Si NWs/p-AGIS solar cell structure. The enhancement in current density is expected outcome for a hybrid nanowire-thin film heterojunction solar cells when compared with its planar counterpart since it provides both light trapping and efficient charge collection to the electrodes for generated free carriers. In addition, in a typical planar thin film based solar cell, there is high probability of formation of recombination and trap centers, reducing the collection efficiency of generated free carriers, due to the polycrystalline structure of films (e.g., grain boundaries, point defects). That's why, although the two architectures have different n-type materials, the observed high short current density (25.38 mA/cm²) for n-Si NWs/p-AGIS relative to planar n-Ag(In_{1-x}Ga_x)Se₂/CdS:ZnO(Al) (14.50 mA/cm²) is quite reasonable, which highlights the unique advantages of our constructed solar cell based on using Si nanowires.

Although the energy conversion efficiency obtained from n-Si NWs/p-AGIS solar cell is lower than that reported for its planar counterparts constructed with silicon and some Ag-chalcopyrite thin films, it is possible to realize more efficient NW based

solar cells by optimization of some fabrication steps including NW synthesis, thin film layer deposition, contact issues, and surface passivation processes.

Among these optimization processes, for instance, surface passivation alone would increase the energy conversion efficiency of devices to over 5.50 %. Up to now, many research groups have focused on the effect of passivation on efficiency of solar cells based on Si NWs. Gunavan and Guha reported that with incorporation of 10 nm thick Al_2O_3 conformal passivation layer onto Si NWs' surface the solar efficiency increased from 1 to 1.8 % by increment of 47 mV in V_{oc} and 5 mA/cm^2 in J_{sc} [72]. In another major study, Kim et al. reported the introducing a-SiN:H for the passivation of Si micro-wire based solar cell [73]. It was observed that by introducing the passivation the power conversion efficiency was improved from 8.9 to 10.6 %.

In our study, although the AGIS has been employed as absorber thin film layer in which Si NWs were embedded, the same device architecture can be used with a wide variety of other PV materials including CdTe, CdS, and CuInSe_2 to improve the energy conversion efficiency beyond 5.50 % achieved in this work. In addition, as we demonstrated in our previous studies, Si NWs, micro-walls and micro-pillars transfer printed from the single crystal wafer to any arbitrary substrate [55, 74, 75] via bottom-up methods may considerably reduce the cost of such solar cells, as shown in Fig. 15.14.

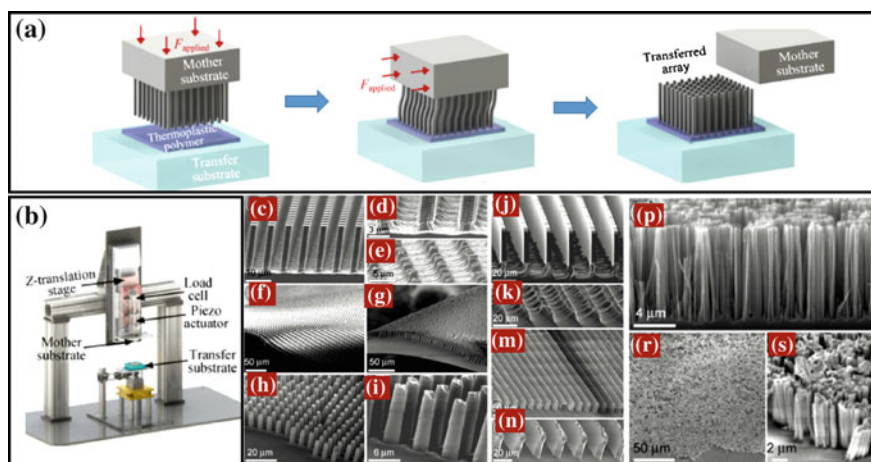


Fig. 15.14 **a** Schematic representation of vertical transfer printing process for 1-D and 2-D Si micro/nano structures, **b** the vertical embossing tool used during transfer printing, **c** Si micro pillars (MPLs) produced by deep reactive ion etching (DRIE), **d** and **e** SEM images showing root location of the MPLs on the mother substrate before and after transferring process, respectively, **f**–**i** transferred MPLs embedded in polymer, **j** Si microwalls (MWAs) produced by DRIE, **k** roots of MWAs following the transfer printing, **m** and **n** low and high magnification SEM images of MWAs transferred into polymer, respectively, **p** Si NWs on Si wafer synthesized by using electroless etching technique, **r** and **s** show the high magnification SEM images of the electroless-etched Si NWs transferred from mother substrate

15.4.1.2 Si-NWs/PCBM Structured Hybrid Solar Cell

In the past few years, an intense interest has emerged in combining an organic material with NWs for the fabrication of hybrid solar cells due to their significant advantages over the conventional organic solar cells [76–80].

The requirement of charge separation at an interface and low diffusion lengths of excitons in a typical organic material limits the efficiency of organic solar cells. A reasonable approach to tackle these problems could be to integrate nanowires (NWs) into a photoactive organic material. In recent years, there has been an increasing amount of literature on incorporation of NWs, such as Si, CdS, ZnO, and TiO_2 , into an organic film [81]. For instance, recently, we have reported a hybrid solar cell based on integration of Si NWs into PCBM thin film layer [82]. The aim of this study was to increase diffusion length of excitons in the organic layer and enhancing the interface area that can be realized through the integration of NWs due to their large surface/volume ratio and ability to provide single crystalline channels for photo-carrier charge collection. It is the first study reporting the fabrication of a hybrid solar cell based on PCBM coated vertically well-aligned p-Si NW arrays.

p-type Si nanowires (NWs) were synthesized using a p-type Si wafer (with (100) orientation and $(1\text{--}10\ \Omega\ \text{cm})$ resistivity) by electroless etching technique. On completion of Si NWs synthesis, p-type PCBM (120 nm) polymer solution was spin coated onto the substrate. Following this, the device annealed at $120\ ^\circ\text{C}$ for 30 min under vacuum. For the top contact electrodes dot patterned ($1.6\ \text{mm}^2$) ITO (200 nm thick) were deposited using RF sputtering, while for the back contacts a 250 nm thick Al was deposited onto the back side of the Si wafer. The schematic representation of the device structure and its energy level diagram are shown in Fig. 15.15 [83].

The morphology of synthesized Si nanowires (NWs), and the structure of the fabricated hybrid solar cell were revealed via the SEM micrographs, shown in Fig. 15.16. It can be seen from the SEM images that average length and diameter of

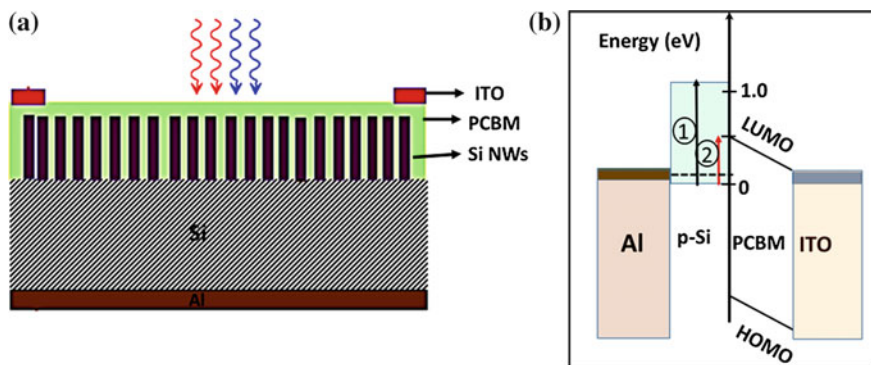


Fig. 15.15 a Schematic representation of p-Si NWs/PCBM/ITO structured hybrid solar cell and b its energy level diagram

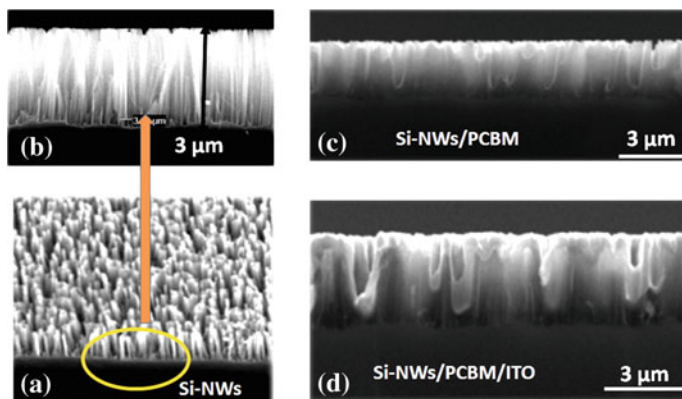
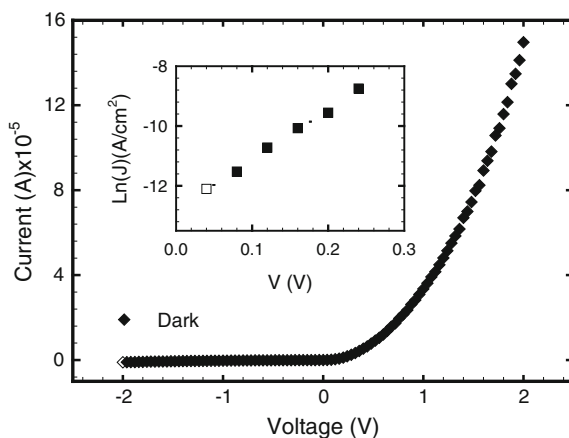


Fig. 15.16 **a** Tilted and **b** cross sectional view of synthesized Si NWs by electroless etching technique. Cross sectional SEM images of Si NWs first coated with **c** PCBM organic thin film and then with **d** transparent conductive oxide (ITO) top contacts

Fig. 15.17 The I–V characteristic of the Al/p-Si NWs/PCBM/ITO structured hybrid solar cell measured under dark condition

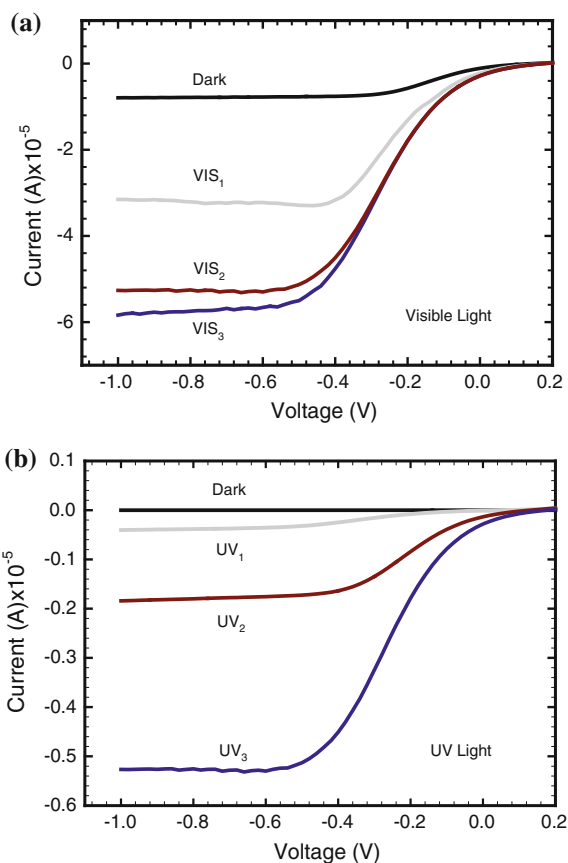


Si NWs are around $3\ \mu\text{m}$ and $\sim 70\ \text{nm}$, respectively. It is also apparent from these images that the deposited PCBM by spin coating resulted a conformal thin film layer over the Si NWs matrix. The component layers of p-Si NWs/PCBM/ITO structured hybrid solar cell can be easily distinguished from the cross-sectional view shown in Fig. 15.16d.

Figure 15.17 presents the I–V characteristic of the Al/p-Si NWs/PCBM/ITO structured solar cell under dark condition. It is apparent from the graph that the device exhibits a standard rectifying diode characteristic with a current rectification ratio of 78 at 0.5 V, verifying the heterojunction between Si NWs and PCBM organic layer.

The I–V characteristics of the solar cell under visible (VIS, 1050 nm) and ultra-violet (UV, 350 nm) light illuminations are presented in Fig. 15.18a, b,

Fig. 15.18 The I–V characteristics of the Al/p-Si NWs/PCBM/ITO structured hybrid solar cell measured under different **a** visible and **b** UV light illumination intensities

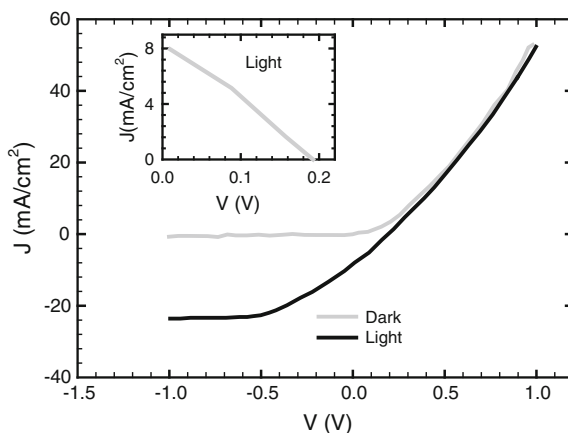


respectively. From these figures we can see that the measured current gradually increases when the intensity of both UV ($UV_3 > UV_2 > UV_1$) and VIS ($VIS_3 > VIS_2 > VIS_1$) lights is increased, a clear verification of that both p-Si NW and PCBM components in the device structure contribute the photocurrent. In other words, it is reasonable to expect a contribution to photocurrent in UV and VIS part of spectrum from Si NW and PCBM since the respective band gaps of these materials are 1.1 and 2 eV (energy difference between the PCBM-LUMO and PCBM-HOMO) [83]. These results suggest that this type of a hybrid heterostructure of p-Si NW/PCBM can be utilized to fabricate portable, low cost and high efficient UV or VIS sensitive photodiodes.

In order to determine the photovoltaic behavior of the p-Si NW/PCBM hybrid solar cell, the J–V (current density vs. voltage) characteristic was tested under AM 1.5G illumination, as presented in Fig. 15.19.

The V_{oc} , I_{sc} , FF and η of the fabricated hybrid solar cell were found as 0.19 V, 8.1 A/cm², 27 and 0.41 %, respectively. To our best knowledge, this is the first study reported on integration of Si NWs into the PCBM organic thin film. When the

Fig. 15.19 The photovoltaic behavior of the p-Si NWs/PCBM/ITO structured hybrid solar cell tested under AM 1.56 illumination



performance of this solar cell is compared with its planar p-Si/PCBM hybrid solar cell (tested under 6 mW/cm^2), with $V_{oc} = 0.12 \text{ V}$ and $I_{sc} = 0.1 \text{ } \mu\text{A}$ reported in literature [76, 83], there is significant improvement in all parameters following the integration of Si NWs into PCBM layer. This improvement can be attributed to high efficiency of carrier collection, light trapping and formation of high interface between Si NW and PCBM provided by unique feature of NWs. The power conversion efficiency of 0.41 % recorded with Al/p-Si NW/PCBM/ITO structured hybrid solar cell can be further improved by several optimization process including the geometry and size of nanowires, passivation process and some contact issues.

15.4.2 Zinc-Oxide

In this part of chapter, we will mainly focus on our recent study based on synthesis of ZnO NWs by hydrothermal technique and solar cells based on their integration into a chalcopyrite thin film ($\text{p-AgGa}_{0.5}\text{In}_{0.5}\text{Se}_2$ (AGIS)) and decoration around silicon-micro pillars/walls. These are the first results on fabrication and characterization of n-ZnO NWs/p-AGIS heterojunction based thin film solar cells. In addition, we will briefly review the research conducted on ZnO one-dimensional nanostructures based solar cells with different architectures.

15.4.2.1 n-ZnO NWs/p-AGIS Heterojunction Based Thin Film Solar Cells

ZnO is an n-type wide band gap (3.37 eV) semiconductor oxide with a large exciton binding energy (60 meV). It can be employed for the fabrication of a wide range of high-performance opto-electronic devices because of its unique properties.

In particular, zinc-oxide (ZnO) nanowires (NWs) have been extensively studied for a large number of applications including field ionization gas sensors, light emitting diodes, solar cells, photodetectors, chemical sensors and field effect transistors due to their unique physical properties as well as their excellent chemical and thermal properties [84–90].

So far, many works have been reported on one-dimensional ZnO nanostructures based solar cells [91–98]. We reported a nanowires embedded in a thin film type solar cell consisted of ZnO NWs and a chalcopyrite thin film [33]. Using hydrothermal technique, vertically well-oriented ZnO NWs were grown onto both a glass substrate pre-coated with indium thin oxide (ITO) and n-Si wafer substrate. In order to synthesize ZnO NWs, the required ZnO seed layers with a wide variety of thicknesses (10–750 nm) were deposited onto ITO/glass and n-Si wafer substrates using radio-frequency (RF) magnetron sputtering. Following this, seed layer pre-coated substrates were immersed in a solution consisting of equimolar 25 mM ($\text{Zn}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$) (Zinc nitrate hexahydrate, Sigma Aldrich) and HMTA (Hexamethylenetetramine, Sigma Aldrich) as precursors in deionized (DI) water. Growth parameters such as solution temperature and the growth time were selected to be 90 °C and 2–3 h, respectively. For the purpose of construction of a p-n heterojunction based solar cell, ZnO-NWs/ITO/glass-substrates (with 750 nm ZnO seed layer) were then coated with a 800 nm thick $\text{AgGa}_{0.5}\text{In}_{0.5}\text{Se}_2$ (AGIS) thin films, details of deposition and structural and optical properties of which was discussed in previous section and elsewhere [64].

To investigate the morphology and some physical parameters of grown ZnO NWs such as length, diameter and density, scanning electron microscopy (SEM) measurements were conducted. The effect of substrate on the quality of grown ZnO NWs was investigated by growing NWs on both glass and Silicon substrates, pre-coated with 10 nm ZnO seed layers. Figure 15.20 shows SEM images recorded for ZnO NWs synthesized on these substrates. It can be seen from the images that highly dense and vertically oriented ZnO NWs with a length of about 0.6–1.0 μm and diameters in the range of 70–125 nm were grown on both substrates. It can also be seen that the preferential growth direction of ZnO is along the c-axis of the substrate.

To reveal the effect of the seed layer thickness on the quality of grown ZnO NWs, ZnO seed layers with various thicknesses (10–750 nm) were prepared on glass substrates, SEM images of which are presented in Fig. 15.21.

It was observed that when the thickness of the seed layer was increased, there was significant improvement in quality of NWs in terms of uniformity and orientation in the vertical direction. Moreover, it was deduced that when thickness was increased, NWs with larger diameters were grown, which was attributed to the coalescence of vertically-oriented NWs grown on the thicker seed layer. These findings were in a good agreement with results reported in some previous studies [99]. In general, the low-quality ZnO NWs are expected when they are grown on thin ZnO seed layer due to the poor crystal characteristic of seed layer. It can thus be suggested that the density and diameter of ZnO NWs can be tuned by adjusting the thickness of ZnO seed layer.

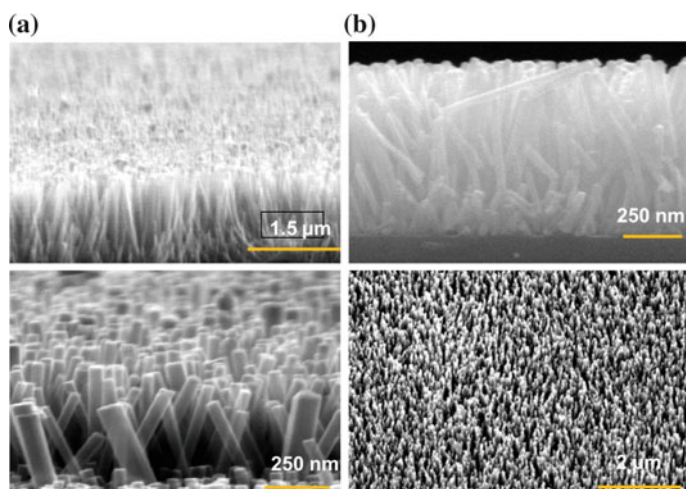
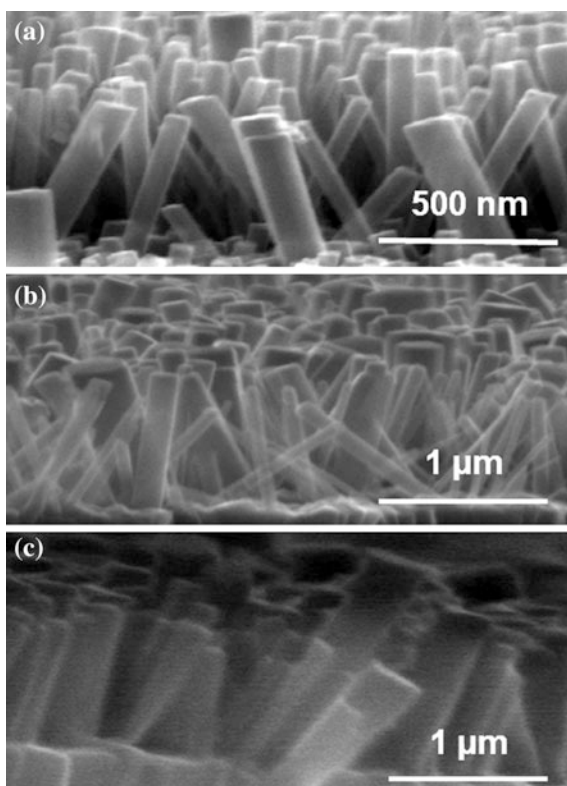


Fig. 15.20 SEM images recorded for ZnO NWs grown on **a** soda-lime glass and **b** silicon substrates

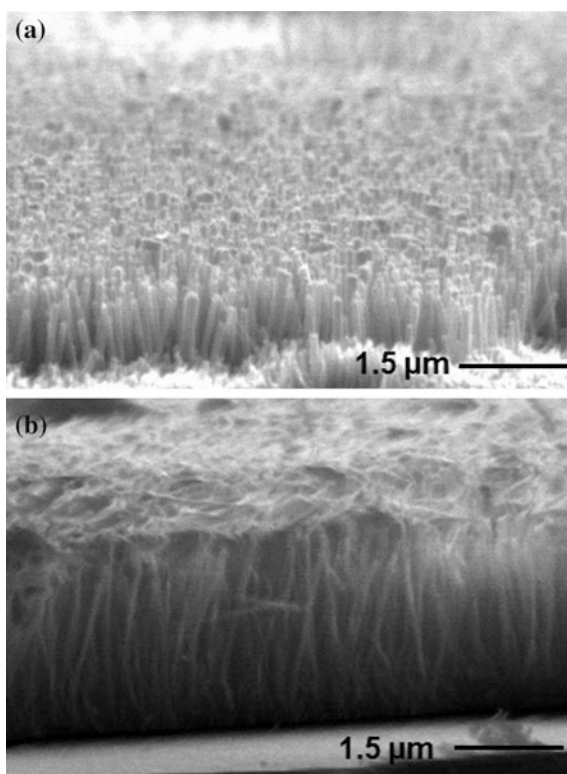
Fig. 15.21 ZnO NWs grown on soda-lime glass substrates pre-coated with **a** 10 nm, **b** 500 nm, and **c** 750 nm thick ZnO seed layers deposited by RF sputtering technique



The effect of the growth time on the length of synthesized ZnO NWs by hydrothermal technique was also studied by keeping all growth parameters constant except the growth time. The comparison NWs synthesized for growth times of 2 and 3 h is shown in Fig. 15.22. Results demonstrated that once the growth time exceeded 2 h, the ZnO thin film grew instead of nanowires due to the coalescence of longer NWs grown at the end of longer time growth period.

ZnO NWs based solar cell was constructed by integration of n-ZnO NWs into a p-AGIS poly-crystalline thin film. The device then annealed at 400 °C in N₂ gas atmosphere for 30 min to enhance the crystallinity of absorber layer (AGIS), which is a heat-energy driven process. The structure of constructed solar cell, AgGa_{0.5}In_{0.5}Se₂/n-ZnO (NWs)/ZnO (seed layer)/ITO/glass, was viewed through the recorded SEM images, presented in Fig. 15.23. The top and bottom contacts of the device for current collection were thermal evaporated indium (with 1 mm diameter dots) and RF sputtered ITO layer, respectively. The dark I–V characteristic of the device at room temperature is shown in Fig. 15.24a. It can be seen the figure, the device exhibits a typical diode behaviour (p-n junction characteristics) with rectification ratio ($I_{\text{forward}}/I_{\text{reverse}}$) of 10 at 1 V, verifying the formation of p-n junction between ZnO NWs and AGIS thin film.

Fig. 15.22 The comparison of ZnO NWs synthesized for growth times of **a** 2 h and **b** 3 h



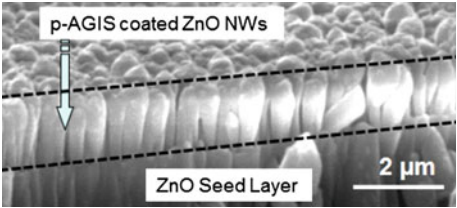
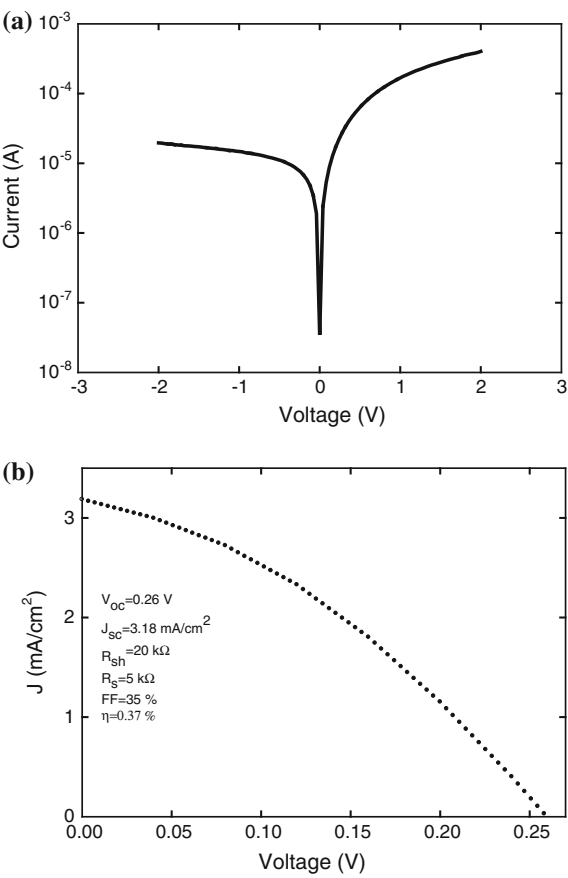


Fig. 15.23 The cross sectional SEM image of Glass/ITO/ZnO (seed layer)/n-ZnO NWs/p-AGIS structured heterojunction solar cell

Fig. 15.24 The I–V characteristics of the fabricated solar cell under **a** dark and **b** light illumination (AM 1.56)



The photovoltaic performance of the solar cell was determined under standard test conditions using a solar simulator (under AM 1.5 illumination (100 mW/cm^2)), J–V characteristics of which is shown in Fig. 15.24b.

The power conversion efficiency was found to be 0.37 %. There are a wide variety of potential factors that can determine the efficiency of a solar cell. In our case, the most probable factors behind the observed low efficiency of the ZnO NW cell are very low shunt resistance ($R_{\text{sh}} = 20 \text{ k}\Omega$), high series resistance ($R_{\text{s}} = 5 \text{ k}\Omega$) and some fabrication related shortcomings.

These have a direct effect on obtaining a low short circuit current density ($J_{\text{sc}} = 3.18 \text{ mA/cm}^2$) and open circuit voltage ($V_{\text{oc}} = 0.26 \text{ V}$). In addition, the geometry of the NWs (length, diameter and vertically orientation), the embedded length of NWs in the thin film, the NW-thin film interface and the crystal quality of the deposited AGIS thin film can be regarded as the potential factors responsible for the high series and low shunt resistances. In particular, the co-existence of two separate phases (AGIS and Ag) in thin film structure of absorber layer deduced from XRD study may result in the formation of disorder related local defects which promote the recombination of photogenerated free carriers.

15.4.3 *Titanium-Dioxide*

In recent years, TiO_2 (Titanium Dioxide), one of the most important wide band gap materials ($\sim 3.0 \text{ eV}$ for rutile TiO_2 structure), has received a considerable attention for their potential applications in various areas including solar cells [100], water photoelectrolysis [101], field emission devices [102], and gas sensors [103].

In particular, intense efforts have been devoted to TiO_2 nanorods (NRs) based solar cells to realize high performance solar cells at lower cost.

Among the aforementioned one-dimensional nanostructures based solar cell architectures (axial, core-shell (or radial), and NRs embedded in thin films), the core-shell is of particular interest to improve the power conversion efficiency over the traditional planar thin film based solar cell due to its enhanced interaction junction area, improved carrier transport to electrodes, and ability to orthogonalize the light absorption and charge collection.

In other words, with this device architecture not only enhanced light absorption but also efficient charge collection to electrodes are accomplished on condition that core radius and shell thickness are optimized properly. To date a considerable amount of literature has been published on TiO_2 NRs based dye-sensitized solar cells [104–115]. Few studies, however, have focused on TiO_2 NRs based semiconductor-sensitized solar cells [116]. Due to superior properties of TiO_2 NRs and excellent properties of CdTe as an absorber layer with an ideal band gap matching the most abundant part of solar spectrum (1.45 eV), TiO_2/CdTe is a one of the most promising material combinations in the fabrication of high-efficient solar cells at lower cost [117]. In this section of chapter, we introduce our recent

study based on an n-TiO₂/p-CdTe core-shell structured solar cell constructed by a two-step route combining hydrothermal method and sputtering technique.

15.4.3.1 n-TiO₂/p-CdTe Core-Shell Structured Solar Cells

Vertically-well oriented TiO₂ nanorods (NRs) were grown on a fluorine-doped thin oxide (FTO)-pre-coated glass substrate using hydrothermal growth technique, detail of which was given elsewhere [118, 119]. After synthesis of T TiO₂ NRs, in order to form an n-TiO₂/p-CdTe core-shell type solar cell, NRs were decorated with a ~ 250 -nm-thick CdTe thin film deposited by sputtering technique. Following this, 10 nm Cu and 40 nm Au metallic layers of top contacts with dot pattern were deposited by thermal evaporation. The device was then annealed at 120 °C for 30 min to form ohmic type contacts.

The SEM images of synthesized TiO₂ nanorods (NRs) grown on a FTO glass substrate recorded with different magnifications are presented in Fig. 15.25. As can be seen from Fig. 15.25a, uniformly distributed TiO₂ NRs were grown successfully on FTO substrate. It can also be seen from Fig. 15.25b that the dimension of NRs is around 1.8 μm for the length and 100 nm for the average diameter.

Figure 15.25c, d show cross-sectional and tilted view of SEM images of fabricated TiO₂/CdTe structured core-shell solar cell. It is worth to remark that the diameter of core-shell TiO₂ increases from bottom to top, resulting in formation of a

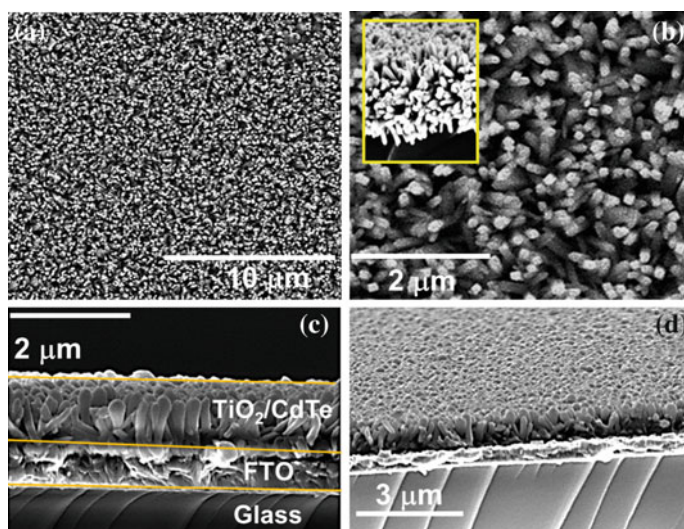


Fig. 15.25 a Low and b high magnification SEM images of synthesized TiO₂ NRs on FTO pre-coated glass substrates, c cross sectional and d tilted view of the completed solar cell device structure

baseball bat-like core shell solar cell following the deposition of CdTe thin film layer.

TEM images of TiO_2 NRs coated with CdTe thin film are shown in Fig. 15.26. The baseball bat-like solar cell structure is deduced from these images as well, mainly attributable to the deposition parameters of CdTe. In particular, deposition rate is a key parameter determining the form of core-shell due to the necessity of a balanced crystal growth rate and deposition rate for a conformal coating of 1-D nanostructures. It is usually expected that in the case of a high crystal growth rate compared to deposition rate by sputtering process there is a sufficient time required for the effective pre-cursor nucleation on NRs surface, which then results in this type formations [116].

In order to analyze the interface region formed between TiO_2 and CdTe, a high resolution TEM image was recorded. As can be seen from Fig. 15.26c, the boundary between TiO_2 and CdTe phases are distinguishable as well as crystal planes associated with each phases.

To investigate the optical properties of the constructed device and TiO_2 NRs the ultraviolet-visible-near infrared (UV-VIS-NIR) absorption spectra were recorded in the 350–1000 nm wavelength range, as shown in Fig. 15.27. It is apparent from the absorption spectrum that there are two onsets, which are at 885 (1.4 eV) and 410 nm (3.02 eV) associated with band gaps of CdTe and TiO_2 , respectively. This result can be taken as a strong evidence for the contribution of both materials to the absorption process, resulting in the observed broadening of absorption spectrum for a wide wavelength range including UV, VIS, and NIR. The high sensitivity in these wavelength range under solar spectrum is a crucial necessity for the fabrication of high efficient solar cells since the most intense solar spectrum reaches on the surface of the earth is in the 1.4–1.6 eV photon energy range.

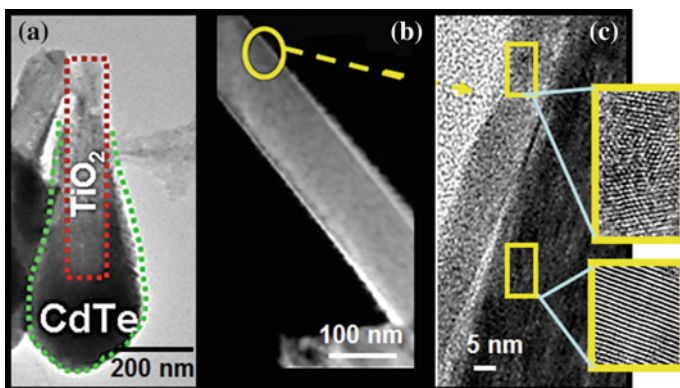


Fig. 15.26 TEM images recorded for TiO_2 NRs decorated with sputtered CdTe thin film layer: **a** low magnification image showing the formation of a core-shell type solar cell, **b** image revealing the systematic variation in thickness of deposited CdTe layer on TiO_2 NRs from *bottom* to *top*, **c** high magnification image revealing the atomic planes of CdTe and TiO_2 phases

Fig. 15.27 Absorption spectra recorded for TiO_2 NRs and TiO_2/CdTe core-shell structure

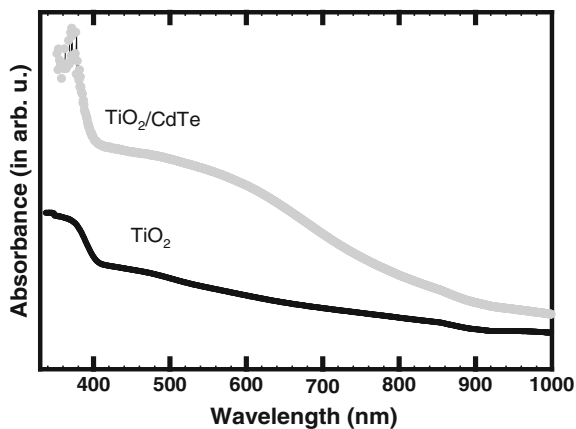


Fig. 15.28 a Responsivity spectrum obtained for TiO_2/CdTe core-shell solar cell, **b** current density (J)–voltage (V) characteristics of the solar cell measured under dark and light illumination (AM 1.56)

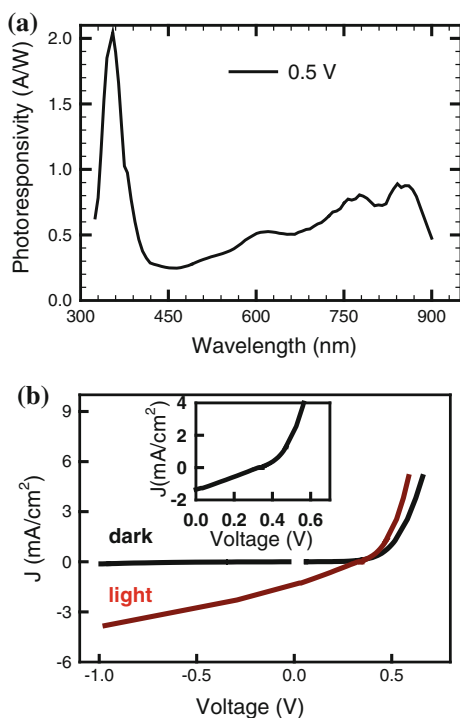


Figure 15.28a, b show the responsivity as a function of wavelength (300–900 nm) at 0.5 V and J – V characteristic of $n\text{-TiO}_2/p\text{-CdTe}$ structured core-shell solar cell, respectively. The variation in photoresponsivity observed in Fig. 15.20a is consistent with the results extracted from absorption spectrum. In other words, there are three humps that can be resolved from the photoresponsivity located at 600 nm

(2.06 eV), 775 nm (1.6 eV), and 854 nm (1.41 eV). Of these peaks, the one observed at 854 nm (1.41 eV) is associated with the optic band gap of CdTe. As for the ones observed at 600 and 775 nm, they are likely to be stemming from the defect states formed either in TiO_2 crystal or at interface formed between TiO_2 and CdTe in core-shell structure.

The J–V characteristic of n- TiO_2 /p-CdTe core-shell solar cell under standard test conditions (A.M. 1.5G) is shown in Fig. 15.28b. Dot patterned Cu/Au and FTO substrates were preferred for top and bottom electrodes of solar cells, respectively during the measurement. It is apparent from the graph that there is a rectification of current, verifying the formation of a p-n heterojunction between CdTe and TiO_2 . It is also clear from the figure that device shows the photovoltaic behavior under illumination. The solar cell parameters, V_{oc} , J_{sc} , FF and power conversion efficiency (η) were found to be as 0.34 V, 1.27 mA/cm², 28 and 0.12 %, respectively.

As it is the first study reporting the solar cell parameters for a TiO_2 -CdTe core-shell solar cell, it is not possible to compare with any study in literature.

Recently, however, a core-shell solar cell based on CdTe decorating TiO_2 nanotube arrays was reported by Zhang et al. [117]. The solar parameters such as the V_{oc} and J_{sc} for this device structure were 0.23 V and 1.23 mA/cm², respectively and with a corresponding power conversion efficiency of 0.10 %, slightly lower than that we obtained with TiO_2 -CdTe core-shell solar cell.

The low power conversion efficiency achieved with n- TiO_2 (NRs)/p-CdTe(thin film) core-shell solar cell, 0.12 %, could be improved by applying several optimization process such as improving the quality of CdTe thin film (e.g., grain size, conductivity, carrier concentration, and mobility), adjusting the morphology (e.g., length, radius, density and orientation) of synthesized TiO_2 NRs properly, electrical contacts to CdTe coated TiO_2 NRs. In addition to these, the post-production optimization process such as CdCl_2 treatment to CdTe thin film may have a pronounceable effect on improving the power conversion efficiency of solar cell over 0.12 % due to its effect on increasing the grain sizes in polycrystalline structure of CdTe.

In addition to our studies, the solar cell device performances of nanowires derived from a wide variety of materials reported in literature are summarized in Table 15.1.

15.4.4 Carbon

Carbon nanotubes (CNTs) present new opportunities to design third-generation photovoltaic devices by enhancing the efficiency and further reducing the production costs to make them competitive in the market. CNTs are investigated in bulk heterojunction [133] as well as in inorganic thin film [134] solar cells. In this section, we review the studies about the utilization of CNTs in dye sensitized solar cells. We describe briefly the material properties of CNTs first. Strategies devoted

Table 15.1 Solar cell device performances of nanowires derived from a wide variety of materials

Solar cell type	Material	^a (J _{sc}) (mA/cm ²)	^b (V _{oc}) (V)	^c (FF)	^d (η) %	References
Core-shell	Si NWs	4.28	0.29	0.33	0.46	[9]
	CdS-Cu ₂ S	–	0.61	0.80	5.40	[120]
	ZnO/CdS	7.23	–1.55	–	3.53	[121]
	ZnO/TiO ₂ /P ₃ HT	0.77	0.51	0.54	0.23	[122]
	p/i/n Si NWs	23.9	0.26	0.55	3.40	[123]
	TCO/a-Si/Si NWs	27	0.476	0.56	7.29	[124]
	amorph-Si NWs	13.9	0.830	0.524	6.00	[125]
	CuO/ZnO	0.89	0.445	–	0.17	[126]
	Si	4.28	0.29	0.33	0.46	[127]
	GaAs	0.20	0.20	0.27	0.83	[128]
	InP	13.72	0.43	0.57	3.37	[8]
	ZnO/ZnS	–	0.09	0.28	–	[129]
Axial	GaAs	21.08	0.565	0.6365	7.58	[23]
	p ⁺ nn ⁺ -Si	13	0.450	0.30	4.40	[130]
Embedded	PEDOT:PSS-Si NWs	9.38	0.43	0.45	1.82	[131]
	ZnO/AgNW/ZnO	31	0.458	0.45	6.37	[132]
	ZnO/Cu ₂ O ₂	1.43	0.15	0.25	0.05	[22]

^aShort current density^bOpen circuit voltage^cFill factor^dEfficiency

to the development of more efficient photoanode materials and to replace the counter electrode (Pt) by carbon nanotubes will be described.

15.4.4.1 Carbon Nanotubes

Carbon nanotubes (CNTs) are cylinder-shaped macromolecules with large aspect ratios. The length axes can grow in a range of 100 nm up to 20 cm depending on the synthesis conditions. In the most general case, a CNT is composed of a concentric arrangement of many cylinders. Such multi-walled nanotubes (MWNTs) can reach diameters of up to 100 nm. Single-walled nanotubes (SWNTs) possess the simplest geometry and have been observed with diameters ranging from 0.4 to 3 nm. The formation of a SWNT can be visualized through the rolling of a graphene sheet. Based on the orientation of the tube axis with respect to the hexagonal lattice, the structure of a nanotube can be completely specified through its chiral vector. A SWNT behaves either as a metal or as a semiconductor, depending on its chiral vector [135].

A critical factor for the application of CNTs in solar cells is their purity. To remove the by-products and catalysts, a detailed purification strategy is a necessity in most cases. Multi-step purification procedures using surfactants etc. are described in detail in [136, 137].

SWNTs as well as MWNTs are investigated as electrode material. Both of them can be an ideal channel for collecting and transporting charges across light-harvesting assemblies depending on the priorities by device construction.

The work function of CNTs with intrinsic values of 4.5–5.1 eV are in the same range to the work function of indium tin oxide (4.4–4.9 eV), the mostly used as a transparent electrode in organic photovoltaics [138]. Recent reports have also shown that CNT films can be produced with sheet resistance values as low as 160 V/m^{-2} along with a transmittance of 87 % at 550 nm [139]. These values are still higher than those of ITO films on glass or plastic with sheet resistances as low as 4 and 10 V/m^{-2} , respectively, but still in the necessary conductivity range for photovoltaic applications.

Additionally, the random network properties of CNT films facilitate maximum mechanical flexibility when compared to ITO films deposited on polymer substrates as demonstrated by Parekh et al. [140].

If the homogeneity and transparency of the carbon nanotube thin films is important, SWNTs are better suited than MWNTs. On the other hand, MWNTs are purely metallic, whereas SWNT synthesis always contain metallic and p-type semiconductive SWNTs as product. To obtain much greater precision over their electronic properties semiconducting and metallic SWNTs should be either chemically separated by using selective surfactants [141] or chemically doped. Heavily doping of SWNTs can be established using strong acids (HNO_3 and/or SOCl_2) [142] or inorganic solvent mixtures (hydrofluoric acid, nitric acid and gold (III) chloride hydrate) [143] by the metallization of p-type SWNTs through effective pinning of the Fermi level inside the valence band. It should be also noted that the metallization of SWNTs is a quasi reversible process [144]. Extended testing carried out in air and at room temperature using a protecting layer of poly (3,4-ethylenedioxythiophene)–poly(styrenesulfonate) aqueous solution (PEDOT: PSS, Baytron) on the top of the SWCNT layer show stable sheet resistances for a aging period of 1400 h at 80°C [145]. In 2012, sheet resistances change less than 10 % over 20 days in ambient and less than 2 % with overnight heating to 300°C in air has been achieved by doping SWNTs using MoO_x [146]. Because of the well defined metallic behaviour, MWNTs are mostly investigated in DSCs.

There are several wet-chemistry methods for fabricating isotropic nanotube films, including drop casting from solvents, spin coating, air brushing, vacuum filtering and quasi Langmuir-Blodgett deposition. Electrospinning and gas flow growth of vertically aligned CNTs on suitable substrates are the mostly used methods for oriented CNT films.

The problems to overcome for isotropic CNT films are the poor uniformity, controllable film thickness and the fluctuations due to the van der Waals interactions

between the nanotubes. The formation of bundled nanotubes could adversely affect the transport and recombination in solar cells. To overcome the carbon nanotube clustering formation surfactants such as sodium dodecylbenzenesulfonate, sodium dodecyl sulphate and sodium carboxymethyl cellulose are used [147, 148]. After washing out the excess amount of surfactants, using sodium dodecyl sulphate for example, Kim et al. obtained a sheet resistance of 57 V/sq at 65 % transmittance and a conductivity in excess of 6700 s cm^{-1} —values that are comparable with those of surfactant-free CNT films [148], process of which is given in Fig. 15.29. Additionally, the strong interaction between surfactant and CNTs allows dispersions to be prepared without excessive sonication produces avoiding surface defects and shortening of the tube length [149].

Although the wet-chemistry methods work well with randomly dispersed CNTs, they are not suitable for vertically aligned CNTs. The aligned structure will be damaged in the wet processing because the upper ends of the neighboring nanotubes have been seen to bundle together and cause some nanotubes to lay down.

Usually, the thin film fabrication process of the as-grown aligned CNTs requires a mechanical densification after monolithically removal of CNTs from the substrate. By controlling the extent of compression the CNT density can be tailored. An UV curable epoxy-precursor contacted to the CNT forest diffuses via capillary action and is then cured under well defined preparation parameter [150].

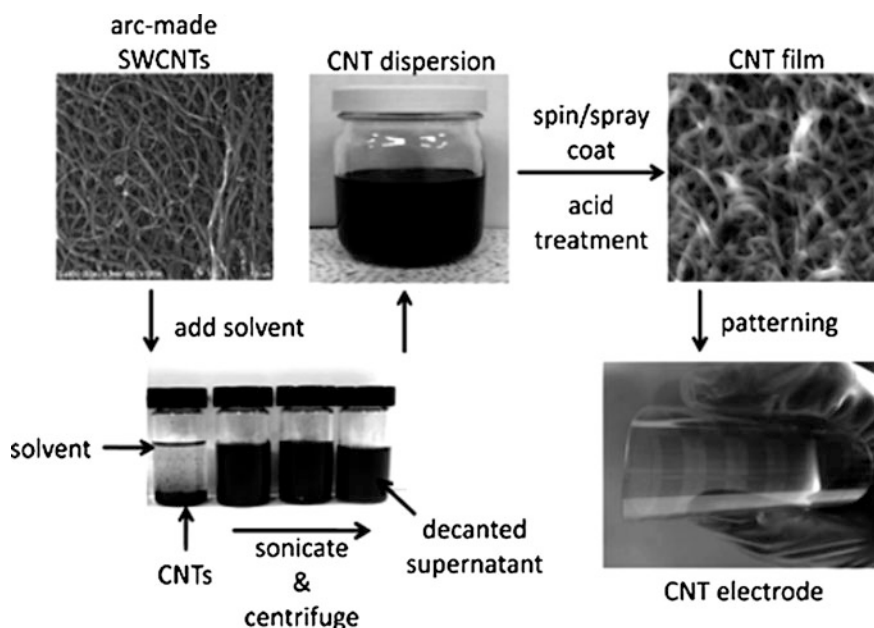


Fig. 15.29 General preparation scheme of CNT thin films as electrodes and its morphology [148]

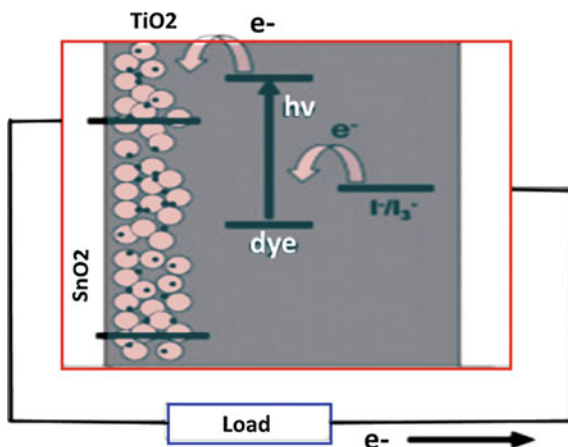
CNTs as Photoanode Material in DSCs

In DSCs, a wide band gap n-type semiconductor material, anatase TiO_2 nanoparticles (generally around 20 nm in diameter) are deposited and interconnected on fluorine-doped tin dioxide FTO (F:SnO_2) coated conducting glass as photo anode, which is attached by dye molecules (sensitizers). The design using mesoporous TiO_2 completely derives from the understanding suggesting that a large interfacial surface area must be averted in solar cell devices in order to enhance the number of dye molecules on the TiO_2 electrode for light harvesting. The electronic transition occurs upon light absorption in the sensitizers, which leads to the formation of an excited state of the dye (exciton). Charge separation occurs at the interface by excited dye injecting photo-induced electrons into the conduction band (CB) of mesoporous TiO_2 working electrode. The injected electrons are afterwards transported through the interconnected mesoporous TiO_2 network towards the FTO substrate of photo anode. After electron injection, the dye molecule in its oxidized form is subsequently regenerated by electron donation from the liquid electrolyte (I^-/I_3^- redox couple) or solid state hole conductor (such as 2,2',7,7'-tetrakis(*N,N*-dimethoxyphenylamine)-9,9'-spiro bifuorene, termed as spiro-OMeTAD [14], P3HT [17] etc.). Meanwhile the holes are transferred via the redox couple towards the counter-electrode (CE) to complete the circuit [151].

The output voltage of a DSC device is determined by the energy difference between the redox potential of the electrolyte and the electron *quasi*-Fermi level (EF) in the TiO_2 under working conditions (Fig. 15.30). The photocurrent output depends on the absorption spectrum of the dye and the absorbed photon to current conversion efficiency, also known as the internal quantum efficiency (IQE).

The intrinsic properties of p-type semiconductor, such as low hole mobility, fast interfacial recombination rate and insufficient infiltration into the mesoporous photo anode are the main limiting factors of the solar cell efficiency.

Fig. 15.30 Schematic operating mechanism of liquid-electrolyte dye-sensitize solar cell based on mesoporous TiO_2 photoanode and I^-/I_3^- redox couple [152]



Using nanoporous materials, a number of the charges are usually trapped by isolated nanoparticles, surface states, or defect states. The so-called trap effects slow down the charge transport through diffusion and limit the device efficiency, which has been also proved by time-resolved photocurrent and photovoltage measurements [153, 154] and modelling studies [155, 156].

An idea to improve the charge collection was realized by integration of metallic carbon nanotubes into the described systems.

Two different strategies were developed:

- (a) Incorporating CNTs into a titanium photoanode
- (b) Core/shell structures incorporating CNT inside TiO_2 nanofiber or decorating CNT walls by TiO_2 nanoparticles.

In Table 15.2, the solar cell parameter of the ‘best’ performing devices together with the references for detailed information are summarized.

A general optimization step for the applications a–b is the necessity to optimize the volume fraction of CNTs in photo anodes to find the optimal balance between improving the electron transport properties by adding CNTs and the dye loading capacity of the titania. It is well recognized that the optimal solar cell efficiencies can be obtained when the dye is chemisorbed in a closely packed monolayer on titania, whereas the chemisorption of multilayers decreases the overall performance. Hereby, the interaction between the hydroxyl group of the dye and titania has a significant importance [155]. Following this concept, MWNTs were oxidized in a hydrogen peroxide (H_2O_2) solution under ultrasonication condition for 24 h at the temperature 50 °C to produce finely dispersed MWNTs terminated with carboxylic acid groups to enhance the adhesion of the titania nanoparticles on the MWNT surface [157]. An overall efficiency of 4.97 % has been achieved by this approach. The efficiency enhancement in comparison to MWNT free reference cell was ca. 49 % (Table 15.2, a1). In another study, using acid treated MWNTs (without OH groups on the side walls) by simply adding into a titania paste and annealing at 150 °C, the efficiency enhancement was only 20 % relative to its MWNT free reference (Table 15.2, a2).

Using the same concept, the efficiency of DSCs was as high as 8.1 % in 2013 (Table 15.2, a3) [159]. Optimizing the photoanode morphology by reducing the grain boundaries using a post-treatment step at 500 °C, hence improving the charge collection and increasing the electron diffusion coefficient was resulted to an improvement of 25 % cell efficiency relative to the efficiency of MWNT free photoanodes.

But the photocurrent or energy conversion efficiency was not monotonously increased with the CNT concentration. An argument is related to the intrinsic feature of the carbon nanostructures, which should be taken into account: As it is well known that carbon species, including carbon black, activated carbon, graphene, as well as CNTs, possesses catalytic activities for conversion of iodine species. For photoelectrode, conversion of triiodide to iodide ions means loss (recombination) of the photogenerated electrons, which is unfavorable to charge collection and should be suppressed.

Table 15.2 The use of carbon nanotubes as photoanode

Photo anode	Device	I–V characteristics	Efficiency	References
0.1 wt% MWNT H ₂ O ₂ treated (a1)	TiO ₂ /MWNT/N719/Pt (liquid electrolyte)	0.63 V/13.5 mA/cm ² / FF: 59.00 %	4.97	[157]
–	TiO ₂ /N719/Pt (liquid electrolyte)	0.60 V/8.5 mA/cm ² / FF: 65.00 %	3.32	
0.1 wt% MWNT Low temperature processing (150 °C) (a2)	TiO ₂ /MWNT/N3/Pt (liquid electrolyte)	0.78 V/9.08 mA/cm ² / FF: 70.80 %	5.02	[158]
–	TiO ₂ /N3/Pt (liquid electrolyte)	0.76 V/7.97 mA/cm ² / FF: 68.50 %	4.15	
0.01 wt% MWNT in TiO ₂ paste/500 °C (a3)	TiO ₂ /MWNT/N719/Pt (liquid electrolyte)	0.72 V/15.6 mA/cm ² / FF: 71.00 %	8.10	[159]
(a3) including light scattering TiO ₂ nanoparticles	TiO ₂ /MWNT/N719/Pt (liquid electrolyte)	0.76 V/16.0 mA/cm ² / FF: 74.00 %	9.00	
0.07 wt% MWNT in TiO ₂ paste/500 °C	TiO ₂ /MWNT/N719/Pt (liquid electrolyte)	0.70 V/11.6 mA/cm ² / FF: 73.00 %	5.90	
–	TiO ₂ /N719/Pt (liquid electrolyte)	0.75 V/12.7 mA/cm ² / FF: 71.0 %	6.50	
TiO ₂ /0.2 MWNT electrospinning rice GRAIN form (b1)	TiO ₂ /MWNT/N3/Pt (liquid electrolyte)	0.82 V/11.98 mA/cm ² / FF: 62.0 %	6.12	[160]
–		0.81 V/9.40 mA/cm ² / FF: 60.0 %	4.61	
TiO ₂ /0.1 wt% MWNT electrospinning elongated nanowires 15 µm layer thickness (b2)	TiO ₂ /MWNT/N7/Pt (liquid electrolyte)	0.75 V/18.53 mA/cm ² / FF: 73.6 %	10.24	[161]
TiO ₂ /0.1 wt% MWNT electrospinning 6 µm layer thickness	TiO ₂ /MWNT/N7/Pt (liquid electrolyte)	0.72 V/11.10 mA/cm ² / FF: 73.0 %	5.84	
TiO ₂ /0.15 wt% MWNT electrospinning 13 µm layer thickness	TiO ₂ /MWNT/N7/Pt (liquid electrolyte)	0.78 V/12.11 mA/cm ² / FF: 75.6 %	7.13	
–	TiO ₂ /N7/Pt (liquid electrolyte)	0.73 V/15.4 mA/cm ² / FF: 50.0 %	5.63	

Additionally, an important parameter was the loss of transparency of the photoanode. For higher CNT concentrations (above 0.01 wt%), the light absorption of the CNTs play a role, resulting in the loss of solar radiation available for photocurrent generation. By using additional titania to enhance the light scattering in the photo active area, the efficiency was as high as 9.0 % (Table 15.2, a3).

Given the MWCNT is incorporated fully inside the titania according to the concept (b), this can rapidly capture and transport the photogenerated electrons and concurrently reduce the undesirable recombination and back-reaction.

Rice-grain shaped TiO_2/MWNT nanocomposite photoanodes were prepared by electrospon with efficiencies of 6.1 % (Table 15.2, b1).

In a similar approach, $\text{TiO}_2/\text{PVP}/\text{MWCNT}$ composite nanofibers were first electrospon at 70 kV on fluorine-doped tin oxide (FTO) glass from a precursor solution that contained titanium isopropoxide, polyvinylpyrrolidone, acetic acid and ethanol. A layer of $\text{TiO}_2/\text{MWCNT}$ nanorod were peeled off from the original FTO glass after calcination and then transferred to another FTO glass with an ultra-thin adhesive TiO_2 paste. The $\text{TiO}_2/\text{MWCNT}$ nanorod photoanode was obtained after calcinating at 450 °C for 2 h [161].

Upon optimization of the parameters for the photoanode, an efficiency as high as 10.24 % has been attained for a photoanode with CNT nanorods that incorporates 0.1 % MWCNTs in the precursor solution. It is worthy to note that the improved efficiency for the photoanode incorporating MWCNTs is largely due to the improved FF, an indication that shows the improvement of charge transport properties.

But also using this concept, the porosity of the titania layer should be high enough enabling sufficient light harvesting by dye molecules. The best solar cell parameters were reported for a mesoporous photo anode with a thickness of ca. 15 μm . Decreasing the layer thickness to 6 μm under the same preparation conditions, the solar cell efficiency was only 5.8 % by constant fill factor. However, upon further increase the amount of MWNT in the precursor solution to a value of 0.15 %, J_{sc} decreased in both sets of photoanode devices, demonstrating the fact that loading of the dye molecules was not sufficient enough in these devices. Despite this reverse trend, the FF kept increasing with increasing the amount of MWNT, an indication suggests that the recombination is also suppressed well in the later case.

CNTs as Counter Electrode Material in DSCs

In DSCs, the redox reaction occur at the interface of electrolyte/counter electrode and electrolyte/dye molecule. The reaction on the counter electrode is the reduction of tri-iodide to iodine. Because of the catalytic activity of CNTs for the reduction of DSC electrolyte tri-iodide [162], there is great effort to replace the costly counter electrode Au/Pt by solution processable cheaper carbon based alternatives [163].

The carbon nanotubes that show results that are closest to Pt are usually single-walled carbon nanotubes, with the MWNTs showing lower overall power conversion efficiency when compared to SWNTs. While there is a clear advantage to using carbon SWNTs in the counter electrodes of DSSCs, there is greater development in the method of mass production of carbon MWNTs, which makes it valuable to investigate methods to improve the reductive properties of MWNTs in the tri-iodide/iodide redox couple in the DSSCs.

The effect of different CNT blends as cathode on the DSC performance is listed in Table 15.3.

Table 15.3 The use of carbon nanotubes as counter electrode materials in dye-sensitized solar cells

Cathode	I–V characteristics	Efficiency (%)	References
SWNTs (c1)	0.68 V/9.51 mA/cm ² /FF: 62.0 %	4.03	[164]
MWNTs (c2)	0.69 V/10.55 mA/cm ² /FF: 60.0 %	4.36	
–	0.63 V/8.11 mA/cm ² /FF: 63.0 %	3.22	
MWNT-blended incarboxymethylcellulose (c3)	0.80 V/11.99 mA/cm ² /FF: 74.0 %	7.13	[165]
–	0.80 V/11.99 mA/cm ² /FF: 68.0 %	6.52	
MWNT (c4)	0.74 V/15.54 mA/cm ² /FF: 57.6 %	6.62	[166]
MWNT/Graphene (c4)	0.74 V/16.04 mA/cm ² /FF: 62.7 %	7.55	
Graphene (c4)	0.75 V/13.09 mA/cm ² /FF: 47.3 %	4.64	
Pt as reference	0.79 V/16.27 mA/cm ² /FF: 68.8 %	8.80	
MWNT (c5)	0.70 V/9.55 mA/cm ² /FF: 52.0 %	3.50	[166]
MWNT-TiN	0.75 V/12.74 mA/cm ² /FF: 57.0 %	5.40	
TiN nanoparticle	0.66 V/9.28 mA/cm ² /FF: 35.0 %	2.10	
Pt as reference	0.73 V/12.83 mA/cm ² /FF: 60 %	5.68	
Vertically aligned CNTs (c6)	0.66 V/13.9 mA/cm ² /FF: 58.0 %	5.20	[167]
Isotropic CNTs	0.61 V/11.8 mA/cm ² /FF: 600 %	3.70	
Pt as reference	0.61 V/17.2 mA/cm ² /FF: 64.0 %	6.60	

In 2010, SWNTs as well as MWNTs were investigated as counter electrode materials with slightly better performance (Table 15.3-c1) for MWNTs in comparison to PtCl₄ electrode [164].

Highly disordered multiwalled CNTs prepared from solution using doctor blade technique were successfully used as counter electrode for DSCs through a simple low-temperature fabrication process. The CNT layer acts hereby as a better catalyst, enhancing the fill factor of the devices (Table 15.3, c3). Variations in electrolyte solutions (an electrolyte containing 1-butyl-3-methylimidazolium instead of Li ion in acetonitrile) demonstrated that electrolytes containing larger molecules perform better using CNT than Pt electrode [165].

The highly catalytic graphene is functionalized onto the carbon MWNTs to support the reduction of the tri-iodide in the redox couple that mediates the charge between the two electrodes of the DSSC [166]. The study of these cells supports the idea that for efficient performance of a carbon-based counter electrode, there needs to be many sharp atomic edges exposed to the ions in the electrolyte for increased catalytic effect. Graphene devices only shows a very poor performance because of forming smooth layers on top of one another, limiting the total surface area for the electrolyte interaction. MWNTs alone, are better suited than graphene, but the best morphology is given by blending the materials together leading to a nanostructured surface area with lowest sheet resistance of the cathode (Table 15.3, c4).

Similarly, titanium nitrides (TiN) with intrinsic electro catalytic activity for the triiodide reduction is blended with MWNTs to form DSC cathodes [168]. A DSC composed of the highly ordered TiN nanotube arrays shows comparable performance with typical Pt counter electrode [169]. However, TiN nanoparticle film electrode alone has lower fill factor (FF) owing to the poor electron transport efficiency across nanoparticles. Furthermore, the introduction of conducting paths of CNTs into TiN can improve electrical conductivity and capacitance of the cathode. The surface of CNTs was functionalized by TiN—by thermal hydrolysis of TiOSO_4 on CNTs. As shown in Table 15.2, c4, the synergic effect of CNTs and TiN is verified. The performance of the TiN-CNT counter electrode is comparable with that of the typical Pt counter electrode. In the TiN anchored CNTs, the superior electrocatalytic activity of TiN nanoparticles is undoubtedly the driving force for the high photovoltaic performance, demonstrated by cyclic voltammetric measurements in the reference paper. The introduction of TiN nanoparticles on CNTs not only decreases the charge-transfer resistance in the electrolyte/electrode interface, but also reduces the diffusion impedance of triiodide ions, leading to the high J_{sc} and FF for DSSCs. Meanwhile, the electron-transport network formed by CNTs also plays an indispensable role in the high photovoltaic performance (Table 15.3, c5).

CNTs have been also blended in conducting polymers such as poly 3,4-ethylenedioxythiophene:polystyrenesulfonate, but the effect of MWNTs was rather insignificant in comparison to PEDOT:PSS only devices [170].

Also vertically aligned carbon nanotubes are investigated as counter electrodes [167]. Ultrathin-alumina coated stainless steel substrates decorated with Fe nanoparticles were used as substrates for the CNT growth. DSCs using vertically aligned CNTs had a better performance than DSCs using randomly dispersed CNTs (without any surface modification). But the DSCs using vertically aligned CNT electrodes do exhibit an increased series resistance compared to traditional Pt cells (C6). The source of this increased resistance has been demonstrated to be the charge transfer resistance at the CNT/electrolyte interface.

Functionalization of the CNT surfaces as well as post-growth surface treatments using inorganic nanoparticles as given in examples above might reduce the activation barrier for the triiodide reduction and thereby reduce the charge transfer resistance.

15.5 Summary and Future Outlook

The huge demand for clean energy can be supplied from the sun due to its abundance and free from contaminations and carbon emissions. However, owing to the high cost, imposed by quantity and quality of potential materials, and inadequate power conversion efficiencies of existing conventional solar cells it is not possible to compete with the energy supplied from fossil fuel sources. To meet this energy requirement, there has been an increasing interest in recent years in both reducing the price and increasing the efficiency of solar cells by use of one-dimensional (1-D)

nanostructures, including nanowires (NWs), quantum wells (QWs), and quantum dots (QDs), in solar cells. 1-D nanostructured solar cells, particularly nanowires and nanopillars based ones, address the main requirement of high-efficient solar cells at lower cost in several ways including high carrier collection efficiency by decoupling the light absorption and carrier collection directions, the ability to grow single-crystalline nanowires on inexpensive substrates without resorting to complex epitaxial routes, and minimization of optical losses (e.g., reducing reflection and enhancing absorption) by light trapping. In the past years, several solar cell device architectures based on nanowires/nanorods/nanopillars have been developed to achieve this goal. In particular, device architectures of nanowire/nanorods array with radial, axial, and nanowire/nanorods array embedded in thin film geometries are the most widely studied ones, all of which offer significant advantages over the traditional planar thin-film or wafer-based solar cells in terms of cost, effective charge separation, strain relaxation and some electrical and optical properties.

Over the last two decades, a considerable amount of literature has been published on one-dimensional nanostructure based solar cells by the use of a wide variety of materials (e.g., poly-crystalline Si, amorphous Si, polymers, II–VI and I–III–VI₂ chalcopyrite (I–III–VI₂) compounds. In this chapter, we mainly focused on several commonly used materials such as silicon, zinc-oxide and titanium-dioxide that we have employed in our recent studies for the fabrication of one-dimensional nanostructures based solar cells as well as the studies reported in literature on the utilization of carbon nanotubes (CNTs) in dye sensitized solar cells. In one of these studies, for instance, we demonstrated a solar cell based on ZnO and Si Nanowires and their integration into a chalcopyrite thin film (AGIS). Photovoltaic properties were determined under AM 1.5G illumination. For the ZnO NWs based solar cell, 0.37 % efficiency was measured, while it was 5.5 % for the Si NWs based solar cell. While at present, the energy conversion efficiencies of n-ZnO-NWs/p-AGIS and n-Si NWs/p-AGIS solar cells are lower than the conventional planar silicon and some Ag-chalcopyrite solar cells, to further improve the efficiency of the device many optimization processes are required such as nanowire growth, thin film layer deposition and contact formation, and surface passivation processes to inhibit surface recombination. It is reasonable to expect that surface passivation alone would increase the energy conversion efficiency of devices to over 5.50 % that we have obtained for n-Si NWs/p-AGIS solar cell. In addition, although the AGIS has been used as the thin film in which Si and ZnO NWs were embedded, the same device architecture can be employed for other potential PV materials (e.g., CdTe, CdS, and CuInSe₂) to enhance the energy conversion efficiency. Moreover, Si NWs transfer printed from the single crystal wafer to any arbitrary substrate through several bottom-up processes may drastically lower the cost of such solar cells.

Carbon nanotube/titania blends function well as transparent, working electrodes and as such could offer improved performance of DSSCs through the reduction of recombination processes by more effective transport of electrons. The most critical factor hereby is to find the balance between the concentration of light harvesting dyes attached onto the titania and the number of carbon nanotube charge transport channels on porous titania electrode. Performance gains up to 9–10 % for DSSCs

employing liquid electrode has been reported using CNT blends in titania. Also as counter electrode, the replacement of Pt by 'potentially' low-cost carbon nanotubes as alternative materials, is progressing well. However, the mass-production technologies have to be further improved to obtain low-cost carbon nanotube materials for large area solar cell applications. The polychirality of nanotubes together with purity requirements reported so far certainly confounds this issue and thus new ideas for the development of highly uniform carbon nanotube mass production has to be realized.

In summary, the use of one-dimensional nanostructures in several type of solar cell device architecture is a potential approach for the realization of cost effective high efficiency solar cells. However, to fully exploit the unique properties of these structures much work is needed. In other words, several fabrication and optimization stages, such as surface passivation to eliminate surface/interface recombination centers arising from large surface/volume ratio, the uniform doping concentration for homo-junction solar cells, appropriate material systems to meet lattice-matching for hetero-junction solar cells, chemical and mechanical stability, and some practical issues (e.g., rapid scaling, device packaging, and integration into device modules) are required to realize one-dimensional nanostructures based solar cells with efficiencies comparable to those achieved with their planar counterparts at much lower production cost.

References

1. Energy BSRoW, Technical Report, BP (2010)
2. P. Würfel, *Frontmatter*. Physics of solar cells (Wiley-VCH Verlag GmbH, 2007), pp. I–XII. doi:[10.1002/9783527618545.fmatter](https://doi.org/10.1002/9783527618545.fmatter)
3. J.S. Li, H.Y. Yu, Y.L. Li, Aligned Si nanowire-based solar cells. *Nanoscale* **3**(12), 4888–4900 (2011)
4. Y.N. Xia, P.D. Yang, Y.G. Sun, Y.Y. Wu, B. Mayers, B. Gates, Y.D. Yin, F. Kim, Y.Q. Yan, One-dimensional nanostructures: synthesis, characterization, and applications. *Adv. Mater.* **15**(5), 353–389 (2003)
5. Z.Y. Fan, D.J. Ruebusch, A.A. Rathore, R. Kapadia, O. Ergen, P.W. Leu, A. Javey, Challenges and prospects of nanopillar-based solar cells. *Nano Res* **2**(11), 829–843 (2009)
6. P.V. Kamat, Meeting the clean energy demand: nanostructure architectures for solar energy conversion. *J. Phys. Chem. C* **111**(7), 2834–2860 (2007)
7. M. Law, L.E. Greene, J.C. Johnson, R. Saykally, P.D. Yang, Nanowire dye-sensitized solar cells. *Nat. Mater.* **4**(6), 455–459 (2005)
8. J.A. Czaban, D.A. Thompson, R.R. LaPierre, GaAs core-shell nanowires for photovoltaic applications. *Nano Lett.* **9**(1), 148–154 (2009)
9. E.C. Garnett, P.D. Yang, Silicon nanowire radial p-n junction solar cells. *J. Am. Chem. Soc.* **130**(29), 9224–9225 (2008)
10. E.C. Garnett, M.L. Brongersma, Y. Cui, M.D. McGehee, Nanowire solar cells. *Annu. Rev. Mater. Res.* **41**, 269–295 (2011)
11. R. Kapadia, Z.Y. Fan, K. Takei, A. Javey, Nanopillar photovoltaics: materials, processes, and devices. *Nano Energy* **1**(1), 132–144 (2012)
12. Three-Dimensional Nano Architectures: Designing Next Generation Devices (Springer, New York, London, 2011). doi:[10.1007/978-1-4419-9822-4](https://doi.org/10.1007/978-1-4419-9822-4)

13. B.Z. Tian, X.L. Zheng, T.J. Kempa, Y. Fang, N.F. Yu, G.H. Yu, J.L. Huang, C.M. Lieber, Coaxial silicon nanowires as solar cells and nanoelectronic power sources. *Nature* **449**(7164), 885–890 (2007)
14. Y. Zhang, L.W. Wang, A. Mascarenhas, Quantum coaxial cables for solar energy harvesting. *Nano Lett.* **7**(5), 1264–1269 (2007)
15. B.M. Kayes, H.A. Atwater, N.S. Lewis, Comparison of the device physics principles of planar and radial p-n junction nanorod solar cells. *J. Appl. Phys.* **97**(11), 114302–114313 (2005)
16. L. Tsakalakos, J. Balch, J. Fronheiser, B.A. Korevaar, O. Sulima, J. Rand, Silicon nanowire solar cells. *Appl Phys Lett* **91**(23) (2007)
17. B.D. Yuhas, P.D. Yang, Nanowire-based all-oxide solar cells. *J. Am. Chem. Soc.* **131**(10), 3756–3761 (2009)
18. Z.Y. Fan, H. Razavi, J.W. Do, A. Moriwaki, O. Ergen, Y.L. Chueh, P.W. Leu, J.C. Ho, T. Takahashi, L.A. Reichertz, S. Neale, K. Yu, M. Wu, J.W. Ager, A. Javey, Three-dimensional nanopillar-array photovoltaics on low-cost and flexible substrates. *Nat. Mater.* **8**(8), 648–653 (2009)
19. R.R. LaPierre, Numerical model of current-voltage characteristics and efficiency of GaAs nanowire solar cells. *J. Appl. Phys.* **109**(3), 034311–034316 (2011)
20. C. Colombo, M. Heiss, M. Gratzel, A.F.I. Morral, Gallium arsenide p-i-n radial structures for photovoltaic applications. *Appl. Phys. Lett.* **94**(17), 173108 (2009)
21. Y.B. Tang, Z.H. Chen, H.S. Song, C.S. Lee, H.T. Cong, H.M. Cheng, W.J. Zhang, I. Bello, S.T. Lee, Vertically aligned p-type single-crystalline gan nanorod arrays on n-type si for heterojunction photovoltaic cells. *Nano Lett.* **8**(12), 4191–4195 (2008)
22. V. Sivakov, G. Andra, A. Gawlik, A. Berger, J. Plentz, F. Falk, S.H. Christiansen, Silicon nanowire-based solar cells on glass: synthesis, optical properties, and cell parameters. *Nano Lett.* **9**(4), 1549–1554 (2009)
23. M.Q. Yao, N.F. Huang, S. Cong, C.Y. Chi, M.A. Seyed, Y.T. Lin, Y. Cao, M.L. Povinelli, P.D. Dapkus, C.W. Zhou, GaAs nanowire array solar cells with axial p-i-n junctions. *Nano Lett.* **14**(6), 3293–3303 (2014)
24. A. Mews, *Nanomaterials handbook*, edited by Yury Gogotsi. *Angew. Chem. Int. Ed.* **46**(13), 2143 (2007). doi:[10.1002/anie.200685445](https://doi.org/10.1002/anie.200685445)
25. L.N. Dem'yanets, L.E. Li, T.G. Uvarova, Zinc oxide: hydrothermal growth of nano- and bulk crystals and their luminescent properties. *J. Mater. Sci.* **41**(5), 1439–1444 (2006). doi:[10.1007/s10853-006-7457-z](https://doi.org/10.1007/s10853-006-7457-z)
26. H.-E. Wang, Z. Chen, Y.H. Leung, C. Luan, C. Liu, Y. Tang, C. Yan, W. Zhang, J.A. Zapien, I. Bello, S.-T. Lee, Hydrothermal synthesis of ordered single-crystalline rutile TiO₂ nanorod arrays on different substrates. *Appl. Phys. Lett.* **96**(26), 263104 (2010). doi:[10.1063/1.3442913](https://doi.org/10.1063/1.3442913)
27. X. Wang, Y. Li, Selected-control hydrothermal synthesis of α - and β -MnO₂ single crystal nanowires. *J. Am. Chem. Soc.* **124**(12), 2880–2881 (2002). doi:[10.1021/ja0177105](https://doi.org/10.1021/ja0177105)
28. H. Zhitao, L. Sisi, C. Jinkui, C. Yong, Controlled growth of well-aligned ZnO nanowire arrays using the improved hydrothermal method. *J. Semiconduct.* **34**(6), 063002 (2013)
29. M. Ahmad, M.A. Iqbal, J. Kiely, R. Luxton, M. Jabeen, Low temperature hydrothermal synthesis of ZnO nanowires for nanogenerator: effect of gold electrode on the output voltage of nanogenerator. *Indian J. Eng. Mater. S* **21**(6), 672–676 (2014)
30. S.N. Bai, S.C. Wu, Synthesis of ZnO nanowires by the hydrothermal method, using sol-gel prepared ZnO seed films. *J. Mater. Sci-Mater El* **22**(4), 339–344 (2011)
31. R. Hao, X. Deng, Y.B. Yang, D.Y. Chen, Research progress in preparation and applications of ZnO nanowire/rod arrays by hydrothermal method. *Acta Chim. Sinica* **72**(12), 1199–1208 (2014)
32. H.S. Jang, B. Son, H. Song, G.Y. Jung, H.C. Ko, Controlled hydrothermal growth of multi-length-scale ZnO nanowires using liquid masking layers. *J. Mater. Sci.* **49**(23), 8000–8009 (2014)

33. H. Karaagac, M. Parlak, E. Yengel, M.S. Islam, Heterojunction solar cells with integrated Si and ZnO nanowires and a chalcopyrite thin film. *Mater. Chem. Phys.* **140**(1), 382–390 (2013)
34. D.P. Neveling, T.S. van den Heever, R. Bucher, W.J. Perold, L.M.T. Dicks, Effect of seed layer deposition, au film layer thickness and crystal orientation on the synthesis of hydrothermally grown ZnO nanowires. *Curr. Nanosci.* **10**(6), 827–836 (2014)
35. I.J. No, S. Lee, S.H. Kim, J.W. Cho, P.K. Shin, Morphology control of ZnO nanowires grown by hydrothermal methods using Au nanodots on Al doped ZnO seed layer. *Jpn. J. Appl. Phys.* **52**(2), 025003 (2013)
36. Y.K. Tseng, M.C. Hung, S.L. Su, S.K. Li, Using the hydrothermal method to grow p-type ZnO nanowires on Al-doped ZnO thin film to fabricate a homojunction diode. *J. Nanosci. Nanotechnol.* **14**(10), 7907–7910 (2014)
37. H. Karaagac, V.J. Logeeswaran, M.S. Islam, Fabrication of 3D-silicon micropillars/walls decorated with aluminum-ZnO/ZnO nanowires for optoelectric devices. *Phys. Status Solidi A* **210**(7), 1377–1380 (2013)
38. K. Peng, Y. Xu, Y. Wu, Y. Yan, S.-T. Lee, J. Zhu, Aligned single-crystalline Si nanowire arrays for photovoltaic applications. *Small* **1**(11), 1062–1067 (2005). doi:[10.1002/sml.200500137](https://doi.org/10.1002/sml.200500137)
39. K.J. Morton, G. Nieberg, S.F. Bai, S.Y. Chou, Wafer-scale patterning of sub-40 nm diameter and high aspect ratio (>50:1) silicon pillar arrays by nanoimprint and etching. *Nanotechnology* **19**(34), 345301 (2008)
40. Z. Huang, N. Geyer, P. Werner, J. de Boor, U. Gösele, Metal-assisted chemical etching of silicon: a review. *Adv. Mater.* **23**(2), 285–308 (2011). doi:[10.1002/adma.201001784](https://doi.org/10.1002/adma.201001784)
41. H. Karaagac, M.S. Islam, Enhanced field ionization enabled by metal induced surface states on semiconductor nanotips. *Adv. Funct. Mater.* **24**(15), 2224–2232 (2014)
42. M. Meyyappan MS, *Inorganic Nanowires: Applications, Properties and Characterization*. (CRC Press, 2010)
43. B. Gates, B. Mayers, A. Grossman, Y. Xia, A sonochemical approach to the synthesis of crystalline selenium nanowires in solutions and on solid supports. *Adv. Mater.* **14**(23), 1749–1752 (2002)
44. R.V. Kumar, Y. Koltypin, X.N. Xu, Y. Yeshurun, A. Gedanken, I. Felner, Fabrication of magnetite nanorods by ultrasound irradiation. *J. Appl. Phys.* **89**(11), 6324–6328 (2001)
45. A.P. Nayak, A.M. Katzenmeyer, J.-Y. Kim, M.K. Kwon, Y. Goshō, M. Saif Islam, Purely sonochemical route for oriented zinc oxide nanowire growth on arbitrary substrate. *Proc. SPIE* **7683**, 738312. doi:[10.1117/12.851755](https://doi.org/10.1117/12.851755)
46. Y.Y. Wu, P.D. Yang, Direct observation of vapor-liquid-solid nanowire growth. *J. Am. Chem. Soc.* **123**(13), 3165–3166 (2001)
47. M.H. Huang, S. Mao, H. Feick, H.Q. Yan, Y.Y. Wu, H. Kind, E. Weber, R. Russo, P.D. Yang, Room-temperature ultraviolet nanowire nanolasers. *Science* **292**(5523), 1897–1899 (2001)
48. W. Lu, C.M. Lieber, Semiconductor nanowires. *J. Phys. D Appl. Phys.* **39**(21), R387–R406 (2006)
49. S. Han, R.S. Wagner, Grain boundary effects on carrier transport in undoped polycrystalline chemical-vapor-deposited diamond. *Appl. Phys. Lett.* **68**(21), 3016–3018 (1996)
50. C.C. Chen, C.C. Yeh, Large-scale catalytic synthesis of crystalline gallium nitride nanowires. *Adv. Mater.* **12**(10), 738 (2000)
51. Y.J. Chen, J.B. Li, Y.S. Han, X.Z. Yang, J.H. Dai, The effect of Mg vapor source on the formation of MgO whiskers and sheets. *J. Cryst. Growth* **245**(1–2), 163–170 (2002)
52. X.F. Duan, C.M. Lieber, General synthesis of compound semiconductor nanowires. *Adv. Mater.* **12**(4), 298–302 (2000)
53. Y.W. Wang, L.D. Zhang, C.H. Liang, G.Z. Wang, X.S. Peng, Catalytic growth and photoluminescence properties of semiconductor single-crystal ZnS nanowires. *Chem. Phys. Lett.* **357**(3–4), 314–318 (2002)

54. J. Zhang, X.S. Peng, X.F. Wang, Y.W. Wang, L.D. Zhang, Micro-Raman investigation of GaN nanowires prepared by direct reaction Ga with NH_3 . *Chem. Phys. Lett.* **345**(5–6), 372–376 (2001). doi:[10.1016/S0009-2614\(01\)00905-8](https://doi.org/10.1016/S0009-2614(01)00905-8)
55. M. Triplett, H. Nishimura, M. Ombaba, V.J. Logeeswarren, M. Yee, K.G. Polat, J.Y. Oh, T. Fuyuki, F. Leonard, M.S. Islam, High-precision transfer-printing and integration of vertically oriented semiconductor arrays for flexible device fabrication. *Nano Res.* **7**(7), 998–1006 (2014)
56. C.N.R. Rao, F.L. Deepak, G. Gundiah, A. Govindaraj, Inorganic nanowires. *Prog. Solid State Ch* **31**(1–2), 5–147 (2003)
57. P.L. Dong, X.D. Wang, M. Zhang, M. Guo, S. Seetharaman, The preparation and characterization of beta-SiAlON nanostructure whiskers. *J. Nanomater.* **2008**, 282187–282192 (2008)
58. Y.J. Hsu, S.Y. Lu, Vapor-solid growth of Sn nanowires: growth mechanism and superconductivity. *J. Phys. Chem. B* **109**(10), 4398–4403 (2005)
59. Z.W. Pan, Z.R. Dai, Z.L. Wang, Nanobelts of semiconducting oxides. *Science* **291**(5510), 1947–1949 (2001)
60. Y.J. Zhang, N.L. Wang, S.P. Gao, R.R. He, S. Miao, J. Liu, J. Zhu, X. Zhang, A simple method to synthesize nanowires. *Chem. Mater.* **14**(8), 3564–3568 (2002)
61. S.S. Fan, M.G. Chapline, N.R. Franklin, T.W. Tombler, A.M. Cassell, H.J. Dai, Self-oriented regular arrays of carbon nanotubes and their field emission properties. *Science* **283**(5401), 512–514 (1999)
62. S. Sakurai, M. Inaguma, D.N. Futaba, M. Yumura, K. Hata, A fundamental limitation of small diameter single-walled carbon nanotube synthesis—a scaling rule of the carbon nanotube yield with catalyst volume. *Materials* **6**(7), 2633–2641 (2013)
63. M. Xu, D.N. Futaba, M. Yumura, K. Hata, Alignment control of carbon nanotube forest from random to nearly perfectly aligned by utilizing the crowding effect. *ACS Nano* **6**(7), 5837–5844 (2012)
64. H. Karaagac, M. Kaleli, M. Parlak, Characterization of $\text{AgGa}_{0.5}\text{In}_{0.5}\text{Se}_2$ thin films deposited by electron-beam technique. *J. Phys. D-Appl. Phys.* **42**(16) (2009). doi:Artn 165413; doi:[10.1088/0022-3727/42/16/165413](https://doi.org/10.1088/0022-3727/42/16/165413)
65. I. Repins, M.A. Contreras, B. Egaas, C. DeHart, J. Scharf, C.L. Perkins, B. To, R. Noufi, 19.9 %-efficient $\text{ZnO}/\text{CdS}/\text{CuInGaSe}_2$ solar cell with 81.2 % fill factor. *Prog. Photovolt.* **16** (3), 235–239 (2008). doi:[10.1002/Pip.822](https://doi.org/10.1002/Pip.822)
66. K. Yamada, N. Hoshino, T. Nakada, Crystallographic and electrical properties of wide gap $\text{Ag}(\text{In}_{1-x}\text{Ga}_x)\text{Se}_2$ thin films and solar cells. *Sci. Technol. Adv. Mat.* **7**(1), 42–45 (2006). doi:[10.1016/j.stam.2005.11.016](https://doi.org/10.1016/j.stam.2005.11.016)
67. P.P. Ramesh, O.M. Hussain, S. Uthanna, B.S. Naidu, P.J. Reddy, Photovoltaic performance of p- $\text{AgInSe}_2/\text{n-CdS}$ thin film heterojunctions. *Mater. Lett.* **34**(3–6), 217–221 (1998)
68. Y.S. Murthy, O.M. Hussain, B.S. Naidu, P.J. Reddy, Characterization of p- $\text{AgInSe}_2/\text{n-CdS}$ thin-film heterojunction. *Mater. Lett.* **10**(11–12), 504–508 (1991)
69. G.H. Chandra, O.M. Hussain, S. Uthanna, B.S. Naidu, Characterization of p- $\text{AgGa}_{0.25}\text{In}_{0.75}\text{Se}_2/\text{n-Zn}_{0.35}\text{Cd}_{0.65}\text{S}$ polycrystalline thin film heterojunctions. *Mat. Sci. Eng. B-Solid* **86**(1), 60–63 (2001)
70. B. Ozdemir, M. Kulakci, R. Turan, H.E. Unalan, Effect of electroless etching parameters on the growth and reflection properties of silicon nanowires. *Nanotechnology* **22**(15), 155606 (2011). doi:Artn 155606; doi:[10.1088/0957-4484/22/15/155606](https://doi.org/10.1088/0957-4484/22/15/155606)
71. K.Q. Peng, J.J. Hu, Y.J. Yan, Y. Wu, H. Fang, Y. Xu, S.T. Lee, J. Zhu, Fabrication of single-crystalline silicon nanowires by scratching a silicon surface with catalytic metal particles. *Adv. Funct. Mater.* **16**(3), 387–394 (2006). doi:[10.1002/adfm.200500392](https://doi.org/10.1002/adfm.200500392)
72. O. Gunawan, S. Guha, Characteristics of vapor-liquid-solid grown silicon nanowire solar cells. *Sol. Energy Mat. Sol. C* **93**(8), 1388–1393 (2009). doi:[10.1016/j.solmat.2009.02.024](https://doi.org/10.1016/j.solmat.2009.02.024)
73. D.R. Kim, C.H. Lee, P.M. Rao, I.S. Cho, X.L. Zheng, Hybrid Si microwire and planar solar cells: passivation and characterization. *Nano Lett.* **11**(7), 2704–2708 (2011). doi:[10.1021/NI2009636](https://doi.org/10.1021/NI2009636)

74. V.J. Logeeswaran, A.M. Katzenmeyer, M.S. Islam, Harvesting and transferring vertical pillar arrays of single-crystal semiconductor devices to arbitrary substrates. *IEEE T Electron. Dev.* **57**(8), 1856–1864 (2010). doi:[10.1109/Ted.2010.2051195](https://doi.org/10.1109/Ted.2010.2051195)
75. M.M. Ombaba, L.V. Jayaraman, M.S. Islam, Precision stress localization during mechanical harvesting of vertically oriented semiconductor micro- and nanostructure arrays. *Appl. Phys. Lett.* **104**(24), 243109 (2014)
76. H. Bi, R.R. LaPierre, A GaAs nanowire/P3HT hybrid photovoltaic device. *Nanotechnology* **20**(46), 465205 (2009). doi:Artn 465205; doi:[10.1088/0957-4484/20/46/465205](https://doi.org/10.1088/0957-4484/20/46/465205)
77. J. Davenas, S.B. Dkhil, D. Cornu, A. Rybak, Silicon nanowire/P3HT hybrid solar cells: effect of the electron localization at wire nanodiameters. *Energy Proc.* **31**, 136–143 (2012). doi:[10.1016/j.egypro.2012.11.175](https://doi.org/10.1016/j.egypro.2012.11.175)
78. B. Eisenhawer, S. Sensfuss, V. Sivakov, M. Pietsch, G. Andra, F. Falk, Increasing the efficiency of polymer solar cells by silicon nanowires. *Nanotechnology* **22**(31), 315401 (2011). doi:Artn 315401; doi:[10.1088/0957-4484/22/31/315401](https://doi.org/10.1088/0957-4484/22/31/315401)
79. L.N. He, C.Y. Jiang, H. Wang, D. Lai, Rusli, Si nanowires organic semiconductor hybrid heterojunction solar cells toward 10 % efficiency. *ACS Appl. Mater. Inter.* **4**(3), 1704–1708 (2012). doi:[10.1021/Am201838y](https://doi.org/10.1021/Am201838y)
80. Y. Kang, D. Kim, Well-aligned CdS nanorod conjugated polymer solar cells. *Sol. Energy Mat. Sol. C* **90**(2), 166–174 (2006). doi:[10.1016/j.solmat.2005.03.001](https://doi.org/10.1016/j.solmat.2005.03.001)
81. J.S. Huang, C.Y. Hsiao, S.J. Syu, J.J. Chao, C.F. Lin, Well-aligned single-crystalline silicon nanowire hybrid solar cells on glass. *Sol. Energy Mat. Sol. C* **93**(5), 621–624 (2009). doi:[10.1016/j.solmat.2008.12.016](https://doi.org/10.1016/j.solmat.2008.12.016)
82. H. Karaagac, A hybrid solar cell based on silicon nanowire and organic thin film. *Physica Status solidi (a)* **211**(11), 2503–2508 (2014). doi:[10.1002/pssa.201431320](https://doi.org/10.1002/pssa.201431320)
83. G.J. Matt, T. Fromherz, M. Bednorz, S. Zamiri, G. Goncalves, C. Lungenschmied, D. Meissner, H. Sitter, N.S. Sariciftci, C.J. Brabec, G. Bauer, Fullerene sensitized silicon for near-to mid-infrared light detection. *Adv. Mater.* **22**(5), 647 (2010). doi:[10.1002/adma.200901383](https://doi.org/10.1002/adma.200901383)
84. J. Bae et al., Si nanowire metal–insulator–semiconductor photodetectors as efficient light harvesters. *Nanotechnology* **21**(9), 095502 (2010)
85. Y. Cui, Q. Wei, H. Park, C.M. Lieber, Nanowire nanosensors for highly sensitive and selective detection of biological and chemical species. *Science* **293**(5533), 1289–1292 (2001). doi:[10.1126/science.1062711](https://doi.org/10.1126/science.1062711)
86. T.C.H. Yamada, S. Ishidac, Y. Arakawac, Si-nanowire optical waveguide devices for optical communications. *Proc. SPIE* **6019**, 60192X (2005)
87. L. Hu, G. Chen, Analysis of optical absorption in silicon nanowire arrays for photovoltaic applications. *Nano Lett.* **7**(11), 3249–3252 (2007). doi:[10.1021/nl071018b](https://doi.org/10.1021/nl071018b)
88. L. Liao, H.B. Lu, M. Shuai, J.C. Li, Y.L. Liu, C. Liu, Z.X. Shen, T. Yu, A novel gas sensor based on field ionization from ZnO nanowires: moderate working voltage and high stability. *Nanotechnology* **19**(17), 175501 (2008). doi:Artn 175501; doi:[10.1088/0957-4484/19/17/175501](https://doi.org/10.1088/0957-4484/19/17/175501)
89. F. Qian, S. Gradečak, Y. Li, C.-Y. Wen, C.M. Lieber, Core/multishell nanowire heterostructures as multicolor, high-efficiency light-emitting diodes. *Nano Lett.* **5**(11), 2287–2291 (2005). doi:[10.1021/nl051689e](https://doi.org/10.1021/nl051689e)
90. V. Schmidt, H. Riel, S. Senz, S. Karg, W. Riess, U. Gösele, Realization of a silicon nanowire vertical surround-gate field-effect transistor. *Small* **2**(1), 85–88 (2006). doi:[10.1002/sml.200500181](https://doi.org/10.1002/sml.200500181)
91. P.D. Li, C.M. Sun, T.G. Jiu, G.J. Wang, J. Li, X.F. Li, J.F. Fangt, High-performance inverted solar cells based on blend films of ZnO nanoparticles and TiO₂ nanorods as a cathode buffer layer. *ACS Appl. Mater. Inter.* **6**(6), 4074–4080 (2014)
92. J.P. Liu, S.S. Wang, Z.Q. Bian, M. Shan, C.H. Huang, Organic/inorganic hybrid solar cells with vertically oriented ZnO nanowires. *Appl. Phys. Lett.* **94**(17), 173107 (2009)

93. O. Lupan, V.M. Guerin, I.M. Tiginyanu, V.V. Ursaki, L. Chow, H. Heinrich, T. Pauporte, Well-aligned arrays of vertically oriented ZnO nanowires electrodeposited on ITO-coated glass and their integration in dye sensitized solar cells. *J. Photoch. Photobio A* **211**(1), 65–73 (2010)
94. V. Strano, E. Smecca, V. Depauw, C. Trompoukis, A. Alberti, R. Reitano, I. Crupi, I. Gordon, S. Mirabella, Low-cost high-haze films based on ZnO nanorods for light scattering in thin c-Si solar cells. *Appl. Phys. Lett.* **106**(1), 013901 (2015)
95. D.I. Suh, S.Y. Lee, T.H. Kim, J.M. Chun, E.K. Suh, O.B. Yang, S.K. Lee, The fabrication and characterization of dye-sensitized solar cells with a branched structure of ZnO nanowires. *Chem. Phys. Lett.* **442**(4–6), 348–353 (2007)
96. M.T. Tsai, Z.P. Yang, T.S. Jing, H.H. Hsieh, Y.C. Yao, T.Y. Lin, Y.F. Chen, Y.J. Lee, Achieving graded refractive index by use of ZnO nanorods/TiO₂ layer to enhance omnidirectional photovoltaic performances of InGaP/GaAs/Ge triple-junction solar cells. *Sol. Energy Mat. Sol C* **136**, 17–24 (2015)
97. J. Zhang, W.X. Que, P. Zhong, G.Q. Zhu, p-Cu₂O/n-ZnO nanowires on ITO glass for solar cells. *J. Nanosci. Nanotechnol.* **10**(11), 7473–7476 (2010)
98. Y.F. Zhu, W.Z. Shen, Synthesis of ZnO nanoplates decorated rhombus-shaped ZnO nanorods and their application in solar cells. *Physica E* **59**, 110–116 (2014)
99. L.W. Ji, S.M. Peng, J.S. Wu, W.S. Shih, C.Z. Wu, I.T. Tang, Effect of seed layer on the growth of well-aligned ZnO nanowires. *J. Phys. Chem. Solids* **70**(10), 1359–1362 (2009). doi:[10.1016/j.jpcs.2009.07.029](https://doi.org/10.1016/j.jpcs.2009.07.029)
100. L. De Marco, M. Manca, R. Giannuzzi, F. Malara, G. Melcarne, G. Ciccarella, I. Zama, R. Cingolani, G. Gigli, Novel preparation method of TiO₂-nanorod-based photoelectrodes for dye-sensitized solar cells with improved light-harvesting efficiency. *J. Phys. Chem. C* **114**(9), 4228–4236 (2010). doi:[10.1021/jp910346d](https://doi.org/10.1021/jp910346d)
101. I.S. Cho, Z.B. Chen, A.J. Forman, D.R. Kim, P.M. Rao, T.F. Jaramillo, X.L. Zheng, Branched TiO₂ nanorods for photoelectrochemical hydrogen production. *Nano Lett.* **11**(11), 4978–4984 (2011). doi:[10.1021/Nl2029392](https://doi.org/10.1021/Nl2029392)
102. J.B. Chen, C.W. Wang, Y.M. Kang, D.S. Li, W.D. Zhu, F. Zhou, Investigation of temperature-dependent field emission from single crystal TiO₂ nanorods. *Appl. Surf. Sci.* **258**(20), 8279–8282 (2012). doi:[10.1016/j.apsusc.2012.05.037](https://doi.org/10.1016/j.apsusc.2012.05.037)
103. H.W. Lin, Y.H. Chang, C. Chen, Facile fabrication of TiO₂ nanorod arrays for gas sensing using double-layered anodic oxidation method. *J. Electrochem. Soc.* **159**(1), K5–K9 (2012). doi:[10.1149/2.013201jes](https://doi.org/10.1149/2.013201jes)
104. J.K. Chen, W.Y. Fu, G.Y. Yuan, A. Runa, H. Bala, X.D. Wang, G. Sun, J.L. Cao, Z.Y. Zhang, Fabrication of TiO₂ nanocrystals/nanorods composites thin film electrode: Enhanced performance of dye-sensitized solar cells. *Mater. Lett.* **135**, 229–232 (2014)
105. K. Fan, W. Zhang, T.Y. Peng, J.N. Chen, F. Yang, Application of TiO₂ fusiform nanorods for dye-sensitized solar cells with significantly improved efficiency. *J. Phys. Chem. C* **115**(34), 17213–17219 (2011)
106. Z.M. He, J. Liu, J.W. Miao, B. Liu, T.T.Y. Tan, A one-pot solvothermal synthesis of hierarchical microspheres with radially assembled single-crystalline TiO₂-nanorods for high performance dye-sensitized solar cells. *J. Mater. Chem. C* **2**(8), 1381–1385 (2014)
107. Y.H. Jung, K.H. Park, J.S. Oh, D.H. Kim, C.K. Hong, Effect of TiO₂ rutile nanorods on the photoelectrodes of dye-sensitized solar cells. *Nanoscale Res. Lett.* **8**, 37 (2013)
108. S.H. Kang, Thickness effect of single crystalline TiO₂ nanorods for dye-sensitized solar cells. *J. Nanosci. Nanotechnol.* **14**(8), 6318–6321 (2014)
109. S. Kathirvel, C.C. Su, H.C. Lin, B.R. Chen, W.R. Li, Facile non-hydrolytic solvothermal synthesis of one dimensional TiO₂ nanorods for efficient dye-sensitized solar cells. *Mater. Lett.* **129**, 149–152 (2014)
110. P.L. Kuo, T.S. Jan, C.H. Liao, C.C. Chen, K.M. Lee, Syntheses of size-varied nanorods TiO₂ and blending effects on efficiency for dye-sensitized solar cells. *J. Power Sources* **235**, 297–302 (2013)

111. J. Liu, J. Luo, W.G. Yang, Y.L. Wang, L.Y. Zhu, Y.Y. Xu, Y. Tang, Y.J. Hu, C. Wang, Y.G. Chen, W.M. Shi, Synthesis of single-crystalline anatase TiO_2 nanorods with high-performance dye-sensitized solar cells. *J. Mater. Sci. Technol.* **31**(1), 106–109 (2015)
112. Y.D. Park, K. Anabuki, S. Kim, K.W. Park, D.H. Lee, S.H. Um, J. Kim, J.H. Cho, Fabrication of stable electrospun TiO_2 nanorods for high-performance dye-sensitized solar cells. *Macromol. Res.* **21**(6), 636–640 (2013)
113. M.K. Wang, J. Bai, F. Le Formal, S.J. Moon, L. Cevey-Ha, R. Humphry-Baker, C. Gratzel, S. M. Zakeeruddin, M. Gratzel, Solid-state dye-sensitized solar cells using ordered TiO_2 nanorods on transparent conductive oxide as photoanodes. *J. Phys. Chem. C* **116**(5), 3266–3273 (2012)
114. Y.L. Xie, P.C. Lin, S.Q. Hu, Y.C. Lu, L. Li, H. Wang, Growth of ZnO nanorods on TiO_2 nanoparticles films and their application to the electrode of dye-sensitized solar cells. *J. Mater. Sci-Mater El* **25**(6), 2665–2670 (2014)
115. W.J. Zhang, Y. Xie, D.H. Xiong, X.W. Zeng, Z.H. Li, M.K. Wang, Y.B. Cheng, W. Chen, K.Y. Yan, S.H. Yang, TiO_2 nanorods: a facile size- and shape-tunable synthesis and effective improvement of charge collection kinetics for dye-sensitized solar cells. *ACS Appl. Mater. Inter.* **6**(12), 9698–9704 (2014)
116. B.W. Luo, Y. Deng, Y. Wang, Z.W. Zhang, M. Tan, Heterogeneous flammulina velutipes-like CdTe/TiO_2 nanorod array: A promising composite nanostructure for solar cell application. *J. Alloy. Compd.* **517**, 192–197 (2012). doi:[10.1016/j.jallcom.2011.12.090](https://doi.org/10.1016/j.jallcom.2011.12.090)
117. H.B. Zhang, M.J. Zhang, C.B. Tian, N. Li, P. Lin, Z.H. Li, S.W. Du, An effective method for the synthesis of 3D inorganic Ln(III)-K(I) sulfate open frameworks with unusually high thermal stability: in situ generation of sulfate anions. *J. Mater. Chem.* **22**(14), 6831–6837 (2012). doi:[10.1039/C2jm16779d](https://doi.org/10.1039/C2jm16779d)
118. M. Ghaffari, M.B. Cosar, H.I. Yavuz, M. Ozenbas, A.K. Okyay, Effect of Au nano-particles on TiO_2 nanorod electrode in dye-sensitized solar cells. *Electrochim. Acta* **76**, 446–452 (2012). doi:[10.1016/j.electacta.2012.05.058](https://doi.org/10.1016/j.electacta.2012.05.058)
119. H. Karaagac, L.E. Aygun, M. Parlak, M. Ghaffari, N. Biyikli, A.K. Okyay, Au/ TiO_2 nanorod-based schottky-type UV photodetectors. *Phys. Status Solidi-R* **6**(11), 442–444 (2012). doi:[10.1002/pssr.201206379](https://doi.org/10.1002/pssr.201206379)
120. J. Tang, Z. Huo, S. Brittman, H. Gao, P. Yang, Solution-processed core-shell nanowires for efficient photovoltaic cells. *Nat. Nano* **6**(9), 568–572 (2011)
121. Y. Tak, S.J. Hong, J.S. Lee, K. Yong, Fabrication of ZnO/CdS core/shell nanowire arrays for efficient solar energy conversion. *J. Mater. Chem.* **19**(33), 5945–5951 (2009). doi:[10.1039/B904993B](https://doi.org/10.1039/B904993B)
122. L.E. Greene, M. Law, B.D. Yuhas, P. Yang, ZnO-TiO_2 core-shell nanorod/P3HT solar cells. *J. Phys. Chem. C* **111**(50), 18451–18456 (2007). doi:[10.1021/jp077593l](https://doi.org/10.1021/jp077593l)
123. B. Tian, X. Zheng, T.J. Kempa, Y. Fang, N. Yu, G. Yu, J. Huang, C.M. Lieber, Coaxial silicon nanowires as solar cells and nanoelectronic power sources. *Nature* **449** (7164):885–889 (2007)
124. G. Jia, M. Steglich, I. Sill, F. Falk, Core-shell heterojunction solar cells on silicon nanowire arrays. *Sol Energy Mat. Sol. C* **96**, 226–230 (2012). doi:[10.1016/j.solmat.2011.09.062](https://doi.org/10.1016/j.solmat.2011.09.062)
125. M.M. Adachi, M.P. Anantram, K.S. Karim, Core-shell silicon nanowire solar cells. *Sci. Rep.* **3** (2013). doi:<http://www.nature.com/srep/2013/130326/srep01546/abs/srep01546.html#supplementary-information>
126. X. Zhao, P. Wang, Y. Gao, X. Xu, Z. Yan, N. Ren, CuO/ZnO core/shell nanowire arrays and their photovoltaics application. *Mater. Lett.* **132**, 409–412 (2014). doi:[10.1016/j.matlet.2014.06.124](https://doi.org/10.1016/j.matlet.2014.06.124)
127. B.M. Kayes, H.A. Atwater, N.S. Lewis, Comparison of the device physics principles of planar and radial p-n junction nanorod solar cells. *J. Appl. Phys.* **97**(11), 114302 (2005). doi:[10.1063/1.1901835](https://doi.org/10.1063/1.1901835)
128. S.E. Han, G. Chen, Optical absorption enhancement in silicon nanohole arrays for solar photovoltaics. *Nano Lett.* **10**(3), 1012–1015 (2010). doi:[10.1021/nl904187m](https://doi.org/10.1021/nl904187m)

129. L. A  , D. Kieven, J. Chen, R. Klenk, T. Rissom, Y. Tang, M.C. Lux-Steiner, ZnO nanorod arrays as an antireflective coating for Cu(In, Ga)Se₂ thin film solar cells. *Prog. Photovoltaics Res. Appl.* **18**(3), 209–213 (2010). doi:[10.1002/pip.946](https://doi.org/10.1002/pip.946)
130. T. Stelzner, M. Pietsch, G. Andra, F. Falk, E. Ose, S. Christiansen, Silicon nanowire-based solar cells. *Nanotechnology* **19**(29), 295203 (2008)
131. S.A. Moiz, A.M. Nahhas, H.D. Um, S.W. Jee, H.K. Cho, S.W. Kim, J.H. Lee, A stamped PEDOT:PSS-silicon nanowire hybrid solar cell. *Nanotechnology* **23**(14), 145401 (2012)
132. A. Kim, Y. Won, K. Woo, C.-H. Kim, J. Moon, Highly transparent low resistance ZnO/Ag nanowire/ZnO composite electrode for thin film solar cells. *ACS Nano* **7**(2), 1081–1091 (2013). doi:[10.1021/nm305491x](https://doi.org/10.1021/nm305491x)
133. S. Cataldo, P. Salice, E. Menna, B. Pignataro, Carbon nanotubes and organic solar cells. *Energy Environ. Sci.* **5**(3), 5919–5940 (2012)
134. P.-L. Ong, W.B. Euler, I.A. Levitsky, Hybrid solar cells based on single-walled carbon nanotubes/Si heterojunctions. *Nanotechnology* **21**(10), 105203 (2010)
135. M.S. Dresselhaus, G.D. P. Avouris, *Carbon Nanotubes: Synthesis, Properties and Applications*, vol. 80 (Springer, Berlin, 2001). doi:[10.1007/3-540-39947-X](https://doi.org/10.1007/3-540-39947-X)
136. P.X. Hou, C. Liu, H.M. Cheng, Purification of carbon nanotubes. *Carbon* **46**(15), 2003–2025 (2008)
137. L.A. Montoro, J.M. Rosolen, A multi-step treatment to effective purification of single-walled carbon nanotubes. *Carbon* **44**(15), 3293–3301 (2006)
138. C.M. Aguirre, S. Auvray, S. Pigeon, R. Izquierdo, P. Desjardins, R. Martel, Carbon nanotube sheets as electrodes in organic light-emitting diodes. *Appl. Phys. Lett.* **88**(18) (2006)
139. D.H. Zhang, K. Ryu, X.L. Liu, E. Polikarpov, J. Ly, M.E. Thompson, C.W. Zhou, Transparent, conductive, and flexible carbon nanotube films and their application in organic light-emitting diodes. *Nano Lett.* **6**(9), 1880–1886 (2006)
140. B.B. Parekh, G. Fanchini, G. Eda, M. Chhowalla, Improved conductivity of transparent single-wall carbon nanotube thin films via stable postdeposition functionalization. *Appl. Phys. Lett.* **90**(12), 121913 (2007)
141. T. Tanaka, Y. Urabe, D. Nishide, H. Kataura, Discovery of surfactants for metal/semiconductor separation of single-wall carbon nanotubes via high-throughput screening. *J. Am. Chem. Soc.* **133**(44), 17610–17613 (2011)
142. P. Havu, M.J. Hashemi, M. Kaukonen, E.T. Sepp  l  , R.M. Nieminen, Effect of gating and pressure on the electronic transport properties of crossed nanotube junctions: formation of a Schottky barrier. *J. Phys.: Condens. Matter* **23**(11), 112203 (2011)
143. Y. Jung, X.K. Li, N.K. Rajan, A.D. Taylor, M.A. Reed, Record high efficiency single-walled carbon nanotube/silicon p-n junction solar cells. *Nano Lett.* **13**(1), 95–99 (2013)
144. F. Hennrich, R. Wellmann, S. Malik, S. Lebedkin, M.M. Kappes, Reversible modification of the absorption properties of single-walled carbon nanotube thin films via nitric acid exposure. *Phys. Chem. Chem. Phys.* **5**(1), 178–183 (2003)
145. R. Jackson, B. Domercq, R. Jain, B. Kippelen, S. Graham, Stability of doped transparent carbon nanotube electrodes. *Adv. Funct. Mater.* **18**(17), 2548–2554 (2008)
146. S.L. Hellstrom, M. Vosgueritchian, R.M. Stoltenberg, I. Irf  n, M. Hammock, Y.B. Wang, C. C. Jia, X.F. Guo, Y.L. Gao, Z.N. Bao, Strong and stable doping of carbon nanotubes and graphene by MoO_x for transparent electrodes. *Nano Lett.* **12**(7), 3574–3580 (2012)
147. H.-Z. Geng, K.K. Kim, K.P. So, Y.S. Lee, Y. Chang, Y.H. Lee, Effect of acid treatment on carbon nanotube-based flexible transparent conducting films. *J. Am. Chem. Soc.* **129**(25), 7758–7759 (2007). doi:[10.1021/ja0722224](https://doi.org/10.1021/ja0722224)
148. S. Kim, J. Yim, X. Wang, D.D.C. Bradley, S. Lee, J.C. Demello, Spin- and spray-deposited single-walled carbon-nanotube electrodes for organic solar cells. *Adv. Funct. Mater.* **20**(14), 2310–2316 (2010)

149. M.S. Strano, V.C. Moore, M.K. Miller, M.J. Allen, E.H. Haroz, C. Kittrell, R.H. Hauge, R.E. Smalley, The role of surfactant adsorption during ultrasonication in the dispersion of single-walled carbon nanotubes. *J. Nanosci. Nanotechnol.* **3**(1–2), 81–86 (2003)
150. H. Cebeci, R.G. de Villoria, A.J. Hart, B.L. Wardle, Multifunctional properties of high volume fraction aligned carbon nanotube polymer composites with controlled morphology. *Compos. Sci. Technol.* **69**(15–16), 2649–2656 (2009)
151. A. Hagfeldt, M. Gratzel, Light-induced redox reactions in nanocrystalline systems. *Chem. Rev.* **95**(1), 49–68 (1995)
152. CCISolar, Dye sensitized solar cell—DSCC (2015). <http://www.ccisolar.caltech.edu/index.php?module=webpage&id=113&page=3>
153. A.C. Fisher, L.M. Peter, E.A. Ponomarev, A.B. Walker, K.G.U. Wijayantha, Intensity dependence of the back reaction and transport of electrons in dye-sensitized nanocrystalline TiO₂ solar cells. *J. Phys. Chem. B* **104**(5), 949–958 (2000)
154. T. Oekermann, D. Zhang, T. Yoshida, H. Minoura, Electron transport and back reaction in nanocrystalline TiO₂ films prepared by hydrothermal crystallization. *J. Phys. Chem. B* **108**(7), 2227–2235 (2004)
155. R.C. Nelson, Energy transfers between sensitizer and substrate. III. sensitization by thick dye films. *J. Opt. Soc. Am.* **51**(11), 1182–1186 (1961). doi:[10.1364/JOSA.51.001182](https://doi.org/10.1364/JOSA.51.001182)
156. J. van de Lagemaat, A.J. Frank, Nonthermalized electron transport in dye-sensitized nanocrystalline TiO₂ films: transient photocurrent and random-walk modeling studies. *J. Phys. Chem. B* **105**(45), 11194–11205 (2001). doi:[10.1021/jp0118468](https://doi.org/10.1021/jp0118468)
157. T.Y. Lee, P.S. Alegaonkar, J.B. Yoo, Fabrication of dye sensitized solar cell using TiO₂ coated carbon nanotubes. *Thin Solid Films* **515**(12), 5131–5135 (2007)
158. K.-M. Lee, C.-W. Hu, H.-W. Chen, K.-C. Ho, Incorporating carbon nanotube in a low-temperature fabrication process for dye-sensitized TiO₂ solar cells. *Sol. Energy Mat. Sol. C* **92**(12), 1628–1633 (2008). doi:[10.1016/j.solmat.2008.07.012](https://doi.org/10.1016/j.solmat.2008.07.012)
159. K.T. Dembele, G.S. Selopal, C. Soldano, R. Nechache, J.C. Rimada, I. Concina, G. Sberveglieri, F. Rosei, A. Vomiero, Hybrid carbon nanotubes-TiO₂ photoanodes for high efficiency dye-sensitized solar cells. *J. Phys. Chem. C* **117**(28), 14510–14517 (2013)
160. Z. Peining, A.S. Nair, Y. Shengyuan, P. Shengjie, N.K. Elumalai, S. Ramakrishna, Rice grain-shaped TiO₂-CNT composite—a functional material with a novel morphology for dye-sensitized solar cells. *J. Photochem. Photobiol. A* **231**(1), 9–18 (2012). doi:[10.1016/j.jphotochem.2012.01.002](https://doi.org/10.1016/j.jphotochem.2012.01.002)
161. L.J. Yang, W.W.F. Leung, Electrospun TiO₂ nanorods with carbon nanotubes for efficient electron collection in dye-sensitized solar cells. *Adv. Mater.* **25**(12), 1792–1795 (2013)
162. H.S. Wroblowa, A. Saunders, Flow-through electrodes: II. The I₃⁻/I⁻-redox couple. *J. Electroanal. Chem. Interfacial Electrochem.* **42**(3), 329–346 (1973). doi:[10.1016/S0022-0728\(73\)80323-7](https://doi.org/10.1016/S0022-0728(73)80323-7)
163. K. Suzuki, M. Yamaguchi, M. Kumagai, S. Yanagida, Application of carbon nanotubes to counter electrodes of dye-sensitized solar cells. *Chem. Lett.* **32**(1), 28–29 (2003). doi:[10.1246/cl.2003.28](https://doi.org/10.1246/cl.2003.28)
164. S. Uk Lee, W. Seok Choi, B. Hong, A comparative study of dye-sensitized solar cells added carbon nanotubes to electrolyte and counter electrodes. *Sol. Energy Mat. Sol. C* **94**(4), 680–685 (2010). doi:[10.1016/j.solmat.2009.11.030](https://doi.org/10.1016/j.solmat.2009.11.030)
165. S.H. Seo, S.Y. Kim, B.K. Koo, S.I. Cha, D.Y. Lee, Influence of electrolyte composition on the photovoltaic performance and stability of dye-sensitized solar cells with multiwalled carbon nanotube catalysts. *Langmuir* **26**(12), 10341–10346 (2010)
166. J. Velten, A.J. Mozer, D. Li, D. Officer, G. Wallace, R. Baughman, A. Zakhidov, Carbon nanotube/graphene nanocomposite as efficient counter electrodes in dye-sensitized solar cells. *Nanotechnology* **23**(8), 085201 (2012)
167. H. Anwar, A.E. George, I.G. Hill, Vertically-aligned carbon nanotube counter electrodes for dye-sensitized solar cells. *Sol. Energy* **88**, 129–136 (2013)

168. G.-R. Li, F. Wang, Q.-W. Jiang, X.-P. Gao, P.-W. Shen, Carbon nanotubes with titanium nitride as a low-cost counter-electrode material for dye-sensitized solar cells. *Angew. Chem. Int. Ed.* **49**(21), 3653–3656 (2010). doi:[10.1002/anie.201000659](https://doi.org/10.1002/anie.201000659)
169. Q.W. Jiang, G.R. Li, X.P. Gao, Highly ordered TiN nanotube arrays as counter electrodes for dye-sensitized solar cells. *Chem. Commun.* **44**, 6720–6722 (2009)
170. H.-J. Shin, S.S. Jeon, S.S. Im, CNT/PEDOT core/shell nanostructures as a counter electrode for dye-sensitized solar cells. *Synthetic Met.* **161**(13–14), 1284–1288 (2011). doi:[10.1016/j.synthmet.2011.04.024](https://doi.org/10.1016/j.synthmet.2011.04.024)

Chapter 16

Computational Studies of Bismuth-Doped Zinc Oxide Nanowires

Çetin Kılıç, Mehmet Aras and Sümeyra Güler-Kılıç

Abstract A computational modeling methodology is described, which is based on the total-energy and defect calculations using the density and hybrid functionals in combination with the supercell approach. This methodology is employed to study the structure and energetics of zinc oxide (ZnO) nanowires, and the electrical properties of a bismuth-doped zinc oxide (ZnO:Bi) nanowire. A simple model is developed, where the nanowire formation energy is expressed as a function of the nanowire diameter and cross-sectional morphology (hexagonal or triangular). Defect calculations are employed to characterize the ZnO:Bi nanowire, in regard to the location and charge-state of the dopant. The latter indicates that the dopant atoms are mostly substituted into the Zn sites on the nanowire surface, and therefore Bi in ZnO electrically acts as a donor.

16.1 Introduction

The doping of one-dimensional ZnO nanostructures has been attracting a growing interest in recent decades [1], which is driven by not only scientific curiosity but also potential device applications [2, 3]. In accord with this research trend, an increasing number of computational studies, e.g., [4–10], has been devoted to the modeling of doped ZnO nanowires in order to elucidate their behavior in regard to a variety of phenomena such as *p*-type electrical conduction [11–13], which were mostly performed within the framework of the density functional theory (DFT). In this chapter, we utilize a modeling methodology for studying the doping of ZnO

Ç. Kılıç (✉) · M. Aras · S. Güler-Kılıç

Department of Physics, Gebze Technical University, 41400 Gebze, Kocaeli, Turkey
e-mail: cetin_kilic@gtu.edu.tr

M. Aras

e-mail: m.aras@gtu.edu.tr

S. Güler-Kılıç

e-mail: sumeyra@gtu.edu.tr

nanowires via computational methods, which is developed in Sect. 16.2. This methodology is exemplified by application to (i) undoped ZnO nanowires in Sect. 16.3, and (ii) a bismuth-doped zinc oxide (ZnO:Bi) nanowire in Sect. 16.4 where we focus on the electrical properties of this nanowire.

Although bismuth is known to act as an *acceptor* in ZnO thin films [14, 15], it has a low solubility in zinc oxide [16] so that Bi-doped ZnO films usually exhibit *n*-type electrical conduction. This is so because undoped ZnO is naturally *n*-type due to the existence of intrinsic donors such as zinc interstitials and/or unintentional dopants such as hydrogen, and therefore the holes produced by the acceptors are *compensated* by the native defects of ZnO, hindering *p*-type doping [17]. On the other hand, a recent experimental study [15] suggests that stable and reproducible *p*-type conductivity could be achieved by doping ZnO thin films by Bi *provided that* the bismuth atoms are incorporated substitutionally into the oxygen sites, forming acceptor-like Bi_O defects. Needless to say, the latter is usually difficult due to the *segregation* behavior of Bi in ZnO, which is a well-known phenomenon owing to its occurrence in zinc oxide varistors [16, 18]. In the latter systems, Bi segregates to the grain boundaries where it leads to the formation of acceptor-like defect complexes [19]. As long as the *p*-doped grain boundaries could act as electron traps, the electron-rich (i.e., *n*-type) interior regions are surrounded by an electrostatic barrier created at the Bi-decorated grain boundaries. Hence, the varistor has a high resistance at low voltages. The nonlinear current-voltage characteristic of the varistor is produced by the release of the electrons at the breakdown voltage. Accordingly, the varistor has a low resistance at high voltages. The electrical properties of undoped and Bi-doped ZnO thin films are compared in Table 16.1. It is seen that doping by Bi results in (i) a significant decrease in the electron concentration of the sample of [14], and (ii) a substantial increase in the hole concentration of the sample of [15]. Both of the latter are indicators for the acceptor behavior originated from doping of ZnO by Bi.

The electrical behavior of ZnO:Bi nanowires should be contrasted with that of thin films since bismuth acts as a *donor* in ZnO nanowires [20]. A comparison of the electrical properties of undoped and Bi-doped ZnO nanowires is given in Table 16.2. It is seen that both undoped and Bi-doped ZnO nanowires have *n*-type conduction, and the electron concentration is increased by one order of magnitude as a result of doping by Bi. Consequently, it is likely that Bi substitutes Zn site to

Table 16.1 Electrical properties of undoped and Bi-doped ZnO thin films

	Conduction type	Carrier concentration (cm ⁻³)	Mobility (cm ² /Vs)	Resistivity (Ω cm)	Thickness (nm)	References
ZnO	<i>n</i>	-2.7×10^{18}	7.4	0.3	660	[14]
ZnO:Bi	<i>n</i>	-9.1×10^{16}	15.9	4.3		[14]
ZnO	<i>n</i>	-6.6×10^{17}	122.1	0.077	200	[15]
ZnO:Bi	<i>p</i>	4.8×10^{19}	5.9	0.022		[15]

Table 16.2 Electrical properties of undoped and Bi-doped ZnO nanowires

	Conduction type	Carrier concentration (cm^{-3})	Mobility (cm^2/Vs)	Resistivity ($\Omega \text{ cm}$)	Radius (nm)	References
ZnO	<i>n</i>	-4.0×10^7	17.2	0.2	60	[21]
ZnO:Bi	<i>n</i>	-3.5×10^8	1.5	0.1	20	[20]

form a donor-like defect Bi_{Zn} in a positive charge state, releasing electrons [20]. Interestingly, the charge state of Bi in ZnO thin films is also *positive* [14], which imply that Bi_{Zn} forms in the sample of [14] too. On the contrary, the acceptor behavior of bismuth in the sample of [15] is attributed to Bi_{O} , as mentioned above. Hence, comparative analysis of experimental findings results in some degree of controversy in regard to the questions: (i) What is the charge state of Bi in ZnO? (ii) Does Bi substitute Zn or O sites? (iii) Does Bi act as a donor or acceptor in ZnO? It is thus fortunate that the state and location of Bi in ZnO could also be characterized by using the aforementioned modeling methodology, notwithstanding the experimental controversies. It is the latter task that is accomplished in this chapter, with focus on the nanowires.

The rest of this chapter is organized as follows. Section 16.2 is devoted to the elements of modeling methodology employed in the following sections, where the computational methods of calculation are described in detail. We present the results of calculations performed for the structure and energetics of the undoped ZnO nanowires in Sect. 16.3. The electrical properties of ZnO:Bi nanowire are studied in Sect. 16.4 with the aid of defect calculations.

16.2 Computational Modeling

In the modeling of nanowires, we employ the total-energy and electronic-structure calculations based on the (hybrid) density functional theory+*U* approach in combination with a variety of periodic supercells, where *U* is the Hubbard on-site repulsion. The supercell approach is also adopted in our *defect calculations* that are utilized to investigate doping of a zinc oxide nanowire with bismuth. Hence, in this section, we first describe the supercell approach, which is followed by a discussion on the defect calculations. Next, we expound on the *density- and hybrid-functional+U* calculations, and lastly, a summary of the computational details are given.

16.2.1 Supercell Approach

The nanowires under consideration here are cut from bulk ZnO in wurtzite structure in such a way that the wire axis coincide with the [0001] direction of wurtzite,

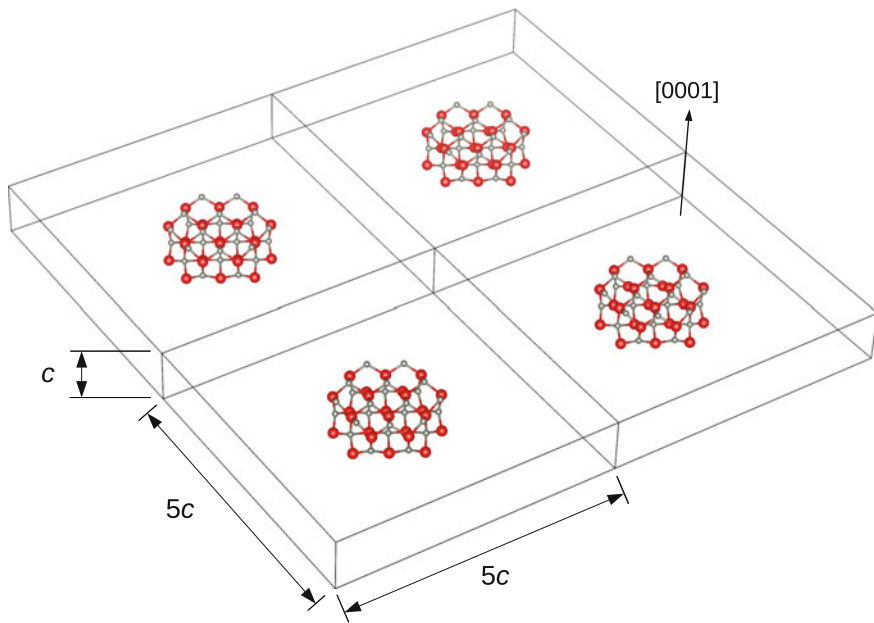


Fig. 16.1 The tetragonal supercell of dimension $5c \times 5c \times c$ used to study *undoped* ZnO nanowires. The [0001] direction of wurtzite is indicated by the *vertical arrow*

which have either a hexagonal or triangular cross section. The supercells containing these nanowires are devised to include vacuum regions in order to make the nanowire extend to infinity *only* along the [0001] direction of wurtzite although the supercells themselves are subject to the Bloch periodicity condition in all directions. Figure 16.1 displays a tetragonal supercell of dimension $5c \times 5c \times c$ that is employed for *undoped* ZnO nanowires, where c denotes the length of periodicity along the wire axis. A cubic supercell of side length $L = 5c$ is used for *doped* ZnO nanowires, which is displayed in Fig. 16.2. Generating a supercell configuration for *substitutional* doping of ZnO nanowire is also schematically illustrated in the latter figure. The relevant supercell configurations (i.e., atomistic structures contained in the aforementioned supercells) are all optimized via minimization of the total energy that is computed as described below in Sect. 16.2.3 (with the computational settings given in Sect. 16.2.4). Prior to the latter, the equilibrium value of c is determined for each nanowire, which corresponds to the energy minimum.

16.2.2 Defect Calculations

It has become customary to perform defect calculations using the supercell approach in combination with the total-energy calculations (cf. Sect. 16.2.3) within

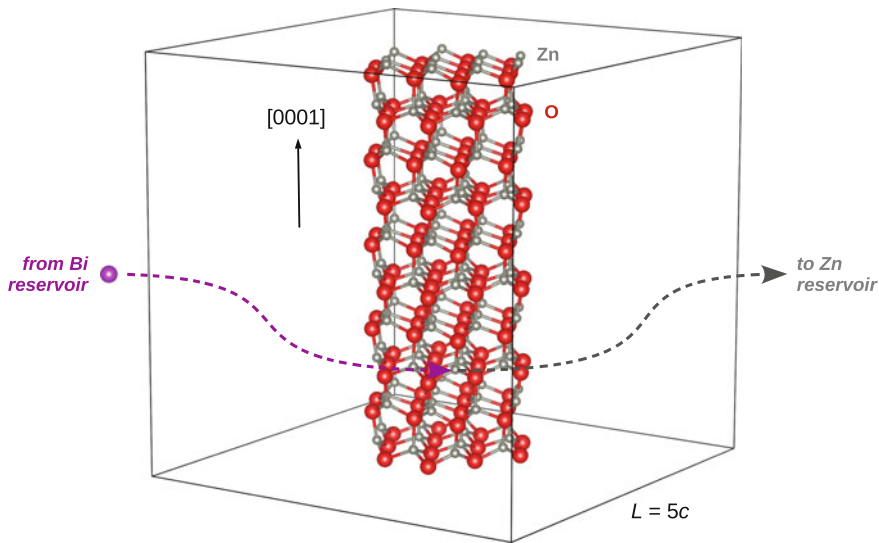


Fig. 16.2 The cubic supercell of length $L = 5c$ employed for doped ZnO nanowires. The vertical arrow indicates the wire axis that coincides with the [0001] direction of wurtzite. Generating a supercell configuration for studying a Bi-doped ZnO nanowire is schematically illustrated, where an extrinsic defect Bi_{Zn} is created

the framework of a simple thermodynamic model where a host material is considered in *equilibrium* with the reservoirs of atoms and the Fermi sea of electrons. The *intrinsic* and/or *extrinsic* defects are created by exchanging atoms and/or electrons between the host and the reservoirs. This is illustrated schematically in Fig. 16.2 for the *substitutional* doping of ZnO nanowire with Bi, where an extrinsic defect is created by transferring a Zn (Bi) atom to (from) the zinc (bismuth) reservoir from (to) the host, which will be denoted as Bi_{Zn} . Note that a *charged* defect Bi_{Zn}^q would be created if q electrons were also transferred to the Fermi reservoir. It is also notable that transferring q electrons from the host to the Fermi reservoir (from the Fermi reservoir to the host) implies using a *positively-charged* (*negatively-charged*) supercell, which enables modeling of *donor-like* (*acceptor-like*) defects. For Bi_{Zn}^q , the energetics of defect formation could be described as follows: Assume that the ZnO nanowire is in equilibrium with the zinc, oxygen, bismuth, and Fermi reservoirs consisting of m_{Zn} , m_{O} , m_{Bi} , and m_e particles with chemical potentials μ_{Zn} , μ_{O} , μ_{Bi} , and μ_e , respectively. The total energy of a combined system made of a defect-free nanowire and all reservoirs is $E_0 + m_{\text{Zn}}\mu_{\text{Zn}} + m_{\text{O}}\mu_{\text{O}} + m_{\text{Bi}}\mu_{\text{Bi}} + m_e\mu_e$, where E_0 denote the total energy of the supercell containing the *defect-free* nanowire. Once Zn and Bi atoms and q electrons are exchanged between the nanowire and the reservoirs to form Bi_{Zn}^q in the nanowire, the total energy becomes $E(\text{Bi}_{\text{Zn}}^q) + (m_{\text{Zn}} + 1)\mu_{\text{Zn}} + m_{\text{O}}\mu_{\text{O}} + (m_{\text{Bi}} - 1)\mu_{\text{Bi}} + (m_e + q)\mu_e$ where $E(\text{Bi}_{\text{Zn}}^q)$ denote the total energy of the supercell containing the nanowire with defect Bi_{Zn}^q . Thus, the *defect formation energy* ΔH_f ,

which is defined as the change in energy associated with the formation of the defect, is given by

$$\Delta H_f(\text{Bi}_{\text{Zn}}^q) = [E(\text{Bi}_{\text{Zn}}^q) - E_0] + \mu_{\text{Zn}} - \mu_{\text{Bi}} + q\mu_e, \quad (16.1)$$

for Bi_{Zn}^q , which is clearly a function of the chemical potentials μ_{Zn} , μ_{Bi} , and μ_e . It is customary to set the *zero* of the atomic chemical potential μ_x ($x = \text{Zn}$, O , or Bi) to the energy per atom $E(x)$ of an elemental form of the respective atomic species. On the other hand, the electrochemical potential of electrons μ_e , viz. the *Fermi level* E_F , is conventionally referenced to the valence band maximum (VBM) energy E_v of the host. Consequently, (16.1) is modified as

$$\Delta H_f(\text{Bi}_{\text{Zn}}^q) = [E(\text{Bi}_{\text{Zn}}^q) - E_0 + E(\text{Zn}) - E(\text{Bi}) + qE_v] + \mu_{\text{Zn}} - \mu_{\text{Bi}} + qE_F. \quad (16.2)$$

A similar derivation for the defect Bi_{O}^q that is formed by substituting the dopant on an O site (to replace a constituent O atom) results in

$$\Delta H_f(\text{Bi}_{\text{O}}^q) = [E(\text{Bi}_{\text{O}}^q) - E_0 + E(\text{O}) - E(\text{Bi}) + qE_v] + \mu_{\text{O}} - \mu_{\text{Bi}} + qE_F, \quad (16.3)$$

which is a function of the chemical potentials μ_{O} , and μ_{Bi} , and E_F . If the dopant Bi is incorporated into an *interstitial* site to form Bi_i^q , then the defect formation energy is given by

$$\Delta H_f(\text{Bi}_i^q) = [E(\text{Bi}_i^q) - E_0 - E(\text{Bi}) + qE_v] - \mu_{\text{Bi}} + qE_F. \quad (16.4)$$

It should be emphasized that the defect formation energy ΔH_f is an indicator for the *abundance* of the defect under given thermodynamic conditions since it is a significant portion of the *Gibbs energy of formation* that determines the *equilibrium defect concentration*. ΔH_f is therefore usually computed for a relevant range of the atomic chemical potentials μ_x as well as the Fermi level E_F . It is further to be reminded that this simple thermodynamic model has long been used to study the defect structure of *solids* [21–23] forasmuch as the derivations given above are generally valid for any material system including solids, which was also employed in former studies [24–27] of one of the authors.

It is clear in the foregoing formulation that the defect formation energy ΔH_f is linearly proportional (with a negative slope) to the Fermi level E_F for $q > 0$ ($q < 0$) while ΔH_f is independent of E_F for $q = 0$, cf. (16.2)–(16.4). Thus ΔH_f is usually studied as a function of E_F in order to determine the *charge state* of the defect (or dopant) under consideration. Note that the value of E_F eventually depends on the *free-carrier concentration*. The *defect transition energy* $\varepsilon(q_1/q_2)$ of defect D is defined as the value of E_F for which $\Delta H_f(D^{q_1}) = \Delta H_f(D^{q_2})$, which is obtained via

$$\varepsilon(q_1/q_2) = \frac{E(D^{q_2}) - E(D^{q_1})}{q_1 - q_2}. \quad (16.5)$$

Note that $\varepsilon(q_1/q_2)$ is *independent* of the atomic chemical potentials μ_α . If E_F is smaller (larger) than $\varepsilon(q_1/q_2)$, then the charge state of D is $q_1(q_2)$. Thus, $\varepsilon(q_1/q_2)$ gives the value of the Fermi level at which the transition from charge state q_1 to q_2 takes place. The defect transition energies correspond to *thermal ionization energies*, and accordingly they are used to determine the position of electrical (donor or acceptor) levels with respect to the band edges in order to reveal the behavior of defects (or dopants). For example, considering a defect with possible charge states $q = -, 0$, and $+$, if the difference $E_v - \varepsilon(0/-) [\varepsilon(+/0) - E_c]$ is positive (or negative with a magnitude within a few $k_B T$), then the defect yields a *shallow* acceptor [donor] level. Here E_c denotes the conduction band minimum (CBM) energy. Otherwise, the electrical level is *deep* since it would require an amount of energy appreciably greater than the thermal energy $k_B T$ to ionize this defect.

As for the atomic chemical potentials $\{\mu_\alpha\}$ appearing in (16.2)–(16.4), it should be reminded that the *actual* value of μ_α depends on the *kinetic* or *thermodynamic* conditions during *growth* or in the course of *equilibration* processes. It is thus conventional to employ a lower and an upper bound for each μ_α , which would enable one to study typical representative cases. For the constituent atoms Zn and O, these lower and upper bounds are determined as follows: It is required to have $\mu_{Zn}, \mu_O \leq 0$ because the host system (i.e., ZnO nanowire) would otherwise decompose into its elemental constituents. Moreover,

$$\mu_{Zn} + \mu_O = \Delta H_N, \quad (16.6)$$

assuming that chemical equilibrium is maintained between the host system and the atomic reservoirs, where ΔH_N denotes the *heat of formation* (per formula unit) of the nanowire made of N Zn–O pairs. The latter implies $\Delta H_N \leq \mu_\alpha$ ($\alpha = \text{Zn or O}$) so that the lower bound is equal to ΔH_N . Hence

$$\Delta H_N \leq \mu_\alpha \leq 0 \quad (\alpha = \text{Zn or O}). \quad (16.7)$$

μ_α is equal to upper (lower) bound if the host system is in an environment that is extremely *rich* (*poor*) in element α . For example, O–poor condition implies $\mu_O = \Delta H_N$, whereas O–rich condition means $\mu_O = 0$. Accordingly, $\mu_{Zn} = 0$ and $\mu_{Zn} = \Delta H_N$ under O–poor (i.e., Zn–rich) and O–rich (i.e., Zn–poor) conditions, respectively.

The value of the Bi chemical potential is set to the adsorption energy of a Bi atom on the outer surface of the ZnO nanowire. The lowest-energy adatom configuration is shown in Fig. 16.3, for which the adsorption energy $E_{ad} = E(\text{Bi/nanowire}) - E(\text{nanowire}) - E(\text{Bi})$ is -0.577 eV. Note that $E_{ad} < 0$ for the adatom configuration displayed in Fig. 16.3, indicating the attractive interaction between the adatom and the nanowire. As for the results presented in Sect. 16.4, we set $\mu_{Bi} = E_{ad}$. The latter means that the zero of μ_{Bi} is taken as the energy of the Bi atom $E(\text{Bi})$. On the other hand, the ground-state phase is used as a *reference* for the constituent elements of the nanowire, viz. Zn and O, i.e., the zero of μ_{Zn} and μ_O are set to the energy per atom $E(\text{Zn})$ and $E(\text{O})$ of bulk Zn metal and O_2 molecule, respectively.

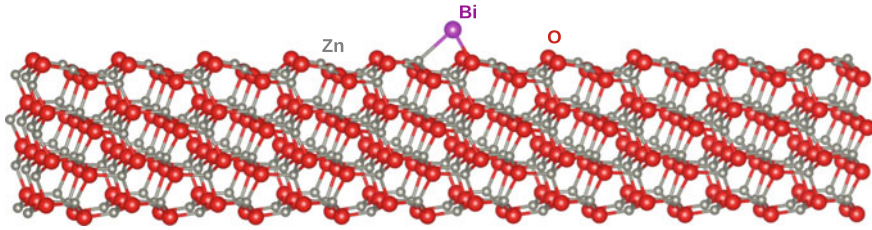


Fig. 16.3 The lowest-energy configuration employed to calculate the adsorption energy Bi on the ZnO nanowire surface

16.2.3 Density- and Hybrid-Functional+U Calculations

First-principles calculations based on the density functional theory [28, 29] (DFT) have been the de facto method of choice in ab initio materials modeling for computational efficiency [30]. It is nevertheless notable that the success (or failure) of DFT in *realistic* modeling depends inherently on the quality of approximation for the exchange-correlation energy E_{xc} functional. The standard DFT calculations performed within the local density approximation [29] (LDA) or the generalized gradient approximation using the Perdew-Burke-Ernzerhof (PBE) functional [31] are not entirely satisfactory for modeling materials containing localized electrons such as some transition-metal or rare-earth compounds. Thus, the local/semilocal functionals are often *either* combined with Hubbard U , as in the LDA+ U method [32], *or* mixed with a fraction of Hartree-Fock exchange energy [33], as in the *hybrid functional* scheme [34]. In a recent study [35] devoted to a set of II–VI semiconductors with localized d states, we found that employing the hybrid functional scheme and the DFT+ U method as *complementary* approaches results in an improved description *not only* of the electronic structure *but also* of the crystal structure and energetics. In contrast, the improvements due to employing a hybrid functional or Hubbard U *alone* are *not* sufficient to make the predictions agree with the experimental data, especially for wide band gap semiconductors such as ZnO that has localized d electrons. The latter is particularly valid in predicting the band gap E_g as long as the band gap is *underestimated* in proportionality with $1/\epsilon_\infty$, cf. the Appendix of [35], where ϵ_∞ denotes the high-frequency dielectric constant. It is demonstrated in [35] that this well-known *band-gap problem* could be cured by adopting a *hybrid-functional+U* approach where the range-separated Heyd-Scuseria-Ernzerhof (HSE) hybrid functional [36] is combined with the Hubbard U . This can be briefly explained as follows: the band-gap correction due to the Hubbard term turns out to be proportional to U/ϵ_∞ , cf. Fig. 4 of [35], which complements the HSE-calculated band gap (being underestimated in proportionality with $1/\epsilon_\infty$) provided that U is set to have a suitable value. Consequently, the suitable U value is determined by simply matching the experimental band gap, which is denoted as U^* . This is illustrated in Fig. 16.4 where the calculated band gap is plotted as a function of U . The variation of the band gap with U is seemingly *linear*.

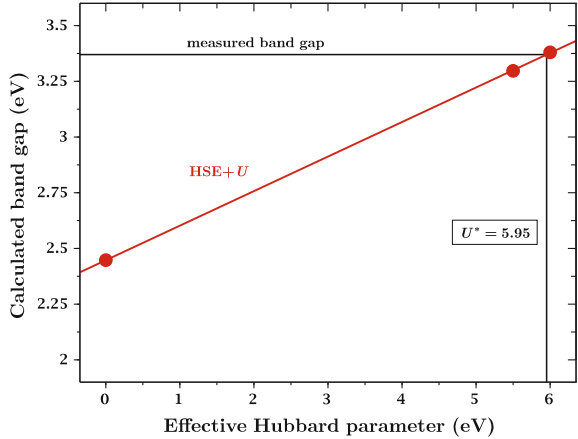


Fig. 16.4 Setting the value of U^* by matching the experimental band gap. For wurtzite ZnO bulk crystal, the band gaps (solid circles) obtained in our HSE+ U calculations is plotted with respect to the effective Hubbard parameter U . The red line represents a linear fit to the calculated points. The vertical line marks U^*

Table 16.3 The calculated (PBE, PBE+ U , HSE, HSE+ U^*) and experimental values of the wurtzite lattice parameters a and c , the wurtzite internal parameter u , the formation energy ΔH_f , the d band position ε_d relative to the valence band maximum, and the band gap E_g for wurtzite ZnO bulk crystal

Method	$a(\text{\AA})$	$c(\text{\AA})$	c/a	u	ΔH_f (eV)	ε_d (eV)	E_g (eV)
PBE	3.2872	5.3059	1.6141	0.3790	-2.8858	-4.66	0.72
PBE+ U	3.1801	5.1208	1.6102	0.3798	-3.6770	-6.25	1.91
HSE	3.2606	5.2332	1.6023	0.3811	-3.1196	-5.71	2.45
HSE+ U^*	3.1850	5.1070	1.6036	0.3812	-3.7409	-7.03	3.37
Experiment	3.2495	5.2069	1.6024	0.3819	-3.6326	-7.50	3.37

Hence, the U value that corresponds to the experimental band gap is obtained via a linear fit, yielding $U^* = 5.95$ eV for ZnO.

It should be emphasized that the results obtained for a set of II–VI semiconductors with d electrons [35] prove that an improved description of the electronic and crystal structure and energetics is obtained via HSE+ U^* calculations, compared to the respective HSE or PBE+ U calculations. Comparative inspection of rows of Table 16.3 reveals that the latter is particularly true for wurtzite ZnO bulk crystal, where the calculated and experimental values are given for the following quantities: the wurtzite lattice parameters a and c , the wurtzite internal parameter u , the formation energy ΔH_f , the d band position ε_d relative to the valence band maximum, and the band gap E_g .

Within the HSE+ U approach, the exchange-correlation energy functional is effectively generalized to

$$\begin{aligned}
E_{\text{xc}}^{\text{HSE}+U} = & \alpha E_{\text{x}}^{\text{HF,SR}}(\omega) + (1 - \alpha) E_{\text{x}}^{\text{PBE,SR}}(\omega) + E_{\text{x}}^{\text{PBE,LR}}(\omega) + E_{\text{c}}^{\text{PBE}} \\
& + \frac{U}{2} \sum_{\sigma} \left(\sum_i \rho_{ii}^{\sigma} - \sum_{i,j} \rho_{ij}^{\sigma} \rho_{ji}^{\sigma} \right), \tag{16.8}
\end{aligned}$$

where the first line gives the range-separated HSE functional, and the simplified (rotationally invariant) formulation [37] of the DFT+ U method is adopted in the second line. Note that the *effective* Hubbard parameter $U = \overline{U} - \overline{J}$ is introduced as the difference between the on-site Coulomb $\overline{U}$ and exchange $\overline{J}$ parameters in the latter formulation [37] where ρ_{ij}^{σ} denotes the density matrix of *localized* electrons with spin projection σ . In (16.8), $E_{\text{x}}^{\text{HF,SR}}$ and $E_{\text{x}}^{\text{PBE,SR}}$ are the *short-range* contribution from the HF and PBE exchange functionals, respectively, while $E_{\text{x}}^{\text{PBE,LR}}$ is the *long-range* contribution from the PBE exchange functional, and $E_{\text{c}}^{\text{PBE}}$ denotes the correlation energy functional within the generalized gradient approximation. Note that range separation is accomplished by splitting the Coulomb operator into two components with the aid of the *screening parameter* ω , i.e., $1/r = [1 - \text{erf}(\omega r)]/r + \text{erf}(\omega r)/r$, and the screened Coulomb potential is applied only to the exchange interaction [36]. The *exchange mixing coefficient* α is then used only for the short-range components of the HF and PBE exchange functionals. Clearly, $E_{\text{xc}}^{\text{HSE}+U}$ reduces to the HSE functional for $U = 0$. On the other hand, the PBE functional is recovered in (16.8) for $U = 0$, $\alpha = 0$, and $\omega \rightarrow \infty$. Hence, (16.8) yields the PBE+ U functional for $U \neq 0$, $\alpha = 0$, and $\omega \rightarrow \infty$.

It should be pointed out from a practical point of view that employing hybrid-functionals for the defect calculations (cf. Sect. 16.2.2) becomes an exhaustive task *once* one needs to use large supercells, e.g., Fig. 16.2, especially if the computational resources are moderate. For this reason and encouraged by the results of a number of test calculations, we devised a computational strategy where the PBE+ U and HSE+ U^* calculations are employed successively: The former is used to perform structure optimizations, which is followed by computing the electronic structure via the latter. We found that an adequate U value for PBE+ U could be determined so that structure optimizations via the PBE+ U and HSE+ U^* calculations yield very similar equilibrium geometries. In practice, the U value for PBE+ U could be set by matching some structural parameter predicted via the PBE+ U and HSE+ U^* calculations. This is demonstrated in Fig. 16.5 for the unit cell volume. It is seen in Fig. 16.5 that the PBE+ U and HSE+ U^* values for the equilibrium unit cell volume turn out to be identical for $U = 7.7$ eV, given that $U^* = 5.95$ eV. Test calculations performed for a thin ZnO nanowire with the foregoing values of U and U^* indicate that employing the PBE+ U and HSE+ U^* calculations successively introduces an uncertainty on the order of 10^{-2} Å in the equilibrium positions of the nanowire atoms.

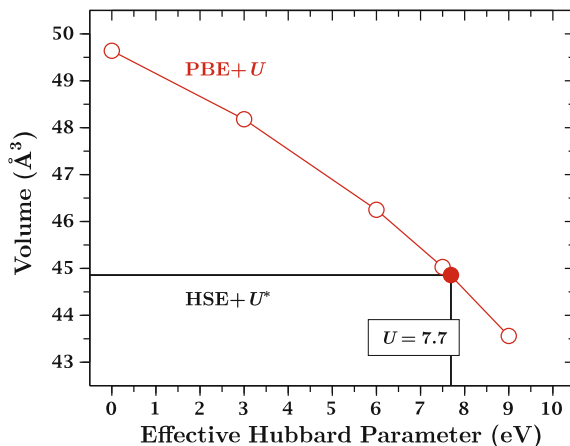


Fig. 16.5 Setting the value of U for PBE+ U by matching the HSE+ U^* -optimized unit cell volume. For wurtzite ZnO bulk crystal, the PBE+ U -optimized volume (*empty circles*) is plotted with respect to the effective Hubbard parameter U . The *red curve* represents a polynomial fit to the calculated points. The vertical line marks the value of U for PBE+ U

16.2.4 Computational Settings

The density-functional (PBE+ U) as well as hybrid-functional (HSE+ U^*) calculations were carried out using the projector augmented-wave (PAW) method [38], as implemented [39] in the Vienna ab initio simulation package [40] (VASP). Plane wave basis sets were used to represent the electronic states, which were determined by imposing a kinetic energy cutoff of 400 eV. The $2s$ and $2p$, $3d$ and $4s$, and $6s$ and $6p$ states were treated as valence states for oxygen, zinc, and bismuth, respectively. The Hubbard U was applied only to the Zn $3d$ states. The Brillouin zone of the tetragonal supercells used for undoped nanowires, cf. Fig. 16.1, was sampled by $1 \times 1 \times 7$ k-point mesh, which were generated according to the Monkhorst-Pack scheme [41]. The Γ point was used for sampling the Brillouin zone of the cubic supercells employed for doped nanowires, cf. Fig. 16.2. Convergence criterion for the electronic self-consistency was set up to 10^{-6} eV. The charged supercell calculations were performed with the aid of a neutralizing jellium background. By studying the supercell energy as a function of supercell size L , we estimated that the finite-size error is on the order of 0.2 eV. In the optimization of the atomistic structures, ionic relaxations were performed for each supercell configuration to minimize the total energy, until the maximum value of residual forces on atoms was reduced to be smaller than 10^{-2} eV/Å.

16.3 Structure and Energetics of ZnO Nanowires

The optimized structures of the ZnO nanowires studied here are displayed in Fig. 16.6, which have either hexagonal (upper panels) or triangular (lower panels) cross sections. The chemical formula $(\text{ZnO})_N$ is used throughout to indicate these structures. Note that $N = 6, 24$, or 54 ($13, 22$, or 33) for the hexagonal (triangular) nanowires. The *inradius* r and the *circumradius* R are introduced to differentiate the interior (i.e., *bulk-like*) and *surface* regions, which are indicated by the green and blue circles, respectively, in Fig. 16.6. The diameter D of the nanowire is given by $D = 2\sqrt{rR}$, which is set to be equal to twice the *midradius* $\sqrt{rR}$. D is plotted with respect to the inverse number of Zn–O pairs ($1/N$) in the upper panel of Fig. 16.7. The middle and lower panels in Fig. 16.7 display the percent changes ΔD and Δc in D and c , respectively, as a function of $1/N$, relative to the *unrelaxed* nanowire structure. It is seen that the former takes only negative values, i.e., $\Delta D < 0$, while the latter takes only positive values, i.e., $0 < \Delta c$, for all nanowires, upon structural relaxation. This means that the equilibrium structures of the nanowires are elongated along the wire axis and at the same time radially shrunk, compared to the unrelaxed structures. Moreover, the variation of the curves in the middle and lower panels of Fig. 16.7 follows a similar trend, indicating that the radial shrinkage is roughly proportional to the axial elongation. The equilibrium structures displayed in Fig. 16.6 are further characterized by measuring the O–Zn interatomic distances that are given in Table 16.4. The optimized (PBE+ U) values for the latter in wurtzite ZnO bulk crystal are 1.94 and 1.95 Å for $\text{O}(n)\text{--Zn}(n)$ and $\text{O}(n)\text{--Zn}(n')$ with $n, n' \rightarrow \infty$, respectively. It is thus clear that the O–Zn interatomic distances are *equal to or slightly larger than* the respective bulk values in the interior region of the nanowires, which indicates a bulk-like coordination. On the contrary, the O–Zn interatomic distances are *significantly smaller* than the respective bulk values for the surface region of the nanowires, which implies the occurrence of substantial *surface reconstruction*. Accordingly, the relaxed nanowires possess a *heterogeneous* structure.

Figure 16.8 shows the variation of the *negative* of the nanowire formation energy $-\Delta H_N$ with the ratio N_s/N where N_s denotes the number of Zn–O pairs residing on the outer surface of the nanowire. The solid symbols in the upper and lower panels of Fig. 16.8 represent the calculated (PBE+ U and HSE+ U^* , respectively) values. In each graph, the solid diamond (square) marks the experimental (calculated) value for $-\Delta H_f$ of wurtzite ZnO bulk crystal. The calculation results (i.e., the solid symbols) are extrapolated to the bulk limit with a model equation given by

$$\Delta H_N = C_b \frac{N - N_s}{N} + C_s \frac{N_s - N_l}{N} + C_l \frac{N_l}{N}, \quad (16.9)$$

which is drawn by the solid line in Fig. 16.8. Here N_l denotes the number of Zn–O pairs located at the edges of the nanowire, C_b , C_s and C_l are the fitting coefficients, whose values are given in Table 16.5. Note that (i) N_l/N , (ii) $(N_s - N_l)/N$, and (iii) $(N - N_s)/N$ give the fraction of Zn–O pairs located (i) at the edges, (ii) on the

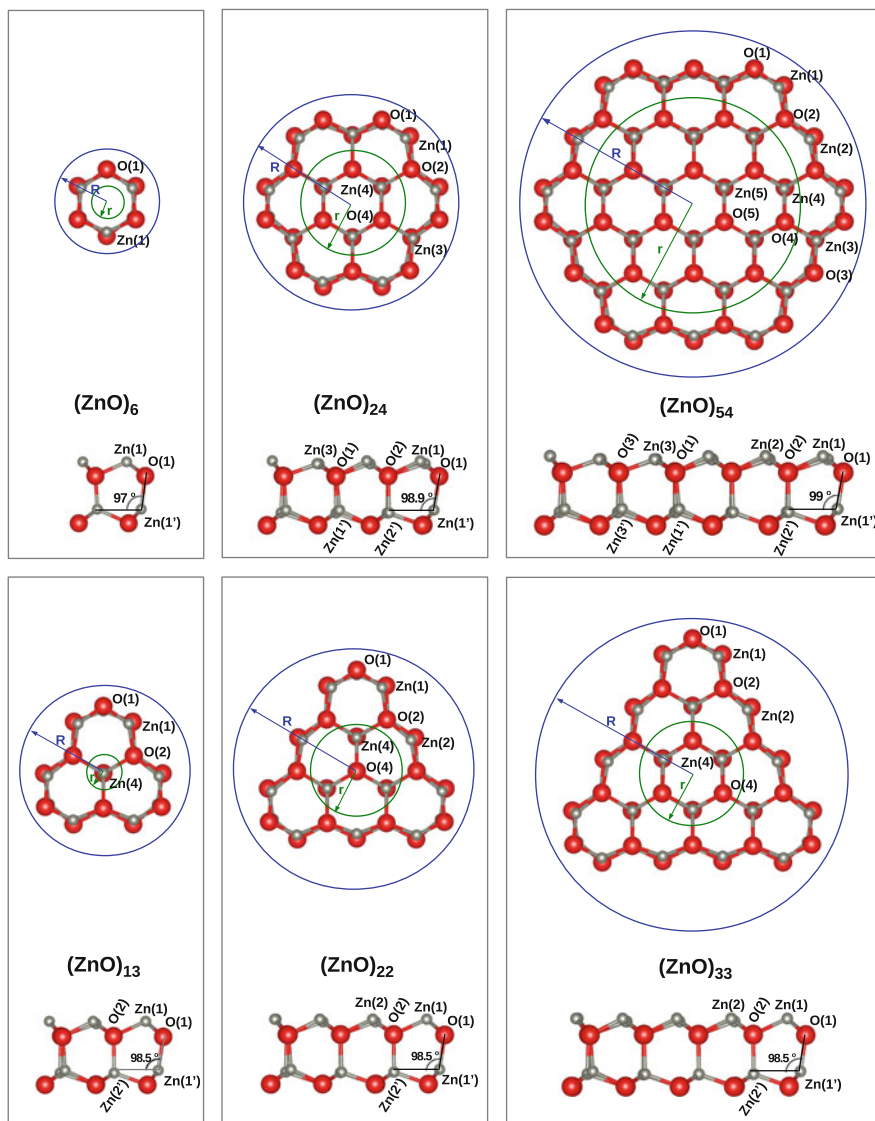


Fig. 16.6 Optimized atomistic structures of ZnO nanowires with hexagonal (*upper panels*) or triangular (*lower panels*) cross sections. The *surface region* of a nanowire consists of atoms confined between the *green* and *blue* circles of radius r and R , respectively, while atoms confined within the *green* circle, with a *bulk-like* environment, reside in the *interior region* of the nanowire

faces, and (iii) in the interior region of the nanowire. In Fig. 16.8, the empty symbols represent the values obtained via (16.9). Hence, the formation energy is obtained for *any* size, i.e., $6 \leq N \leq \infty$, which would enable one to study the size- and shape-dependent energetics of ZnO nanowires. It has recently been

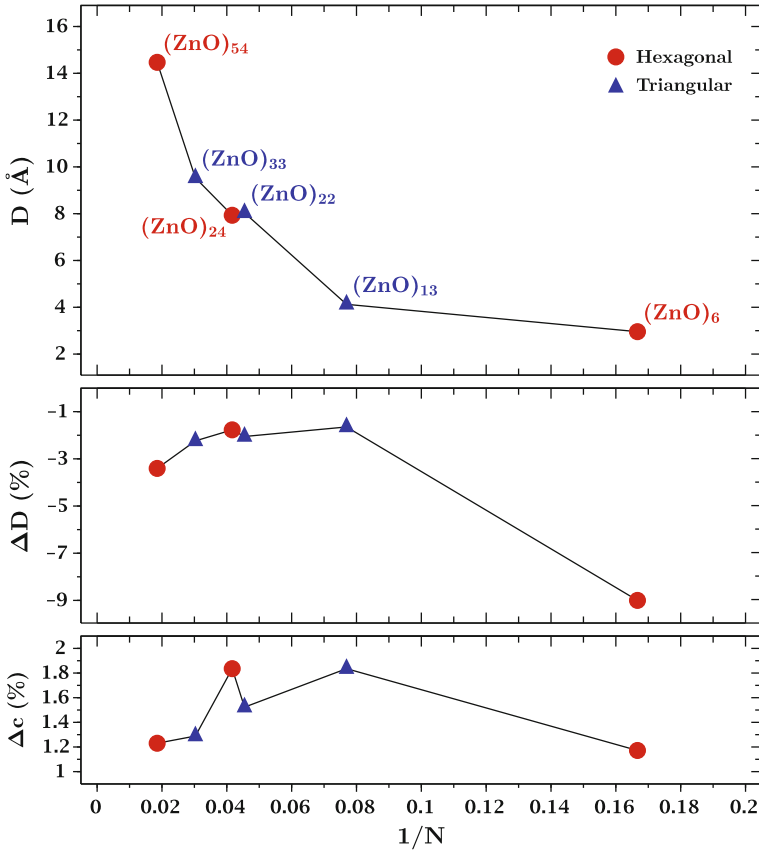


Fig. 16.7 The diameter D (upper panel), the percent change in D (middle panel) and c (lower panel) as a function of the inverse number of Zn–O pairs

demonstrated that the size- and shape-dependent energetics of a variety of nanocrystals [42–44] could be studied in a similar fashion. It is to be noted that this way of studying the nanocrystal energetics makes it possible to extend the computational findings obtained for *small* systems to *large* systems as long as (16.9) is valid not only ultrathin wires such as $(\text{ZnO})_6$ but also ZnO bulk crystal.

16.4 Defect Energetics and Transition Levels in ZnO:Bi Nanowire

In the section, we first employ *defect calculations* to obtain the formation energy for the Bi dopants incorporated into the $(\text{ZnO})_{24}$ nanowire on either the *substitutional* or the *interstitial* sites. We then utilize the latter to determine the charge-state

Table 16.4 The O–Zn interatomic distances (in Å) in $(\text{ZnO})_N$ nanowires

Region		$(\text{ZnO})_6$	$(\text{ZnO})_{13}$	$(\text{ZnO})_{22}$	$(\text{ZnO})_{24}$	$(\text{ZnO})_{33}$	$(\text{ZnO})_{54}$
Surface	O(1)–Zn(1)	1.88	1.85	1.85	1.83	1.85	1.82
	O(1)–Zn(3)				1.89		1.90
	O(2)–Zn(1)		1.92	1.92	1.91	1.92	1.91
	O(2)–Zn(2)			1.88		1.88	1.87
	O(3)–Zn(3)						1.86
	O(1)–Zn(1')	1.83	1.83	1.83	1.81	1.83	1.81
	O(2)–Zn(2')		1.99	1.97	1.96	1.99	1.95
	O(3)–Zn(3')						1.80
Interior	O(4)–Zn(4)			1.95	1.94	1.94	1.92, 1.94
	O(5)–Zn(5)						1.95
	O(5)–Zn(4)						1.96
	O(4)–Zn(4')			2.00	1.97	1.99	1.95, 1.97
	O(5)–Zn(5')						1.96

The respective values in wurtzite ZnO bulk crystal are 1.94 and 1.95 Å for O(n)–Zn(n) and O(n)–Zn(n') with $n, n' \rightarrow \infty$, respectively

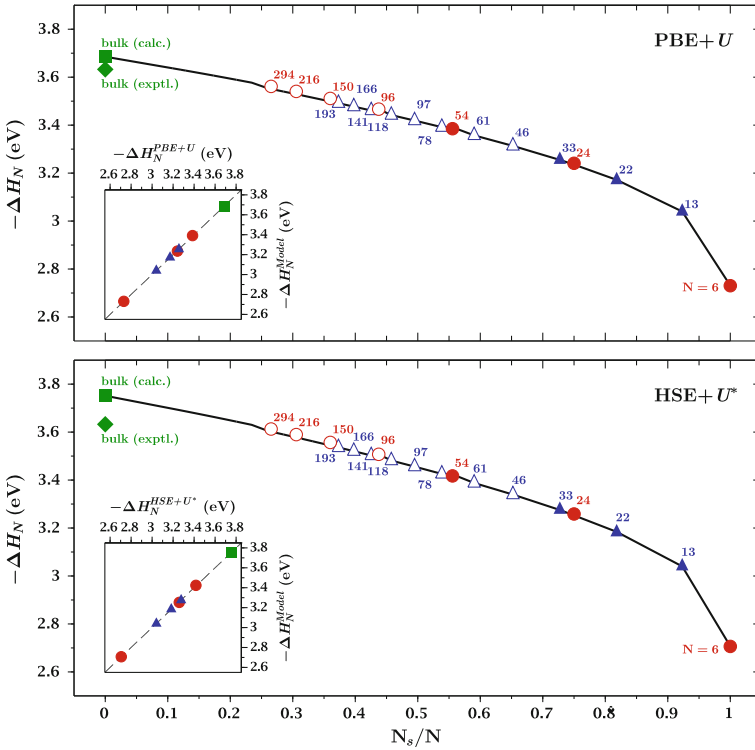
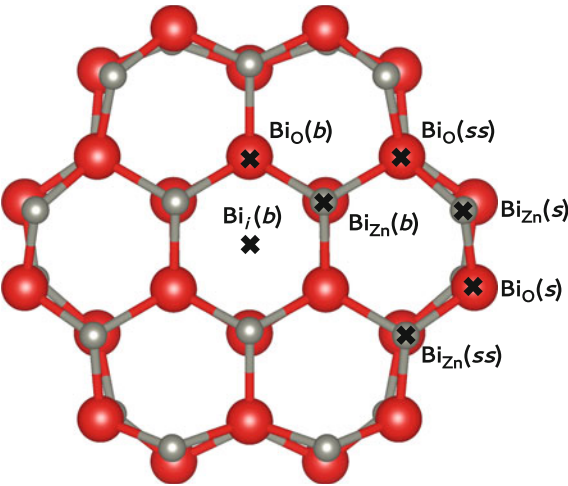


Fig. 16.8 The formation energy ΔH_N of $(\text{ZnO})_N$ nanowires obtained via the PBE+ U (upper panel) and HSE+ U^* (lower panel) calculations versus the ratio of the number of Zn–O pairs located on the outer surface of the nanowire N_s to N

Table 16.5 The values of C_b , C_s and C_l (in eV per formula unit) introduced in (16.9)

Cross section		C_b	C_s	C_l
Hexagonal	PBE+ U	3.684	3.264	2.731
	HSE+ U^*	3.750	3.277	2.707
Triangular	PBE+ U	3.685	3.201	2.340
	HSE+ U^*	3.752	3.213	2.286

Fig. 16.9 The cross marks represent possible doping sites for doping $(\text{ZnO})_{24}$ nanowire with Bi



transition energies in order to reveal the *electrical* behavior of the Bi dopant. For the location of the dopant in the host nanowire, possible doping sites are displayed in Fig. 16.9, which are represented by the cross marks. The latter consist of a number of interior and surface sites where Bi incorporation would cause the following extrinsic defects:

- $\text{Bi}_{\text{Zn}}(b)$, i.e., the dopant replacing a Zn atom in the *bulk-like* region
- $\text{Bi}_{\text{O}}(b)$, i.e., the dopant replacing an O atom in the *bulk-like* region
- $\text{Bi}_i(b)$, i.e., the dopant placed on the *interstitial* site at the center of the nanowire
- $\text{Bi}_{\text{Zn}}(ss)$, i.e., the dopant replacing a Zn atom in the *subsurface* region
- $\text{Bi}_{\text{O}}(ss)$, i.e., the dopant replacing an O atom in the *subsurface* region
- $\text{Bi}_{\text{Zn}}(s)$, i.e., the dopant replacing a Zn atom on the *surface*
- $\text{Bi}_{\text{O}}(s)$, i.e., the dopant replacing an O atom on the *surface*

Note that only *inequivalent* sites are considered. Furthermore, the sites where Bi incorporation is found to lead to an *unstable* configuration are not represented.

Figure 16.10 displays the plots of calculated formation energy as a function of the Fermi level E_F for the foregoing surface (lower panels) and interior (upper panels) defects, cf. Fig. 16.9, under O-poor (left panels) and O-rich (right panels) conditions. The valence band maximum (VBM) and conduction band minimum (CBM) of the nanowire are marked by the vertical lines. The defect (charge-state)

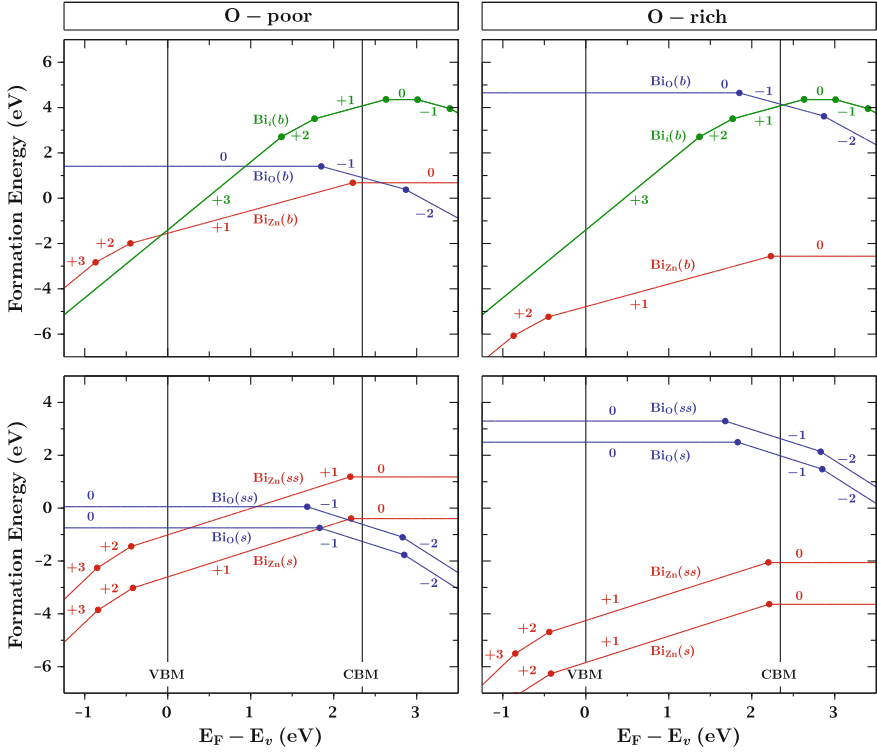


Fig. 16.10 The plots of formation energy as a function of the Fermi level E_F for the surface (lower panels) and interior (upper panels) defects, cf. Fig. 16.9, under O-poor (left panels) and O-rich (right panels) conditions. The valence band maximum (VBM) and conduction band minimum (CBM) are marked by the vertical lines. The defect (charge-state) transition energies are indicated by the solid circles

transition energies are indicated by the solid circles. Figure 16.11 is plotted in order to represent the transition energies (horizontal line segments) relative to the band edge energies. Table 16.6 lists the energy differences $E_c - \varepsilon(q_1/q_2)$ (for donor-like defects) and $\varepsilon(q_1/q_2) - E_v$ (for acceptor-like defects), where E_c and E_v denote the CBM and VBM energies, respectively.

Analyzing the results presented in Fig. 16.10, the formation energy is negative for the various foregoing defects for some range of the Fermi level. Thus, it is likely that many of these defects will form spontaneously under suitable conditions. Note that a defect forms *spontaneously* if its formation energy is *negative*, i.e., $\Delta H_f \leq 0$. However, one should also notice that the formation energy of $Bi_{Zn}(s)$ is almost always *lower* than that of the other defects regardless of the value of the Fermi level as well as the atomic chemical potentials. The only exception to this is that $Bi_O(s)$ has a lower formation energy for a narrow range of E_F around the CBM energy (as in *n*-type ZnO). Accordingly, it is most likely that $Bi_{Zn}(s)$ dominates in the defect

Table 16.6 The calculated (PBE+*U*) values for $E_c - \varepsilon(q_1/q_2)$ and $\varepsilon(q_1/q_2) - E_v$

	Defect	q_1	q_2	$E_c - \varepsilon(q_1/q_2)$ (eV)	$\varepsilon(q_1/q_2) - E_v$ (eV)
Donor-like	Bi _i (<i>b</i>)	+++	++	0.98	
		++	+	0.58	
	Bi _{Zn} (<i>b</i>)	+	0	0.12	
	Bi _{Zn} (<i>s</i>)	+	0	0.14	
				(0.10)	
	Bi _{Zn} (<i>ss</i>)	+	0	0.15	
Acceptor-like	Bi _O (<i>b</i>)	0	–		1.85
	Bi _O (<i>s</i>)	0	–		1.83
					(2.61)
	Bi _O (<i>ss</i>)	0	–		1.68

The values given in parentheses are obtained via the HSE+*U*^{*} calculations

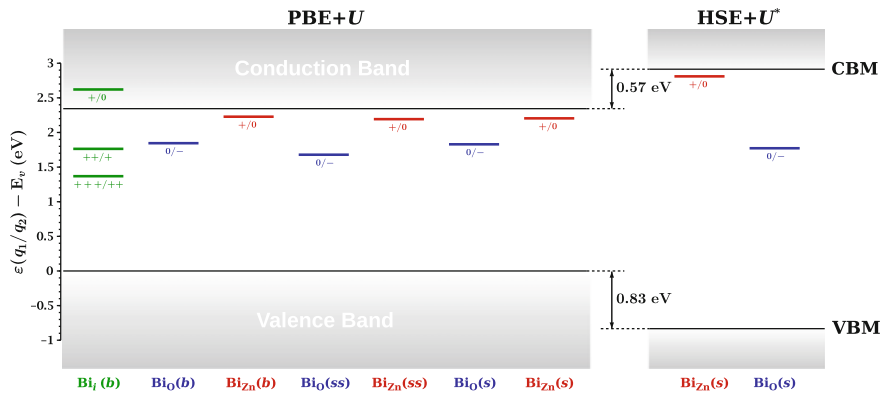


Fig. 16.11 The positions of electrical levels, i.e., the defect transition energies $\varepsilon(q_1/q_2)$, are indicated by the horizontal line segments

structure of ZnO:Bi nanowire. If we restrict our attention to the defects forming in the bulk-like region and compare the interior defects among themselves, cf. the upper panels in Fig. 16.11, we see that Bi_{Zn}(*b*) has the lowest formation energy relative to other defects forming in the bulk-like region. Thus, one could conclude that the incorporation of Bi as a substitutional dopant on the Zn sites is more likely, compared to the rest of the alternatives. This conclusion is clear for O-rich condition. However, the situation is more subtle for *n*-type nanowire under O-poor conditions. In the latter case, both Bi_{Zn}(*s*) and Bi_O(*s*) have *negative* formation energies so that the Bi atoms substitute into not only Zn but also O sites. It is also interesting to note that a segregation behavior could be attributed to Bi in ZnO nanowire because the surface defects have lower formation energies compared to the respective bulk-like defects.

The defect transition energies derived from the formation energies, i.e., the positions of electrical levels, are presented in Fig. 16.11 which is accompanied by

Table 16.6 where the numerical values are given for $E_c - \varepsilon(q_1/q_2)$ and $\varepsilon(q_1/q_2) - E_v$. It is clear that the donor-like defects $\text{Bi}_{\text{Zn}}(b)$, $\text{Bi}_{\text{Zn}}(ss)$, and $\text{Bi}_{\text{Zn}}(s)$ all form *shallow* electrical levels (since $E_c - \varepsilon(q_1/q_2) \sim 0.1$ eV) whereas the acceptor-like defects $\text{Bi}_{\text{O}}(b)$, $\text{Bi}_{\text{O}}(ss)$, and $\text{Bi}_{\text{O}}(s)$ all form *deep* electrical levels (since $\varepsilon(q_1/q_2) - E_v > 1$ eV). Note that the charge state of $\text{Bi}_i(b)$ is + in *n*-type ZnO whose Fermi level is *near* to the CBM, indicating that $\text{Bi}_i(b)$ would act as a *shallow* donor too. Thus, Bi forms either a shallow donor or a deep acceptor in the ZnO nanowire regardless of the location of doping site. This indicates that the incorporation of Bi in ZnO nanowire would not produce *p*-type doping *even if* Bi substitutes into the O site as an acceptor. This analysis makes it clear that Bi in ZnO electrically acts as a donor.

Since the band gap is not well-described in semilocal density approximations [35], which is known to cause problems in the defect calculations performed for bulk ZnO, e.g., [45], we compute the transition energies for $\text{Bi}_{\text{Zn}}(s)$ and $\text{Bi}_{\text{O}}(s)$ within HSE+ U^* description. Here it is of interest to see if the energy differences given in Table 16.6 vary or not while the band gap increases from 2.35 eV (PBE+ U) to 3.75 eV (HSE+ U^*). As seen in Fig. 16.11, the position of $\varepsilon(0/-)$ of $\text{Bi}_{\text{O}}(s)$ remains *relatively unchanged* on the *absolute* scale whereas $\varepsilon(+/0)$ of $\text{Bi}_{\text{Zn}}(s)$ *roughly follows* the upward shift of CBM. Thus, the difference $E_c - \varepsilon(q_1/q_2)$ remains relatively unchanged in Table 16.6 whereas there is a significant increase in the value of $\varepsilon(q_1/q_2) - E_v$. The latter indicates that the band edge correction obtained via the HSE+ U^* description makes the electrical level of $\text{Bi}_{\text{Zn}}(s)$ and $\text{Bi}_{\text{O}}(s)$ become *even* more shallow and deeper, respectively. Hence, the conclusion of the preceding paragraph (i.e., that Bi in ZnO electrically acts as a donor) is *not* susceptible to the band gap error.

Acknowledgments The authors acknowledge financial support from TUBITAK through Grant 114F155.

References

1. C.-L. Hsu, S.-J. Chang, Doped ZnO 1D nanostructures: synthesis, properties, and photodetector application. *Small* **10**, 1613–6829 (2014)
2. L. Schmidt-Mende, J.L. MacManus-Driscoll, ZnO—nanostructures, defects, and devices. *Mater. Today* **10**, 40–48 (2007)
3. S. Xu, Z.L. Wang, One-dimensional ZnO nanostructures: solution growth and functional properties. *Nano Research* **4**, 1013–1098 (2011)
4. G. Wang, Q. Sun, G. Chen, Y. Kawazoe, P. Jena, Vacancy-induced magnetism in ZnO thin films and nanowires. *Phys. Rev. B* **77**, 205411 (2008)
5. R. Qin, J. Zheng, J. Lu, L. Wang, L. Lai, G. Luo, J. Zhou, H. Li, Z. Gao, G. Li, W.N. Mei, Origin of *p*-type doping in zinc oxide nanowires induced by phosphorus doping: A first principles study. *J. Phys. Chem. C* **113**, 9541–9545 (2009)
6. G. Jingyun, Z. Xinzheng, S. Yanghui, Z. Qing, Y. Dapeng, Compensation mechanism in N-doped ZnO nanowires. *Nanotechnology* **21**, 245703 (2010)
7. Y. Li, X. Zhao, W. Fan, Structural, electronic, and optical properties of Ag-doped ZnO nanowires: first principles study. *J. Phys. Chem. C* **115**, 3552–3557 (2011)

8. M.A. Thomas, W.W. Sun, J.B. Cui, Mechanism of Ag doping in ZnO nanowires by electrodeposition: experimental and theoretical insights. *J. Phys. Chem. C* **116**, 6383–6391 (2012)
9. D. Wang, Q. Chen, G. Xing, J. Yi, B.S. Rahman, J. Ding, J. Wang, T. Wu, Robust room-temperature ferromagnetism with giant anisotropy in Nd-doped ZnO nanowire arrays. *Nano Lett.* **12**, 3994–4000 (2012)
10. D.D. Wang, G.Z. Xing, F. Yan, Y.S. Yan, S. Li, Ferromagnetic (Mn, N)-codoped ZnO nanopillars array: Experimental and computational insights. *Appl. Phys. Lett.* **104**, 022412 (2014)
11. Y.-W. Song, K. Kim, J.P. Ahn, G.-E. Jang, S.Y. Lee, Physically processed Ag-doped ZnO nanowires for all-ZnO p–n diodes. *Nanotechnology* **20**, 275606 (2009)
12. J.S. Lee, S.N. Cha, J.M. Kim, H.W. Nam, S.H. Lee, W.B. Ko, K.L. Wang, J.G. Park, J. P. Hong, p-type conduction characteristics of lithium-doped ZnO nanowires. *Adv. Mater.* **23**, 4183–4187 (2011)
13. S.K.S. Parashar, B.S. Murty, S. Repp, S. Weber, E. Erdem, Investigation of intrinsic defects in core-shell structured ZnO nanocrystals. *J. Appl. Phys.* **111**, 113712 (2012)
14. F.X. Xiu, L.J. Mandalapu, Z. Yang, J.L. Liu, G.F. Liu, J.A. Yarmoff, Bi-induced acceptor states in ZnO by molecular-beam epitaxy. *Appl. Phys. Lett.* **85**, 052103 (2006)
15. J.W. Lee, N.G. Subramaniam, J.C. Lee, S. Kumar, T.W. Kang, Study of stable p-type conductivity in bismuth-doped ZnO films grown by pulsed-laser deposition. *EPL* **95**, 47002 (2011)
16. A. Smith, J.F. Baurndard, P. Abelard, M.F. Denanot, ac impedance measurements and V-I characteristics for Co-, Mn-, or Bi-doped ZnO. *J. Appl. Phys.* **65**, 5119–5125 (1989)
17. Ü. Özgür, I. Alivov Ya, C. Liu, A. Teke, M.A. Reshchikov, S. Doğan, V. Avrutin, S.-J. Cho, H. Morkoç, A comprehensive review of ZnO materials, devices. *J. Appl. Phys.* **98**, 041301 (2005)
18. K.-I. Kobayashi, O. Wada, M. Kobayashi, Y. Takada, Continuous existence of bismuth at grain boundaries of zinc oxide varistor without intergranular phase. *J. Am. Ceram. Soc.* **81**, 2071–2076 (1998)
19. H.S. Domingos, J.M. Carlsson, P.D. Bristowe, B. Hellsing, The formation of defect complexes in a ZnO grain boundary. *Interface Sci.* **12**, 227–234 (2004)
20. C. Xu, J. Chun, E. Kim Dong, J.-J. Kim, C B, T. Joo, Electrical properties, near band edge emission of Bi-doped ZnO nanowires. *Appl. Phys. Lett.* **90**, 083113 (2007)
21. G.A. Baraff, M. Schlüter, Electronic structure, total energies, and abundances of the elementary point defects in GaAs. *Phys. Rev. Lett.* **55**, 1327–1330 (1985)
22. S.B. Zhang, J.E. Northrup, Chemical potential dependence of defect formation energies in GaAs: application to Ga self-diffusion. *Phys. Rev. Lett.* **67**, 2339–2342 (1991)
23. C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, C.G. Van de Walle, First-principles calculations for point defects in solids. *Rev. Mod. Phys.* **86**, 253–305 (2014)
24. Ç. Kılıç, A. Zunger, Origins of coexistence of conductivity and transparency in SnO₂. *Phys. Rev. Lett.* **88**, 095501 (2002)
25. Ç. Kılıç, A. Zunger, *n*-type doping of oxides by hydrogen. *Appl. Phys. Lett.* **81**, 73–75 (2002)
26. Ç. Kılıç, A. Zunger, Doping of chalcopyrites by hydrogen. *Appl. Phys. Lett.* **81**, 2007–2009 (2003)
27. Ç. Kılıç, A. Zunger, *n*-type doping and passivation of CuInSe₂ and CuGaSe₂ by hydrogen. *Phys. Rev. B* **68**, 075201 (2003)
28. P. Hohenberg, W. Kohn, Inhomogeneous electron gas. *Phys. Rev.* **136**, B864–B871 (1964)
29. W. Kohn, L.J. Sham, Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133–A1138 (1965)
30. W. Kohn, Nobel lecture: electronic structure of matter—wave functions and density functionals. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1998/kohn-lecture.html. Cited 28 Jan 1999

31. J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996)
32. V.I. Anisimov, F. Aryasetiawan, A.I. Lichtenstein, First-principles calculations of the electronic structure and spectra of strongly correlated systems: the LDA+*U* method. *J. Phys.: Condens. Matter* **9**, 767–808 (1997)
33. A.D. Becke, A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.* **98**, 1372–1377 (1993)
34. A. Seidl, A. Görling, P. Vogl, J.A. Majewski, M. Levy, Generalized Kohn-Sham schemes and the band-gap problem. *Phys. Rev. B* **53**, 3764–3774 (1996)
35. M. Aras, Ç. Kılıç, Combined hybrid functional and DFT+*U* calculations for metal chalcogenides. *J. Chem. Phys.* **141**, 044106 (2014)
36. J. Heyd, G.E. Scuseria, M. Ernzerhof, Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* **118**, 8207–8215 (2003)
37. S.L. Dudarev, G.A. Botton, S.Y. Savrasov, C.J. Humphreys, A.P. Sutton, Electron-energy-loss spectra and the structural stability of nickel oxide: an LSDA+*U* study. *Phys. Rev. B* **57**, 1505–1509 (1998)
38. P.E. Blöchl, Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994)
39. G. Kresse, D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999)
40. G. Kresse, F. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996)
41. H.J. Monkhorst, J.D. Pack, Special points for brillouin-zone integrations. *Phys. Rev. B* **13**, 5188–5192 (1976)
42. Ç. Kılıç, Size- and shape-dependent energetics of transition-metal nanocrystals. *Solid State Commun.* **150**, 2333–2336 (2010)
43. Ç. Kılıç, Energy-distance relation for fcc transition metal nanocrystals. *EPL* **93**, 26004 (2011)
44. M.İ.T. Balta, Ç. Kılıç, Nanocrystal energetics via quantum similarity measures. *Modelling Simul. Mater. Sci. Eng.* **22**, 025009 (2014)
45. A. Alkauskas, A. Pasquarello, Band-edge problem in the theoretical determination of defect energy levels: The O vacancy in ZnO as a benchmark case. *Phys. Rev. B* **84**, 125206 (2011)
46. Z. Fan, D. Wang, P.-C. Chang, W.-Y. Tseng, J.G. Lu, ZnO nanowire field-effect transistor, oxygen sensing property. *Appl. Phys. Lett.* **85**, 5923–5925 (2004)

Chapter 17

Mixed-Phase TiO₂ Nanomaterials as Efficient Photocatalysts

Juying Lei, Hong Li, Jinlong Zhang and Masakazu Anpo

Abstract TiO₂, as one of the most promising photocatalysts, exists different phases such as anatase, rutile and brookite. These phases exhibit different properties and consequently different photocatalytic performances. In addition, mixed-phase TiO₂ have been demonstrated to have enhanced photocatalytic activity relative to pure-phase TiO₂. In the past two decades, many research works have been done on the synthesis of different kinds of mixed-phase TiO₂ and their applications to photocatalysis. In this review, we firstly give an introduction of three main types of TiO₂ phases as mentioned above, including their structural properties, stability, phase transformation and photocatalytic activity. And then we pay more attention on the synthesis of the mixed-phase TiO₂. Six preparation methods are introduced in details, which are hydrothermal method, solvothermal method, microemulsion-mediated solvothermal method, sol-gel method, solvent mixing and calcination method and high-temperature calcination method. Following this, three kinds of applications of the mixed-phase TiO₂ in the photocatalysis field are comprehensively highlighted, including photocatalytic production of hydrogen, reduction of CO₂ and degradation of organic pollutants. As the photocatalytic activity of the mixed-phase TiO₂ is usually higher than the single phase TiO₂, the mechanism for the enhancing effects of the mixed phases are discussed. Finally, the existing problems of mixed-phase TiO₂ are summarized and the application prospects of this kind of nanomaterials are outlooked.

H. Li · J. Zhang

Key Lab for Advanced Materials and Institute of Fine Chemicals,
East China University of Science and Technology, 130 Meilong Road,
Shanghai 200237, People's Republic of China
e-mail: jlzhang@ecust.edu.cn

J. Lei

State Environmental Protection Key Laboratory of Environmental Risk Assessment
and Control on Chemical Process, East China University of Science and Technology,
130 Meilong Road, Shanghai 200237, People's Republic of China

M. Anpo (✉)

Department of Applied Chemistry, Graduate School of Engineering,
Osaka Prefecture University, Naka Ku, 1-1 Gakuencho, Sakai 599-8531, Osaka, Japan
e-mail: anpo@chem.osakafu-u.ac.jp

17.1 Introduction

In recent years, there has been great concern over the many serious environmental problems and lack of natural clean energy resources that we face on a global scale. The increase in world population and industrial growth have all led to accelerated energy consumption and the unabated release of toxic agents and industrial wastes into the air and waterways, leading to pollution-related diseases, global warming and abnormal climatic changes. Thus, environmentally harmonious, clean and safe scientific technologies to address energy as well as pollution and climatic change are major challenges facing most scientists nowadays. Our scientists hope to contribute to the development ecologically clean, safe, and sustainable chemical technologies, materials and processes.

Although there are many approaches to challenge these issues, ever since Honda and Fujishima [1] found the UV light-induced cleavage of water into H_2 and O_2 using a TiO_2 photo-electrode (photosensitizing effect of TiO_2 electrode), there has been enormous interest in the use of TiO_2 and other extended oxide and chalcogenide semiconductors. The purpose is environmental remediation, where organic toxic materials at low concentrations are photocatalytically converted to harmless oxidation products [2–29]. TiO_2 as a photocatalyst has many attractive features; a convenient band gap between its valence band and conduction band of 3.2 eV (ca. 400 nm), high stability, low cost, non-toxicity and good performance in the oxidation of organic pollutants to CO_2 and H_2O . For these reasons, TiO_2 -based photocatalysts have received much attention for various applications in the fields of the generation of clean energy and protection of our environment.

As mentioned above, with its various advantages, TiO_2 has attracted many researchers to perform systematic in-depth studies of it, promoting the application process in various areas related to energy and environment. The related research include studies on photocatalytic mechanism [3–5], regulation of crystal structure and morphology [6–15], improvement of photocatalysis and photoelectric conversion efficiency [16–24], expansion of the optical response range [25–29]. However, traditional TiO_2 has defects in the catalytic performance, which are mainly caused by the following two aspects: on one hand, the large band gap limits its adsorption in the UV region ($\lambda \leq 387$ nm), and the absorption of visible light is almost zero, so that the utilization of the sunlight is low efficiency [3–5]; on the other hand, the recombination of photo-generated electrons and holes is facile and serious, which greatly limits the photocatalytic performance of TiO_2 [30, 31].

Researches have demonstrated that, modification of TiO_2 by organic dye photosensitization [32–36], noble metal deposition [37–41] or doping [42–45] and semiconductor compounding [45–49] can effectively improve the photon utilization and reduce the recombination rate of photo-generated electrons with holes, thereby improving its photocatalytic efficiency. However, the above methods have disadvantages such as poor reaction stability and controllability, high cost, complicated operation process amongst others. In these respects, mixed-phase TiO_2 with relatively simple preparation process, low cost and adjustable crystal type has

attracted much attention. In the mixed-phase photocatalyst, the photo-generated electrons and holes can effectively separate in the two-phase interface, thus reducing their recombination rate. Moreover, the rutile in the mixed phases have a narrower band gap, which can enhance the utilization of visible light to some extent. These advantages can improve the comprehensive properties of TiO₂ photocatalyst. Research results have shown that, adding rutile TiO₂ into pure anatase TiO₂ to form mixed-phase TiO₂ of rutile and anatase, can significantly improve the photocatalytic activity of TiO₂ [50–53]. After these studies, other kinds of mixed-phase TiO₂ have also been studied in depth, such as anatase/brookite [54, 55], anatase/brookite/rutile [56] and brookite/TiO₂(B) [57]. The photocatalytic experimental results of these studies showed that, the photocatalytic activity of mixed-phase TiO₂ is higher than that of single-phase TiO₂. This is because different phases of TiO₂ have different band positions; their combination can effectively promote the separation of photo-generated electron-hole pairs. In addition, catalytic “hot spots” are supposed to be created at the interface of different phases, therefore photocatalysis efficiency can be improved [51, 56].

In this review, first, we would like to briefly introduce three main kinds of TiO₂ phases, and then we will focus on recent advances in the development of mixed-phase TiO₂ photocatalysts, including the synthetic methods for the mixed-phase catalysts, the applications of the catalysts to various photocatalytic reaction systems such as photocatalytic hydrogen production, photoreduction of CO₂ and photocatalytic degradation of organic pollutants. Following this the mechanism of enhanced photocatalytic activity of the mixed-phase TiO₂ will be presented. Finally, we will summarize the existing problems and overview the application prospects of the mixed-phase TiO₂.

17.2 Phases of TiO₂

TiO₂ has eight types of crystal phases, which are rutile, anatase, brookite, TiO₂-B, TiO₂-R, TiO₂-H, TiO₂-II, and TiO₂-III [58]. Among them, rutile, anatase and brookite are naturally occurring oxides of titanium at atmospheric pressure and have been researched mostly for applications [59, 60]. The other five phases have also been investigated [61–64], however, because they are high-pressure phases, they are of minor significance for research and development applications. Therefore, we focus on only the phases of rutile, anatase and brookite in this review.

17.2.1 Structure Properties of Rutile, Anatase and Brookite

The three kinds of TiO₂ phases have different crystallographic properties, which are summarized in Table 17.1 [65].

Table 17.1 Crystallographic property of anatase, rutile, and brookite TiO₂

Crystal structure	Density (g/cm ³)	System	Space group	Cell parameters (nm)		
				a	b	c
Rutile	4.240	Tetragonal	D _{4h} ¹⁴ -P4 ₂ /mnm	0.4584	–	0.2953
Anatase	3.830	Tetragonal	D _{4a} ¹⁹ -I4 ₁ /amd	0.3758	–	0.9514
Brookite	4.170	Rhombohedral	D _{2h} ¹⁵ -Pbca	0.9166	0.5436	0.5135

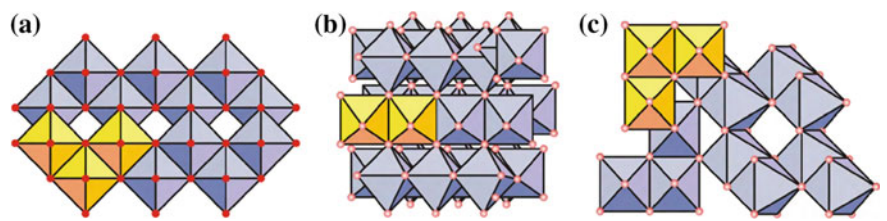


Fig. 17.1 Crystalline structure of **a** anatase, **b** brookite and **c** rutile [68]

In their structures, the basic building block consists of a titanium atom surrounded by six oxygen atoms forming a TiO₆ octahedron in a more or less distorted configuration. In all the three TiO₂ phase structures, the stacking of the TiO₆ octahedra results in threefold coordinated oxygen atoms. The fundamental structural units in these three TiO₂ crystals form from TiO₆ octahedron units and have different lattice configurations as presented in Fig. 17.1. In rutile, TiO₆ octahedra were linked by sharing an edge along the c-axis to form chains. These chains are then interlinked by sharing corner oxygen atoms to form a three-dimensional lattice. Conversely in anatase, the three-dimensional lattice is formed only by edge shared bonding among TiO₆ octahedra. It means that the octahedra in anatase share four edges and are arranged in zigzag chains. In brookite, the octahedra share both edges and corners, forming an orthorhombic structure [66, 67].

In determining these crystal structures and estimating the crystal grain size of anatase, rutile and brookite, the X-ray diffraction (XRD) experimental method is usually used. Anatase peaks in X-ray diffraction occur at $\theta = 12.65^\circ$, 18.9° , and 24.054° , the rutile peaks are found at $\theta = 13.75^\circ$, 18.1° , and 27.2° while brookite peaks are encountered at $\theta = 12.65^\circ$, 12.85° , 15.4° , and 18.1° ; θ represents the X-ray diffraction angle [69, 70].

17.2.2 Stability and Phase Transformation

Among the three types of phases, rutile is the most stable phase, which won't decompose or undergo a phase transformation even at very high temperatures. However, anatase and brookite are metastable; they can be transformed into

thermodynamically stable rutile when calcined to a certain temperature. There has been extensive research into the phase transfer mechanism of TiO₂ during heating. Shannon [71] described the anatase to rutile transformation from crystallography to be a nucleation and growth process, in which rutile phase firstly nucleates on the surface of anatase and then expands to the bulk. Due to the great differences between anatase and rutile, the transformation is reconstructive, which means that the transformation process involves the breaking and reforming of bonds [72]. In the course of the transformation to rutile, the {112} planes in anatase are retained as the {100} planes in the rutile product. In these planes, Ti and O atoms synergistically rearrange. Most of the Ti atoms move to a new location to form rutile by the breaking of two Ti-O bonds in the TiO₆ octahedron. Therefore, the formation of oxygen vacancies can accelerate the phase transformation and the formation of Ti interstitials inhibits the transformation. This anatase to rutile transformation is a non-equilibrium phase transformation, which occurs in a certain range of temperature (400–1000 °C); thus temperature is closely related to impurities, particle size, surface area etc. Impurities and processing atmosphere strongly influence the phase transition temperature and rate because they can result in different defect structure. In general, impurities such as the oxides of Ce, Li, K, Na, Fe, and Mn which can increase the oxygen vacancies usually promote the phase transformation, while impurities like S, P, and W usually inhibit the phase transformation. A reductive atmosphere such as H₂ and Cl₂ accelerate the transformation and gases which are conducive to the formation of Ti interstitials slow the phase transformation.

Artificial preparation favors the synthesis of anatase nanoparticles, especially the preparation of TiO₂ in aqueous solution [73]; this is because the energy of the three types of phases are quite close. When the nanoparticles are small enough (<13 nm), the small surface free energy will become the determinant for the phase transformation [74]. For TiO₂ nanocrystals smaller than 11 nm, anatase is the most stable phase. When the nanocrystals are bigger than 35 nm, the rutile phase reflects the thermodynamic stability. And the stability of brookite lies between the anatase and rutile. In fact, brookite as a metastable phase, most of its physical parameters are between those of anatase and rutile, such as the band width of anatase, brookite and rutile are 3.19, 3.11 and 3.0 eV respectively [75].

17.2.3 Photocatalytic Activity of Rutile, Anatase and Brookite

Generally, anatase exhibits the highest photocatalytic activity among the three types of phases, which are due to the following reasons:

- (1) The band gap of rutile, anatase and brookite are 3.19, 3.11 and 3.0 eV respectively [75], from which we can see that anatase has the widest band gap. Therefore, the electron-hole pair has more positive or more negative potential, leading to higher oxidation ability [76];

- (2) The surface of anatase has a strong adsorption ability for H_2O , O_2 and OH , which is favorable for high photocatalytic activity because the adsorption capacity of the surface has a great influence on the catalytic activity in the photocatalytic reaction and strong adsorption capacity is beneficial to high activity;
- (3) Compared to rutile and brookite, anatase usually has smaller grain size and larger specific surface area in the process of crystallization, which can improve the photocatalytic activity.

However, because the photocatalytic activity is also greatly influenced by the crystallization process, the above rules are not suitable for all situations. Under the same conditions, when the amorphous TiO_2 crystallize, rutile usually forms large grains with poor adsorption ability and thus has low catalytic activity. But if the rutile can have the same grain size and surface properties as anatase, it can also have high photocatalytic activity. For example, Lee et al. [77] found that anatase TiO_2 can transfer to rutile TiO_2 by laser exposure. During the transformation, the specific surface area and grain size remained unchanged. This resultant rutile photocatalyst exhibited considerable high catalytic activity. Tsai et al. [78] prepared anatase and rutile TiO_2 by different methods and investigated their photocatalytic activity toward degradation of phenol. They found the catalytic activity was related to the preparation methods and calcinations temperature. Under certain conditions, rutile TiO_2 exhibited very high catalytic activity. Therefore, whether anatase or rutile, they are both likely to have high photocatalytic activity which greatly depends on the surface properties and the grain size. In addition, Ohno et al. [79] reported that the photocatalytic activity of different phases of TiO_2 is related to the electron acceptors in the system. When the electron acceptor is O_2 , the photocatalytic activity of anatase is high while that of rutile is relatively low. And when Fe^{3+} is the electron acceptor, rutile showed higher catalytic activity. This is because O_2 as an electron acceptor is quite sensitive to properties of the catalyst materials in the photocatalytic reactions. The surface structure of rutile and/or its low energy band may make it have low transfer efficiency of electrons to O_2 and thus rutile has low catalytic activity when O_2 is the electron acceptor. And because most researchers use O_2 for the studies of photocatalytic reactions, rutile usually exhibits low catalytic activity.

The photocatalytic activity of brookite is quite controversial; the main reasons may be that there have been few studies and experiments on brookite and it is difficult to prepare pure brookite TiO_2 . The general products are mixed-phase TiO_2 , including anatase/brookite [80], brookite/rutile [81] or anatase/brookite/rutile [44].

Recent studies have demonstrated that, the mixed phase composited TiO_2 crystals in an appropriate composition ratio have higher photocatalytic activity than that of single phase TiO_2 crystal. Bacsá et al. [82] found that 100 % anatase or 100 % rutile has relatively lower photocatalytic activity, but a mixture of anatase and rutile in various ratios achieved better catalytic activity than pure anatase or pure rutile; the mixture with 30 % rutile and 70 % anatase showed the best catalytic activity. Thus the two kinds of crystals have a certain synergistic effect for

enhanced photocatalysis. And the commercial photocatalyst P25, which has considerable high photocatalytic activity, is also a kind of mixed-phase TiO₂ rather than pure anatase. Therefore, great efforts have been devoted to investigating the synthesis and photocatalysis applications of mixed-phase TiO₂.

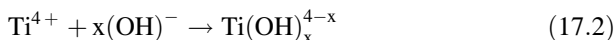
17.3 Synthesis of Mixed-Phase TiO₂ Photocatalysts

It can be observed from the structure and properties of the three types of crystals, that adjusting the experimental parameters of the synthesis can cause conversion among different phase types allowing different mixed-phase TiO₂ photocatalysts to be prepared. At present, there is a variety of methods to prepare mixed-phase TiO₂, such as pulsed laser deposition (PLD), hydrolysis method, hydrothermal method, solvothermal method, micromulsion-mediated solvothermal method, solvent mixing and calcination method (SMC method), high-temperature calcination method, and high temperature gas phase decomposition method as well as others [83]. In the preparation process of mixed-phase TiO₂, with different synthesis methods, the important influence factors for the ratio of mixed phase crystals are different. In all kinds of synthesis methods, the mixed phase ratio, morphology and surface properties can be controlled by optimizing the experimental conditions, such as temperature, pressure, concentration, and types of reagents. The ratio of different crystals in a perfect mixed-phase photocatalyst should be able to be tuned according to the actual needs, and the product should have controllable shape, uniform dispersion, and should be uneasy to agglomerate. The preparation conditions and influence factors will be discussed in the following sub-section, combining concrete preparation methods.

17.3.1 *Hydrothermal Method and Solvothermal Method*

The hydrothermal method is used most commonly in the laboratory preparation of TiO₂ nanomaterials [84–91]. It refers to a method of synthesizing materials in a sealed reaction vessel (such as an autoclave), using aqueous solution as the reaction system, by heating or other means, to create an environment of high temperature and high pressure. In this condition, the vapor pressure of water becomes higher while the density and surface tension decrease and the ion product increases. This leads to dissolution and recrystallization of the substances which are insoluble in water at room temperature. This method belongs to the liquid phase chemical synthesis method. The main factors influencing the properties of the product by the hydrothermal method include vessel volume, reaction temperature, reaction time, and heating rate.

Chen and Sum et al. [92] synthesized highly crystalline pure brookite and two-phase anatase/brookite TiO_2 nanostructures via a simple hydrothermal method with TiS_2 as the precursors in NaOH solutions. The control of the phase composition has been achieved by varying the solution concentration and reaction time. In the reaction system, the TiS_2 underwent a hydrolysis reaction as follows:



At the beginning of the reaction, the hydrolysis of TiS_2 formed Ti^{4+} (17.1). Then in the hydrothermal process, Ti^{4+} reacted with OH^- in the solution (17.2). When the concentration of OH^- is low, the hydrolysis rate of Ti^{4+} is slow and the Ti^{4+} connect with OH^- to form anatase TiO_2 ; when the OH^- concentration is high, the hydrothesis rate of Ti^{4+} is fast leading to the formation of mixed-phase TiO_2 of brookite and anatase. The mass fraction of brookite (B) and anatase (A) can be obtained respectively by calculating with the following formula [93]:

$$W_B = K_B I_B / (K_A I_A + K_B I_B) \quad (17.3)$$

$$W_A = K_A I_A / (K_A I_A + K_B I_B) \quad (17.4)$$

In the formula, I_A and I_B represent the integral intensity of the strongest diffraction peaks in anatase and brookite respectively, and $K_A = 0.886$, $K_B = 2.721$. Figure 17.2a, b show the influences of different NaOH concentrations and different reaction times on TiO_2 crystal. From the graphs, it is observed that with increase of NaOH concentration, the content of brookite increases, and when the reaction time is more than 6 h, the crystalline form of the product did not change significantly. In addition, brookite can be obtained by conversion of sodium titanate. Since Na^+ is able to stabilize the layered structure, when there is no Na^+ in the reaction system, a layered structure is not stable, and the hydrothermal process causes the structure to collapse, forming anatase TiO_2 ; when Na^+ is excessive, the layered structure is very stable and does not easily collapse, ultimately forming titanate; when Na^+ is moderate, the layered structure partially collapses, and part of the layered structure remains unstable, forming brookite (Fig. 17.2c). The above results are consistent with the literature [88].

Zhang et al. [94] introduced NaCl and NH_4OH in the same time into the hydrothermal reaction system, using TiCl_3 as a titanium source. The results showed that the phase composition of the products can be controlled by changing the concentration of NaCl and the volume ratio of NH_4OH to H_2O in the reaction system (Fig. 17.3a). They also believe that Na^+ is able to stabilize the layered titanate structure, which was formed by the hydrolysis of TiCl_3 in aqueous ammonia in their reaction system. In the absence of NaCl, NH_4^+ ions exist in the interlayer of titanate to balance the negative charges of the layered titanate. However, under hydrothermal treatment at high temperature up to 200 °C, NH_4^+ is

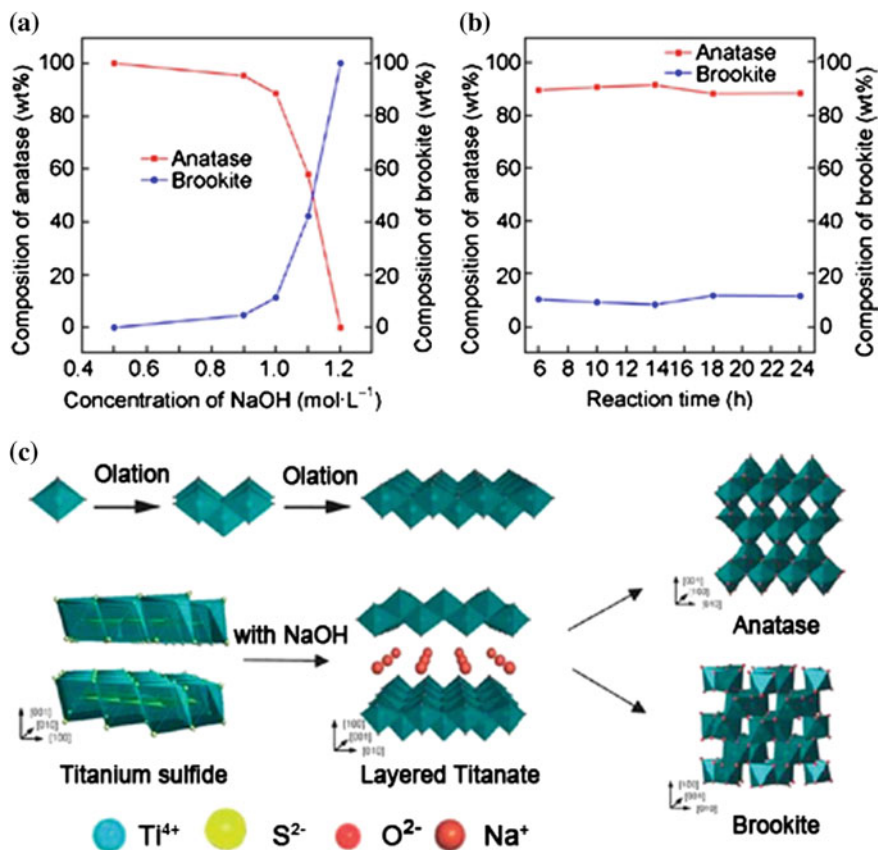


Fig. 17.2 The composition of anatase and brookite in the mixed phase TiO₂ synthesized with **a** various NaOH concentrations and **b** reaction times under hydrothermal reaction [92]; The scheme of formation of anatase and brookite (c) [92]

hydrolyzed to form NH₄OH and is released from the interlayer. The H⁺ ions produced from the hydrolysis of NH₄⁺ ions then induce a hydroxyl condensation reaction to form anatase TiO₂. In the presence of NaCl, Na⁺ can help to stabilize the layered structure with NH₄⁺. Na⁺ and NH₄⁺ exist simultaneously in the interlayer of titanate to balance the negative charges of titanate. When the layered titanate was treated hydrothermally continuously at a high temperature for several hours, the increased hydrolysis of NH₄⁺ led to the collapse of layered structures. However, some bonds in the structure are maintained by Na⁺ ions. Therefore, the structural transformation is delayed which results in the formation of a brookite-like structure [95–97]. The as-formed brookite-like structure expands to form a brookite lattice. Increasing the concentration of NaCl will enhance the powder of the brookite lattice to compete with the anatase lattice. Therefore, the brookite content in the mixed-phase TiO₂ increases with an increase in concentration of NaCl. In addition

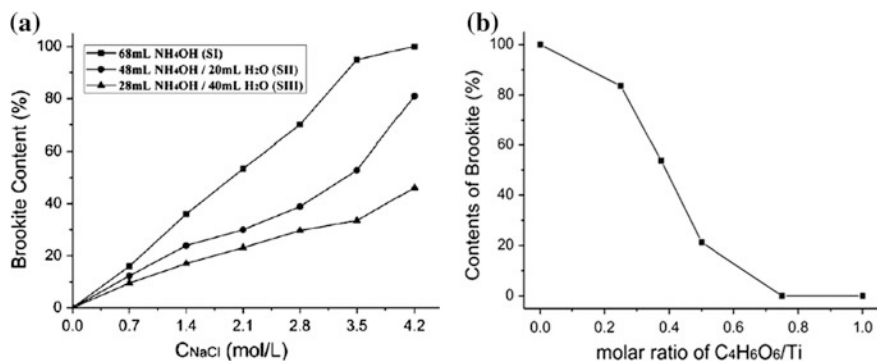


Fig. 17.3 **a** Relationship between the contents of brookite in products and the applied concentration of NaCl [94]; **b** Relationship between the contents of brookite and the applied molar ratio of tartaric acid to $TiCl_3$ [98]

to this research, Zhang et al. [98] developed another facial hydrothermal method for the preparation of anatase/brookite TiO_2 using tartaric acid ($C_4H_6O_6$) as the phase content regulator. Different amounts of $C_4H_6O_6$ were added to the hydrothermal reaction system using $TiCl_3$ as the titanium source and NaOH for adjusting the pH of the reaction solution. The contents of anatase and brookite in the TiO_2 particles were successfully controlled by simply adjusting the molar ratio of $C_4H_6O_6$ to $TiCl_3$ (Fig. 17.3b). The mechanism of the phase evolution between brookite and anatase was explained according to the ligand field theory [99]. They suggested that $C_4H_6O_6$ can chelate with Ti to form a stable titanium complex. When the $C_4H_6O_6/TiCl_3$ molar ratio is less than 0.75 in the reaction system, there are two forms of Ti species. One is the insoluble Ti-contained species $Ti(OH)_4$; the other is the soluble Ti-contained complex $[Ti(OH)_x(C_4H_6O_6)_y]z-$. The amorphous $Ti(OH)_4$ could first transform to layered titanate under hydrothermal treatment, which would be transformed to brookite under a given concentration of Na^+ and OH^- [95]. Due to the large steric hindrance of carboxylic acidic ligands, the $[Ti(OH)_x(C_4H_6O_6)_y]z-$ complexes should combine together by sharing equatorial or apical edges and being arranged in zigzag chains, which benefited the formation of anatase crystallites, thus resulting in an anatase/brookite mixed-phase TiO_2 . They also pointed out that the brookite as the pure phase or the main phase was not obtained when the pH value of the system was less than 9 or the hydrothermal temperature was lower than 180 °C. That is to say, it is difficult to control the contents of anatase and brookite in samples by changing the $C_4H_6O_6/TiCl_3$ molar ratio while $pH < 9$ or temperature < 180 °C in the hydrothermal reaction system.

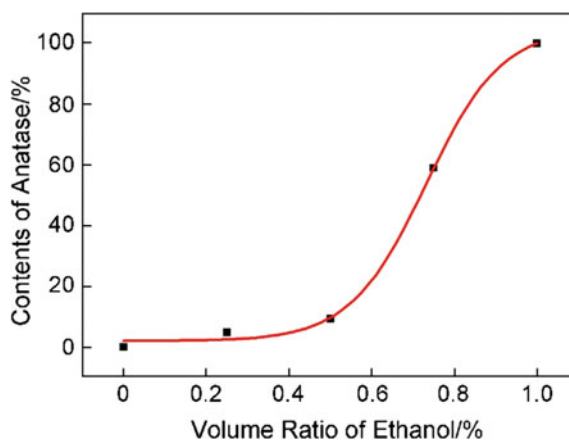
A variation of the hydrothermal method is the solvothermal method wherein organic solvents such as ethanol, glycol and toluene are used instead of an aqueous environment. Similar to the hydrothermal method, in the solvothermal method the crystal types and morphology of the TiO_2 nanomaterials can both be controlled by

regulating parameters such as temperature, pressure inside the system, the reaction time, and the titanium source.

Li et al. [100] synthesized anatase/rutile mixed-phase TiO₂ with different rutile content using the solvothermal method. They obtained mixed-phase TiO₂ crystals by hydrolysis of tetraisopropyltitanate in an acid alcoholic solution and studied the effect of hydrochloric acid on the rutile content in the mixed-phase crystal. The results showed that low H₂O/Ti mole ratio favors the formation of anatase/rutile mixed-phase TiO₂, otherwise it will form brookite/anatase mixed-phase TiO₂. Lei et al. [101] prepared anatase/rutile mixed-phase TiO₂ crystal in a low temperature (80 °C) solvothermal reaction by pre-oxidizing TiCl₃ into Ti⁴⁺ with HNO₃ followed by diluting with urea, water and ethanol. The anatase content in the mixed-phase crystal increases with the increase of ethanol content in the solution (Fig. 17.4). The Scherrer formula is used to calculate the average particle size of anatase and rutile in the mixed-phase crystal; they were both below 10 nm.

The mixed-phase TiO₂ nanomaterials prepared by hydrothermal and solvothermal methods are usually well crystallized and don't need to be calcinated under high temperature. The size and phase type of the product can be regulated by simply changing the experimental parameters such as the type of acid or base in the reaction system, the reaction temperature, the reaction time, the autoclave pressure and so on. Besides these advantages, they also have some drawbacks including the requirement for the equipment to withstand high temperature and high pressure, and experimental safety concerns. Moreover, because the hydrothermal and solvothermal process can't be monitored in real time and usually proceeds without stirring, sometimes the reaction is incomplete and it is difficult to obtain uniform product. Since a variety of organic solvents can be used in the solvothermal method, it is easier to control the morphology and crystalline of the products compared with the hydrothermal method. Thus, the solvothermal method is considered to have better application prospects.

Fig. 17.4 Relationship between the contents of anatase and the volume ratio of ethanol in a solvothermal reaction [101]



17.3.2 *Microemulsion-mediated Solvothermal Method*

The method of microemulsion refers to the process that, firstly, two immiscible solvents form emulsion in the role of surfactants, and then nanomaterials are obtained in the emulsion through nucleation, coalescence, agglomeration, and heat treatment in the oil-in-water or water-in-oil micro-bubbles. In this process, the two immiscible solvents are divided into many micro-reactors (micro-bubbles) by surfactant, in which the reactants are well dispersed and a uniform nucleation occurs. In addition, the growth of the nucleus can be limited in the micro-reactors. Therefore, the particle size and stability of the nanomaterials can be controlled precisely. Compared with other traditional preparation methods, the microemulsion method has been found to have obvious advantages in terms of the synthesis of nanomaterials with superior monodispersity and interfacial properties. Moreover, this method is a very versatile technique which allows the preparation of a great variety of nanomaterials in combination with other techniques. Recently, the combination of the microemulsion method with solvothermal method, namely the microemulsion-mediated solvothermal method, has been used to prepare the mixed-phase TiO_2 crystals [66, 102, 103].

Yan et al. [66] prepared anatase/rutile mixed-phase TiO_2 crystal through the combination of microemulsion and solvothermal method. In the reaction system, tert-octylphenoxypolyethoxyethanol (Triton X-100) as the surfactant and n-hexanol as the co-surfactant were mixed with cyclohexane to serve as the continuous oil phase. Tetrabutyl titanate and $(\text{NH}_4)_2\text{SO}_4$ were dissolved in the hydrochloric acid as the aqueous phase. The aqueous phase was added dropwise to the oil phase, forming a clear microemulsion, which was then solvothermally treated below 120°C for 13 h giving the mixed-phase TiO_2 product. The resultant product has surface area of about $86\text{--}169\text{ m}^2/\text{g}$ and grain sizes of anatase and rutile of about 15 nm and 10 nm respectively. Their results showed that the change of SO_4^{2-} concentration impacts the content of anatase in the mixed-phase crystal (Fig. 17.5). With the increase of SO_4^{2-} concentration, the anatase content in the mixed-phase TiO_2 crystal increased. Different polymorphs can be obtained by affecting the polycondensation of TiO_6 octahedra with SO_4^{2-} (Fig. 17.6a). The SO_4^{2-} interacts with the hydroxy groups on the surface of TiO_6 octahedra to form hydrogen bonds during the hydrothermal reaction process (Fig. 17.6b). Because of the steric effect of SO_4^{2-} , the octahedron with SO_4^{2-} and another octahedron would polycondense along the converse direction in order to decrease the repulsion (Fig. 17.6c), and the orientation of the third octahedron is more conducive to the formation of an anatase nucleus (Fig. 17.6d). When the concentration of SO_4^{2-} stays at a lower level, a rutile structure can be easily formed in order to ensure the lowest free energy in the system. Therefore, controlled adjustment of anatase content in mixed TiO_2 crystal can be achieved by adjusting the concentration of SO_4^{2-} . In the same oil phase system, Zhang et al. [103] prepared mixed-phase TiO_2 crystal with different anatase and rutile content by the microemulsion-mediated solvothermal method, using a mixture of TiCl_3 , urea, and hydrochloric acid solution as the aqueous phase. The results

Fig. 17.5 Relationship between the contents of anatase and the mole percentage of SO₄²⁻ under the micromulsion-mediated hydrothermal method [66]

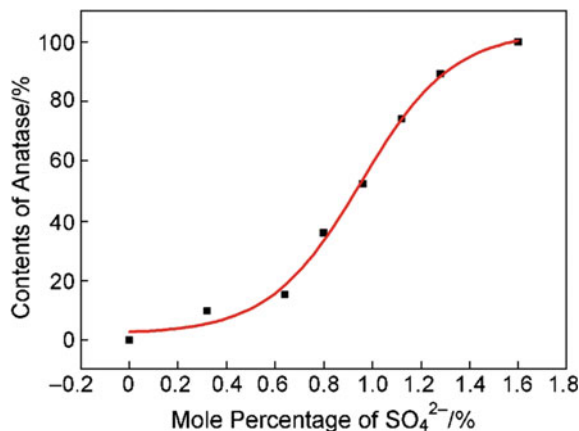
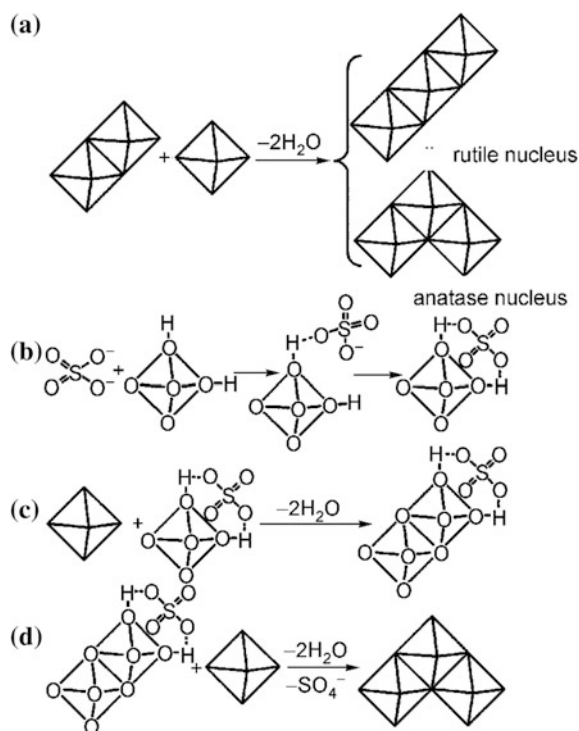


Fig. 17.6 Proposed mechanism: **a** the orientation of the third octahedron determines whether a rutile or an anatase nucleus is formed; **b** interaction between SO₄²⁻ and TiO₆²⁻ octahedral hydroxyls; **c** two TiO₆²⁻ octahedra share edge in the presence of SO₄²⁻; **d** formation of anatase in the presence of SO₄²⁻ [66]



showed that with the increase of urea concentration, anatase content in the mixed-phase TiO₂ increased. The hydrochloric acid in the reaction solution suppressed the generation of anatase and promoted the formation of rutile phase. Different from SO₄²⁻, the Cl⁻ has small radius and weak steric hindrance. TiO₆ octahedra polycondensed in chain to form a rutile structure according to the

minimum energy principle. Therefore, high concentration of Cl^- favors the formation of rutile TiO_2 and a high concentration of SO_4^{2-} favors the formation of anatase TiO_2 .

Although the experiment process is relatively complicated in the preparation of mixed-phase TiO_2 by the microemulsion-mediated solvothermal method, the grain size of the product is small and the phases can be controlled, which gives this method good prospects for development.

17.3.3 Sol-Gel Method

Sol-gel method is one of the commonly used methods for preparation of TiO_2 nanomaterials. The sol-gel method has advantages such as processing at low-temperature and the ability to prepare materials in various shapes, thus, it is one of the most promising techniques to prepare amorphous and crystalline materials. Usually the titanium precursor used in the sol-gel reaction system is titanium alkoxides or titanium halides. Firstly, a sol is prepared by hydrolysis of titanium precursor. After aging the sol for a certain time, a three-dimensional cross-linked gel is obtained; amorphous white powder is obtained after grinding the gel. Finally, crystallized TiO_2 is produced by high temperature calcination.

Scotti et al. [52] synthesized anatase/rutile mixed-phase TiO_2 crystal by a sol-gel method and investigated the influence of HCl/Ti and $\text{H}_2\text{O}/\text{Ti}$ molar ratio on the phase type of the product. The experimental process was as follows: TiCl_4 and tri-block copolymer were dissolved in ethanol then water and HCl were added to adjust the pH of the solution, obtaining a sol. A gel is formed after aging the sol for 3–13 days. After drying and calcination, TiO_2 crystals with different anatase and rutile content were obtained. In the products, pure rutile phase (Fig. 17.7a) showed chestnut burr aggregates of radially growing elongated nanocrystals with average sizes of 10–20 nm in width and 100–200 nm in length; pure anatase phase (Fig. 17.7b) showed aggregates of almost square-ended nanoparticles with average sizes of 5–15 nm. The mixed-phase sample (Fig. 17.7c) was observed to have two phases with the small anatase particles surrounding the chestnut burr aggregates of rutile. The phase content of the products can be controlled by systematically changing the $\text{H}_2\text{O}:\text{Ti}$ (r_w) and $\text{HCl}:\text{Ti}$ (r_a) molar ratios in the reaction system. Results have shown that, in the acidic titanium halide or titanium alkoxide precursor solution, with the increase of H_2O content, concentration of the Ti^{4+} concentration decreased, causing the condensation reaction rate of TiO_6 octahedra to decrease easily forming rutile TiO_2 which has better thermodynamically stability (Fig. 17.7d, e).

Mixed-phase TiO_2 crystals obtained by the sol-gel method are small particles of high purity, but this usually requires long time aging and the particles are easy to agglomerate after high temperature calcination, which may have a significant impact on the photocatalytic activity of TiO_2 .

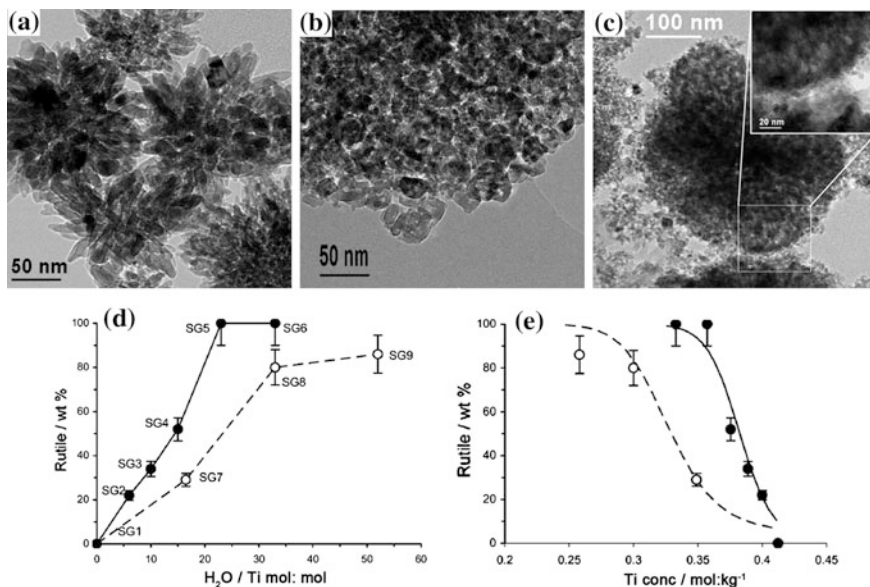


Fig. 17.7 TEM images of **a** pure rutile, **b** pure anatase, and **c** mixed anatase 48 wt% and rutile 52 wt% sample; Plots of rutile content (wt%) in TiO₂ samples: **d** vs. H₂O:Ti molar ratio, **e** vs. titanium concentration [52]

17.3.4 Solvent Mixing and Calcination Method

The solvent mixing and calcination method (SMC method) refers to mixing TiO₂ with different phases in some solvent and then evaporating the solvent completely followed by calcinations at high temperature to make a close contact among different crystal phases. Zachariah et al. [104] synthesized anatase/rutile mixed-phase TiO₂ by the SMC method using anhydrous isopropanol as the solvent. The crystalline size of pure anatase is 10 nm as calculated from the Scherrer equation, and the grain size of pure rutile is 40 nm. In the resultant mixed-phase TiO₂ containing 40 % rutile, the size of anatase is 29 nm, and size of rutile is 47 nm. From Fig. 17.8a, b, pure anatase can be observed consisting of small particles, and the overall particle size is 100–150 nm. The particle size of pure rutile is 300–500 nm (Fig. 17.8c, d). The morphology of the mixed-phase TiO₂ is a mixture of anatase and rutile particles (Fig. 17.8c, f). The phase interface can be seen from the high resolution transmission electron microscopy (HRTEM), indicating there is a close interaction between the two phases. Bojinova et al. [105] prepared an anatase/rutile mixed-phase crystal TiO₂ by the SMC method using ethanol as the solvent. In the mixed-phase TiO₂, the grain size of anatase is 42 nm, and that of rutile is 56 nm, as calculated from the Scherrer formula. Liu et al. [106] made the anatase small particles adsorbed onto rutile nanorods through a layer-by-layer self-assembly technique, using polystyrene sulfonate (PSS) as a medium and finally the PSS was

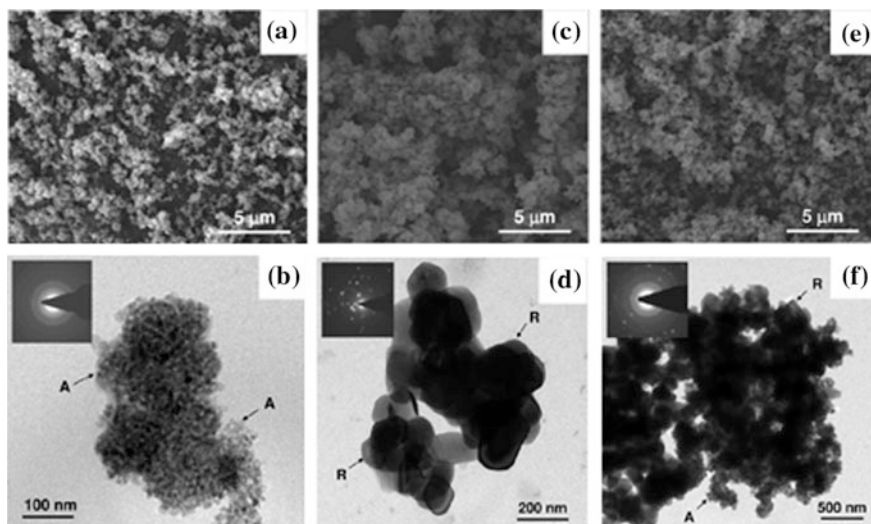


Fig. 17.8 SEM and TEM images of **a, b** anatase, **c, d** rutile, and **e, f** mixed-phase TiO_2 with an optimum rutile content of 40 wt% produced by the SMC method [104]

removed by calcination. The anatase content in the mixed-phase TiO_2 can be controlled by changing the number of loading cycles.

By using the SMC method for the synthesis of mixed-phase TiO_2 , the phase content of the product can be regulated easily. However, because of the final calcination process at high temperature, the TiO_2 is easy to agglomerate. In addition, it's usually difficult to achieve uniform mixing of anatase and rutile, leading to products with many pure anatase aggregates as well as pure rutile aggregates. Therefore, the photocatalytic performance of the mixed-phase TiO_2 prepared by this method is subject to limitations.

17.3.5 High-Temperature Calcination Method

The high-temperature calcination method is one of the earliest methods used to synthesize mixed-phase TiO_2 , which is mostly used to research the phase transition of TiO_2 [84].

Nair et al. [107] obtained mixed-phase TiO_2 crystal with different anatase and rutile contents by the high-temperature calcination method. They firstly prepared small anatase particles by a sol-gel method and then calcinated the anatase particles under elevated temperatures. It was found that the samples calcined at low temperatures ($T < 600\text{ }^\circ\text{C}$) remained pure anatase phase. The phase transformation began at $650\text{ }^\circ\text{C}$. When the temperature was raised to $850\text{ }^\circ\text{C}$, the product becomes

pure rutile. In the mixed phase products, the grain size of anatase was 33.66–51.48 nm, and that of rutile was 45.2–60.6 nm, as calculated by the Scherrer formula. As shown in Fig. 17.9, the particle size of anatase was about 100 nm and rutile was about 200 nm; the mixed-phase particle size was between the above two and increased with calcination temperature. The phase transformation is a process of nucleation and growth. However, there is some dispute about the concrete process. Gouma et al. [108] thought that the rutile firstly formed nuclear on the surface of anatase, then extended to the bulk. Zhang et al. [109] verified the phase transformation of anatase-rutile occurred in the bulk firstly, and then spread to the surface with the increase of calcination temperature by UV Raman spectroscopy. The phase's transformation diagram is shown in Fig. 17.10.

The high-temperature calcination method is very simple and can be used to prepare mixed-phase TiO₂ with perfect polymorphs and tunable phase content. However, the hard agglomeration of TiO₂ is serious, and the particles size is large, resulting in limitations of the photocatalytic activity.

Overall, both in the SMC method and the high-temperature calcination method, a high temperature is required to obtain the mixed-phase TiO₂ or close contact between different crystalline forms. As the electron micrographs and XRD results showed, the grain size of crystal after calcination is significantly larger than that of mixed-phase TiO₂ crystals prepared by the in situ method, and the agglomeration is

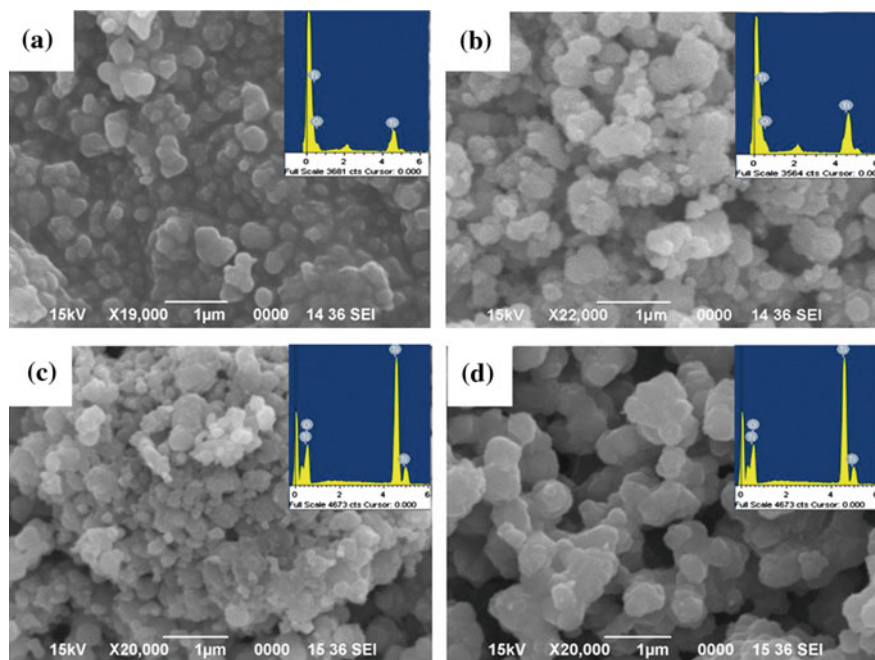


Fig. 17.9 SEM and EDAX images of titania catalysts calcined at different temperature: **a** 600 °C, **b** 700 °C, **c** 750 °C, **d** 800 °C [107]

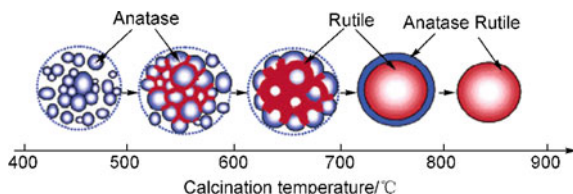


Fig. 17.10 Proposed schemes for the phase transformation of TiO_2 with increasing calcination temperature [109]

Table 17.2 Rate constants for the photocatalytic degradation experiment with mixed-phase TiO_2 synthesized by different methods

Synthesis method	Rate constants/ min^{-1}	References
Hydrothermal method	0.019	[53]
Sol-gel method	2.400	[51]
Microemulsion-mediated	0.030	[75]
Solvothermal method		
SMC method	0.023	[77]
Calcination method	0.003	[79]

also serious, all of which can easily cause decrease of photocatalytic activity. Therefore, in recent years, these methods are used relatively less.

The photocatalytic activity rate constants of mixed-phase TiO_2 prepared by different methods are summarized in Table 17.2. From this table, we can see that the reaction rate of TiO_2 synthesized by in situ methods, such as the sol-gel method and microemulsion-mediated hydrothermal method is generally higher than the high-temperature or SMC method. This is because the latter requires high-temperature calcination, resulting in larger particle size and serious agglomeration that reduce the contact area with the degradation target, and impact the photocatalytic activity. However, the reaction rate of mixed-phase TiO_2 prepared by Zheng et al. [54] was small, which may be because high-temperature calcination leads to a hollow structure, which at the same time, decrease the specific surface area and lower the activity. Photocatalytic activities of mixed-phase TiO_2 prepared under different conditions are also influenced by many other factors: product morphology, particle size, crystal proportion.

In addition to the commonly used approaches mentioned above, there are some new methods to prepare mixed-phase TiO_2 , for example the microwave heating method [110]. The particle size of TiO_2 synthesized by this method is smaller than that prepared by heating with an oil bath. With the development of new technologies, it is believed that there will be many new methods for synthesizing mixed-phase TiO_2 .

17.4 Applications of Mixed-Phase TiO₂ in Photocatalysis

17.4.1 Photocatalytic Hydrogen Production

Hydrogen has been widely recognized as an ideal energy source that could contribute to the replacement of a significant fraction of fossil fuels. However, nowadays the large amount of H₂ demanded is mostly produced from fossil feedstock, mainly by steam reforming of methane. It is therefore mandatory to develop sustainable methods for hydrogen production. Among them, one of the most promising technologies is the photocatalytic hydrogen production, which combines various positive aspects, such as sustainability of the primary energy source (the solar light), the renewability of the starting feedstock and the possible production of by-products with a high added value [111]. And TiO₂ is the most investigated photocatalyst for H₂ production due to its non-toxicity, relatively cheap cost, and excellent stability under the reaction conditions. However, presently, the solar-to-hydrogen energy conversion efficiency of TiO₂ is too low for the technology to be economically sound. The main barriers are the rapid recombination of photo-generated electron-hole pairs as well as backward reaction and the poor activation of TiO₂ by visible light. In response to these deficiencies, many investigators have been conducting research with an emphasis on effective remediation methods like modification of TiO₂ by means of metal loading, metal ion doping, dye sensitization, composite semiconductor, anion doping, and metal ion implantation. In addition to these efforts, mixed-phase TiO₂ have also been explored for hydrogen production [92, 111–115], since multiphasic TiO₂ materials demonstrated higher performances with respect to the correspondent monophasic systems in many photocatalytic processes.

Chen et al. [92] successfully synthesized pure anatase nanoparticles, pure brookite nanoplates, and two-phase anatase/brookite composite by a simple hydrothermal method. Photocatalytic activity of the as-synthesized samples for hydrogen production was investigated in methanol solution. Results have shown that the photocatalytic activity is higher for the two-phase anatase/brookite TiO₂ compared to pure brookite nanoplates, pure anatase nanoparticles and their physically mixed sample, as shown in Fig. 17.11. Moreover, in comparison with the highly active two-phase commercial P25, the synthesized two-phase anatase/brookite TiO₂ is 220 % more active when measured by the H₂ yield per unit area of the photocatalyst surface. Similar results were obtained by Montini et al. [112]. They also prepared TiO₂ nanocomposite with anatase/brookite composition by hydrothermal treatments and the as-prepared anatase/brookite nanocomposites showed higher H₂ production compared to a reference TiO₂ prepared by conventional sol-gel synthesis. Furthermore, they found that the anatase/brookite ratio in the nanocomposite greatly affects the photocatalytic activity in H₂ production.

In addition to the research applying mixed phase anatase/brookite TiO₂ as the photocatalyst for hydrogen production, a great deal of studies have also been done using the mixed phase anatase/rutile TiO₂. Li et al. [113] developed photocatalysts with a tuned anatase/rutile structure by calcination of P25 at different temperatures and

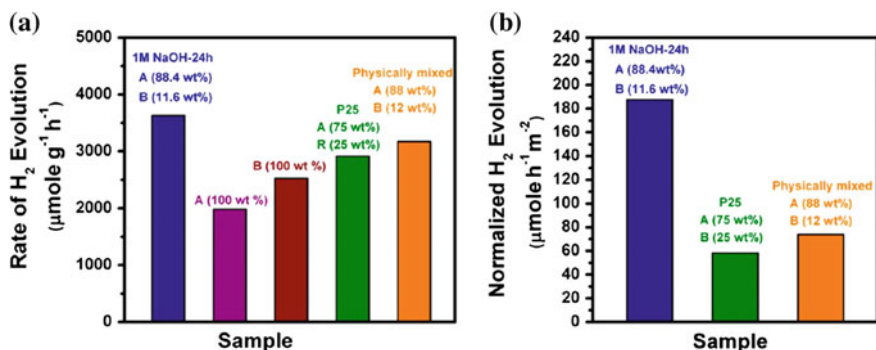


Fig. 17.11 Hydrogen evolved per gram of catalyst per hour under UV–vis irradiation in aqueous methanol solution over 0.3 wt% Pt loaded photocatalysts. A, B and R denote anatase, brookite and rutile respectively [92]

investigated their catalytic activity for hydrogen production by photocatalytic biomass reforming. Surprisingly, they found that the photocatalytic activity of the thermal-treated P25 for hydrogen production can be greatly increased, compared with P25 without any treatment. The overall photocatalytic activity for hydrogen production on thermal-treated P25 can be enhanced up to 3–5 times. It is proposed that the anatase/rutile junction structure is mainly responsible for the improved photocatalytic performance. Their work implies that the photocatalytic performance of TiO₂ could be further improved by elaborately designing the anatase/rutile structure. Similar conclusions were obtained by Amal et al. [114]. They did a systematical study on photocatalytic H₂ evolution over mixed-phase TiO₂ as a function of anatase and rutile phase compositions with methanol as hole scavengers. The TiO₂ nanoparticles they synthesized contain 4–95 mol% anatase, with the remaining being rutile. Synergistic effects in terms of H₂ evolution were observed for a wide range of anatase contents, from 13 to 79 mol%. No synergistic effect was observed for the physically mixed anatase and rutile particles due to insufficient physical contact. Besides the above research on particulate TiO₂ (Fig. 17.12a, b), recently Yu et al. [115] synthesized a kind of anatase/rutile TiO₂ nanofiber photocatalyst (Fig. 17.12b, c). The enhanced H₂ production performance was also observed in the obtained anatase/rutile composite nanofibers. The nanofibers with 45 wt% rutile phase and 55 wt% anatase phase, exhibiting the highest photocatalytic H₂ production rate of 324 mmol h⁻¹ and apparent QE of 20.9 % at 365 nm.

Although the research on the mixed-phase TiO₂ of anatase/rutile all demonstrated that the combination of the two phases can greatly enhance the photocatalytic hydrogen production efficiency, the low utilization rate of visible light is still a big problem which limits its practical application. The relatively narrow band gap of rutile allows some visible light absorption; however this is not viable for practical applications. Further efforts have been devoted to expand the visible light absorption of the mixed-phase TiO₂. Keller et al. [116] reported Au modified TiO₂

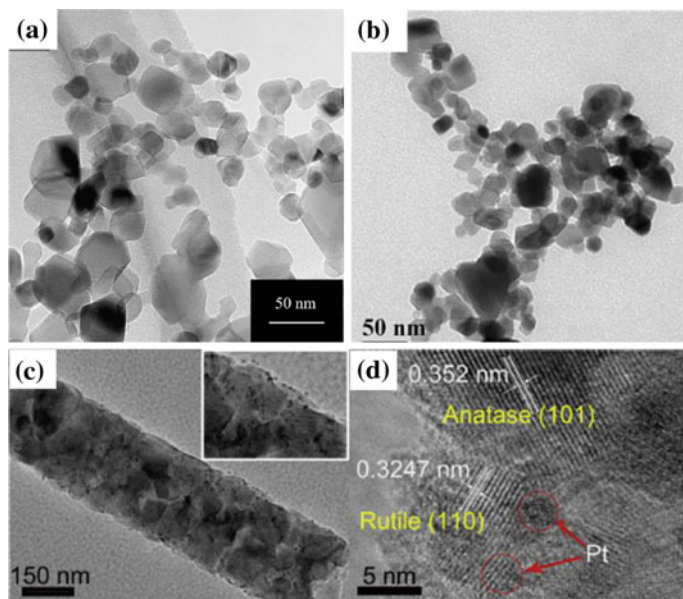


Fig. 17.12 TEM images of particulate anatase and rutile TiO₂; **c, d** TEM and HRTEM images of fibrous anatase and rutile TiO₂ [113]

anatase/rutile catalyst (Fig. 17.13) for hydrogen production. They investigated the light absorption of the catalysts with and without Au deposition by UV–Vis light absorption spectra. Before the Au deposition, the samples with rutile phase whose content was higher than that of P25 showed light absorption extended up to 550 nm, but with a relatively low intensity. After Au deposition, the catalysts showed an obvious modification of the absorption properties. This yielded an additional absorption band around 550 nm, which is attributed to a plasmon resonance phenomenon [117] due to collective oscillations of the conduction electrons located on the 6 s orbital of gold and induced by the incident electromagnetic wave. In addition to this plasmon absorption, deposition of gold also shifted the absorption spectrum deeper into the visible light range, in the 380–450 nm regions. This better light absorption in the visible region was demonstrated to contribute to better results for the hydrogen production under simulated solar light. Furthermore, they pointed out that, there are various factors which were crucial to enhance H₂ evolution efficiency: (i) the surface, crystallographic, and porosity properties of the TiO₂ anatase/rutile catalyst, (ii) the anatase/rutile ratio, (iii) the nature and content of the metallic co-catalyst, (iv) the metal-support interactions, and (v) the relative amount of sacrificial reagent. The influence of these different factors was studied in detail by them. In optimized conditions, important H₂ production efficiency (120 μmol min⁻¹) was obtained over days without deactivation and with very low amounts of sacrificial reagent.

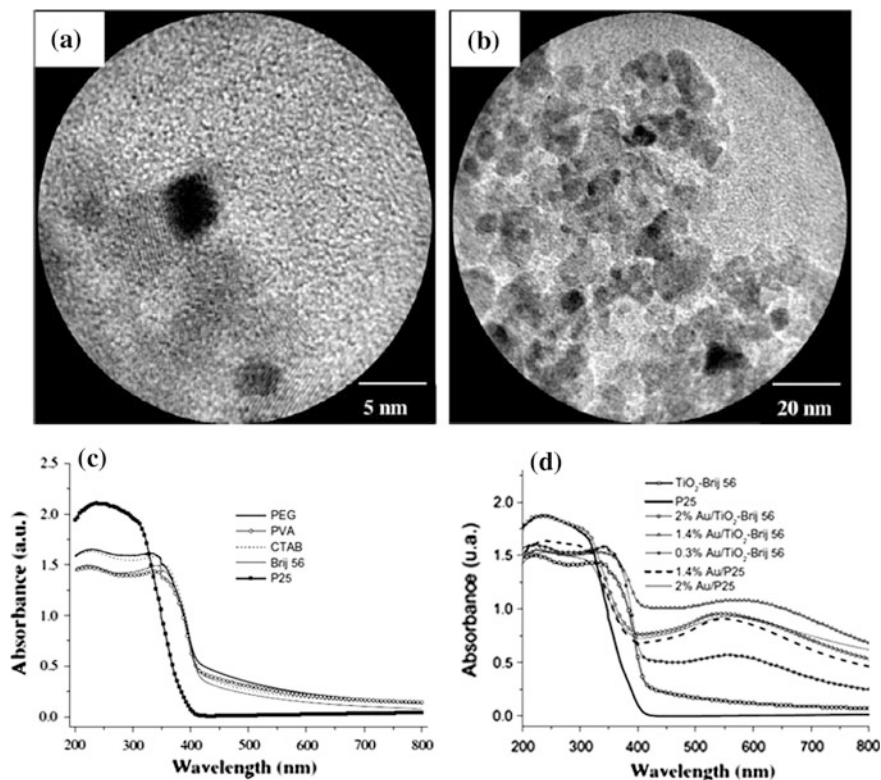


Fig. 17.13 a, b TEM images of mixed-phase TiO_2 photocatalysts with Au deposition; c, d Light absorption properties of mixed-phase TiO_2 photocatalysts without and with Au deposition [116]

17.4.2 Photocatalytic Reduction of CO_2 with Water on Mixed-Phase TiO_2

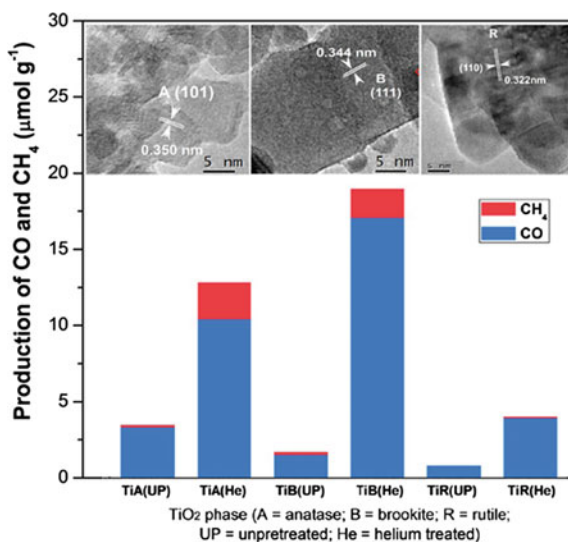
CO_2 is well known to be one of the main causes for the greenhouse effect which leads to global warming. At the same time CO_2 is also a promising carbon resource because it can be converted into various useful chemical compounds and fuels such as CH_4 , CH_3OH , HCOOH . Therefore, in order to reduce the emissions of CO_2 and to achieve a sustainable energy future, novel materials and new technologies have been developed to convert CO_2 . Besides the methods of solar thermo-chemical conversion and electrochemical reduction of CO_2 , solar-activated photocatalytic reduction of CO_2 with water by TiO_2 at room temperature and atmospheric pressure is attractive due to its “green chemistry” and relatively low cost.

During the process of CO_2 photoreduction with H_2O , photo-illumination of the catalyst surface induces the generation of electron-hole pairs in TiO_2 . The excited electrons in the conduction band (CB) of TiO_2 could migrate to the surface and

reduce CO₂. Meanwhile, the holes left in the valence band (VB) of TiO₂ could oxidize H₂O into O₂.

The above photocatalytic process has been demonstrated to be closely related to the phase type of TiO₂. Li et al. [118] investigated the CO₂ photoreduction with water vapor on three TiO₂ nanocrystal polymorphs (anatase, rutile, and brookite). Their experimental results showed that, the photocatalytic reduction activity follows the order: anatase > brookite > rutile. The rutile was the least active, mainly due to the fast e⁻ and h⁺ recombination in rutile. They also investigated the photoreduction activity of the TiO₂ catalysts with a helium treatment. The photoreduction data indicated that helium treated catalysts were more active than non-pre-treated ones. And the catalytic activity for production of CO and CH₄ from photoreduction of CO₂ has the order: brookite > anatase > rutile (Fig. 17.14), where brookite exhibited the highest photocatalytic activity. Thus their study also implies that the brookite phase is a promising material for CO₂ reduction, they performed further research on this topic, including the investigation of brookite-containing mixed phases. They prepared bicrystalline anatase/brookite TiO₂ through a hydrothermal method [119]. The as-prepared bicrystalline TiO₂ were also applied for CO₂ photoreduction in the presence of water vapor for production of CO and CH₄. The photocatalytic activities were compared with those of pure anatase, pure brookite, and a commercial anatase/rutile TiO₂ (P25). The results in Fig. 17.15 showed that bicrystalline anatase/brookite was generally more active than pure anatase, brookite, and P25. The bicrystalline mixture with a composition of 75 % anatase and 25 % brookite showed the highest photocatalytic activity, which was nearly twice as high as that of pure anatase (A100) and three times as high as that of single-phase brookite (B100). The higher activity of bicrystalline anatase/brookite is speculated to be ascribed to the interactions between the anatase and brookite nanocrystals.

Fig. 17.14 The *top* figures are TEM images of different TiO₂ polymorphs: anatase, brookite and rutile (from *left* to *right*). The *bottom* figure is the production of CO and CH₄ on the three different TiO₂ polymorphs [118]



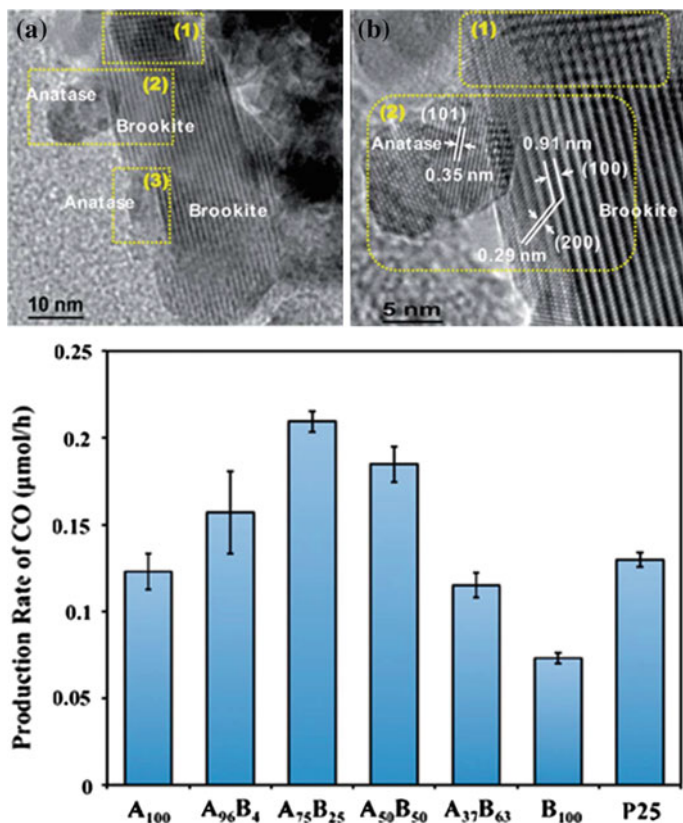


Fig. 17.15 The *top* figures are TEM and HRTEM images of mixed-phase anatase/brookite TiO_2 (A₇₅B₂₅). The *bottom* figure is production of CO on various TiO_2 catalysts (A = anatase, B = brookite, the subscript number are the phase fraction of anatase and brookite) [119]

In addition, the anatase-rich bicrystalline anatase/brookite mixtures are superior to anatase/rutile mixtures P25, indicating the interaction between anatase and brookite seems to be more effective than that between anatase and rutile (as in P25) in the photoreduction of CO_2 .

17.4.3 Photocatalytic Degradation of Organic Pollutants on Mixed-Phase TiO_2

In 1977, Frank and Bard reported for the first time the photocatalytic decomposition of cyanide by TiO_2 in aqueous medium under sunlight [120]. Since then, the application of TiO_2 for photocatalytic degradation of organic pollutants has gained wide attention due to its effectiveness in degrading and mineralizing the recalcitrant

organic compounds as well as the possibility of utilizing the solar UV and visible-light spectrum. Mixed-phase TiO₂ photocatalysts, which usually have better photocatalytic activity than single-phase TiO₂, have also been widely applied for the degradation of organic pollutants.

Many studies have been carried out to examine the photocatalytic degradation of organic dyes in the presence of mixed-phase TiO₂ as photocatalyst. For instance, methyl blue (MB) were degraded in the aqueous solution by anatase/rutile TiO₂ heterojunction nanoflowers under simulated solar light irradiation [121]. It was found that 72 % of MB could be degraded in 120 min with the mixed-phase catalyst prepared in optimized experimental conditions while only 30 % of MB was degraded in the solution with pure anatase. In addition, the TiO₂ nanoflowers showed excellent stability after 9 cycles under the same conditions. These results suggested that the mixed phase anatase/rutile TiO₂ heterojunction nanoflowers have great potential for the future photodegradation of real dye waste water. Since high surface area is conducive to high catalytic activity, mesoporous mixed-phase anatase/rutile TiO₂ were also prepared [122] (Fig. 17.16a) and applied for the

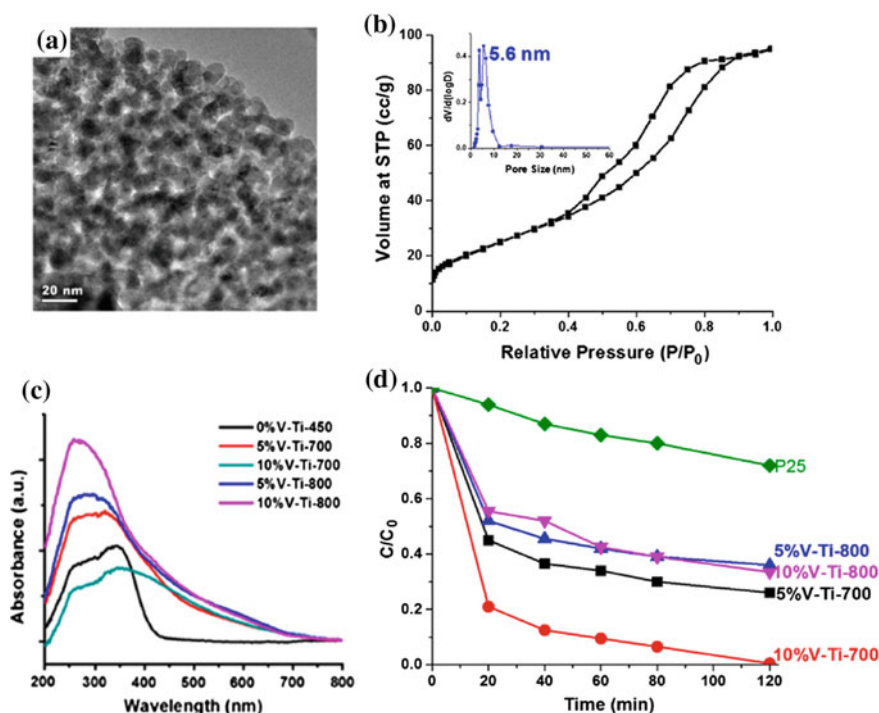


Fig. 17.16 **a** TEM image of mesoporous mixed-phase anatase/rutile TiO₂ (10 %V-Ti-700); **b** N₂ sorption isotherms of 10 %V-Ti-700. The *inset* figure is the BJH desorption pore-size distributions; **c** DR UV-vis spectra for various vanadium doped mixed phase samples; **d** photocatalytic decomposition of 100 mL, 10⁻⁴ MB dye under visible light in 2 h by using 100 mg of 5 % V-Ti-700, 10 %V-Ti-700, 5 %V-Ti-800, 10 %V-Ti-800, and P25 samples, respectively [122]

photocatalytic degradation of MB. The optimum mesoporous catalyst obtained (10 %V-Ti-700) possess surface area up to $94 \text{ m}^2 \text{ g}^{-1}$ and an average pore size of 5.6 nm (Fig. 17.16b). In addition, the catalyst was in situ doped with vanadium which helped to extend the absorption of TiO_2 to the visible light region ($>400 \text{ nm}$) (Fig. 17.16c). As a result, the high surface area, the mixed-phase effect and the vanadium doping qualified the mesoporous mixed-phase anatase/rutile TiO_2 high photocatalytic activity under visible light which are superior to the commercial P25 (Fig. 17.16d).

In addition to the anatase/rutile TiO_2 , brookite/rutile TiO_2 were also explored as photocatalysts for the degradation of dyes by Zhang et al. [81]. They synthesize mixed-phase TiO_2 nanocrystals with a tunable brookite-to-rutile ratio by a facile controllable one pot hydrothermal method. The photocatalytic activities of the resulting TiO_2 nanocrystals were examined in the degradation of Rhodamine B under artificial solar light. The TiO_2 nanocrystals with 38 % brookite and 62 % rutile exhibited the highest photocatalytic activity among the as-prepared samples, which was about 6 times that of the commercial P25. Recently, Chen et al. grew anatase grains on the surface of brookite petals resulting in a brookite/anatase TiO_2 hybrid [55], which appeared to have superior photocatalytic activity for the single phase during the degradation of methyl orange (MO) and 2,4-dichlorophenol (2,4-DCP). The hybrid contains 60 % brookite and 40 % anatase and exhibited the highest activity; the degradation rate constants of which are 2.27 and 1.80 times higher than that of the corresponding physically mixed sample for the degradation of MO and 2,4-DCP, respectively. Following this further Ag modification of the brookite/anatase composite [123] was carried out. Ag^0 clusters with an average diameter of ca. 1.5 nm formed on the surface of the mixed-phase composite (Fig. 17.17a, b). The photocatalytic performance of the as-prepared catalysts was evaluated in terms of the degradation of methyl orange (MO). As shown in Fig. 17.17d, the Ag^0 -loaded brookite/anatase composite had a higher degradation reaction rate constant than that of the pure anatase and brookite, indicating a better photocatalytic activity. Furthermore, by comparing the results in Fig. 17.17c, d, it can be found that the Ag^0 -loaded mixed-phase sample had better catalytic performance than that of Ag free sample. The enhanced photocatalytic reactivity is attributed to the significant improvement in separation of the photo-generated electrons and holes, which can be ascribed to two aspects: (1) synergistic effect of brookite and anatase in the composite, and (2) the schottky barrier at the interface of Ag^0 and TiO_2 .

While most research focused on photocatalytic degradation properties of biphasic TiO_2 catalysts, Que et al. explored brookite/anatase/rutile coexisting TiO_2 composite as the photocatalyst [124]. Their results indicated that in the brookite/anatase/rutile coexisting nanopowders, the brookite and anatase phases were crystallized into irregular nanoparticles $<20 \text{ nm}$ in diameter, whereas the rutile phase was crystallized into single-crystalline nanorods $\sim 20 \text{ nm}$ in diameter and 100–500 nm in length, as shown in Fig. 17.18a. This triphasic TiO_2 catalyst was demonstrated to have better photocatalytic activity during degradation of methyl orange than the biphasic commercial P25 under irradiation of UV light. The sample

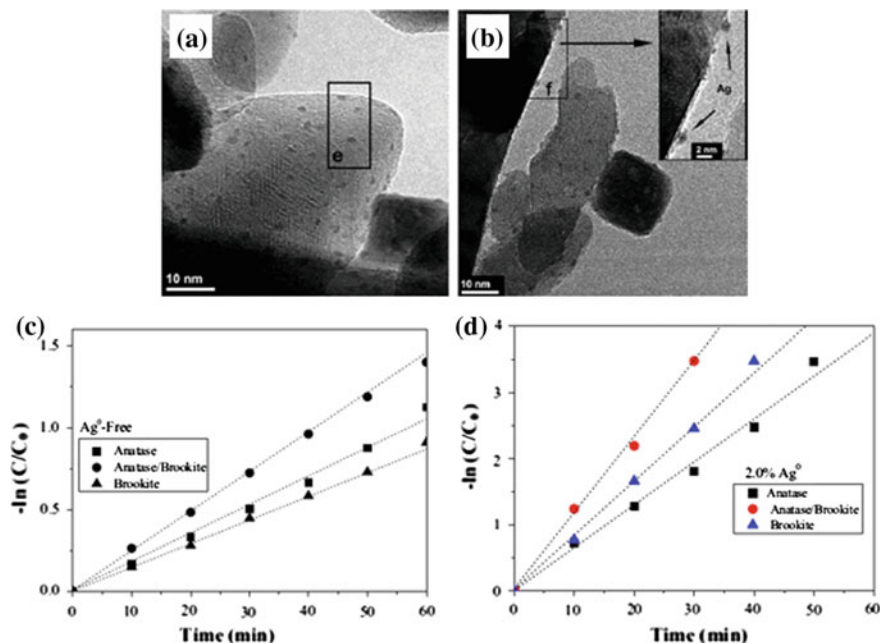


Fig. 17.17 TEM images of 2.0 % Ag⁰-TiO₂ (brookite/anatase) (a, b); photocatalytic degradation of MO with Ag-free (c) and 2.0 % Ag⁰-loaded (d) TiO₂ with different crystal phase composition under UV irradiation [55]

with 29.9 % anatase, 27.9 % brookite, 42.2 % rutile (referred as T2) was shown to have the highest photocatalytic activity, yielding over 90 % bleaching of methyl orange solution in 20 min (Fig. 17.18b). The degradation rate constant k of this sample was 0.10180 min^{-1} , almost twice as high as that of P25 ($k = 0.05397 \text{ min}^{-1}$). In addition, recycle experiments of the photocatalyst were also carried out by using the best photocatalyst sample T2 as shown in Fig. 17.18c. It can be seen that after five recycles, the photocatalytic activity of the sample shows only a slight loss, indicating a good stability of the triphasic TiO₂ catalyst in the photocatalytic degradation process. Similarly, Shao et al. also developed brookite/anatase/rutile triphasic TiO₂ and the catalyst was demonstrated to have excellent photocatalytic performance in the degradation of methylene blue solutions.

Apart from dyes, some other organic compounds have also been used as the model pollutants in the mixed-phase TiO₂ photocatalysis system. For example, phenol and its derivatives have been widely studied as photodegradation target molecules, since they are considered to be one type of primary pollutants due to their danger to organisms at low concentrations [125]. Lu et al. [126] prepared a series of TiO₂ samples with different anatase-to-rutile ratios, and studied the role of crystal phase in photocatalytic oxidation of phenol in water. They found that samples with higher anatase-to-rutile ratios have higher activities for phenol

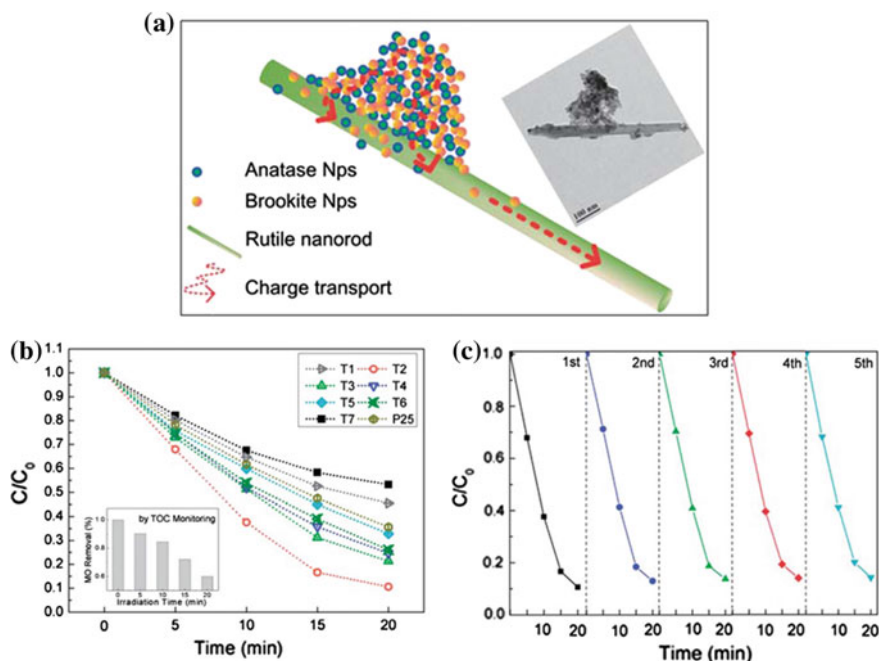


Fig. 17.18 **a** Scheme and TEM image of the brookite/anatase/rutile nanocomposites; **b** Photocatalytic degradation of the MO aqueous solution by using the brookite/anatase/rutile nanocomposites synthesized with various reagent ingredients, and the *inset* shows the removal of MO monitored by TOC with the sample T2 as photocatalysts; **c** Cycling degradation curves of the brookite/anatase/rutile coexisting nanocomposites (sample T2) [124]

degradation. Tian et al. [127] reported Cr-doped TiO_2 (Cr-TiO_2) nanoparticles with anatase and rutile bicrystalline phases for the photocatalytic degradation of 2,4-dichlorophenol (2,4-DCP). It was revealed that Cr^{3+} doping can not only effectively extend the visible light response of TiO_2 (Fig. 17.19a) but also promote the anatase-to-rutile transformation. The photocatalytic activities of different Cr-TiO_2 photocatalysts were evaluated in terms of the photocatalytic degradation of 2, 4-dichlorophenol (2, 4-DCP) under visible light irradiation (Fig. 17.19b). They found appropriate Cr^{3+} doping can significantly enhance the visible-light photocatalytic activity of TiO_2 , which is attributed to the improvement of visible light response as well as suitable anatase-to-rutile ratio. In addition, excess Cr^{3+} doping is detrimental to the improvement of visible light photocatalytic activity, due to the formation of Cr_2O_3 clusters as well as too high a rutile content.

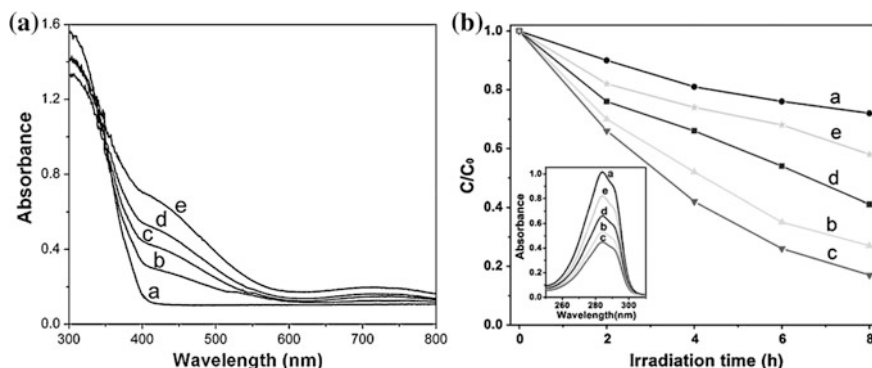
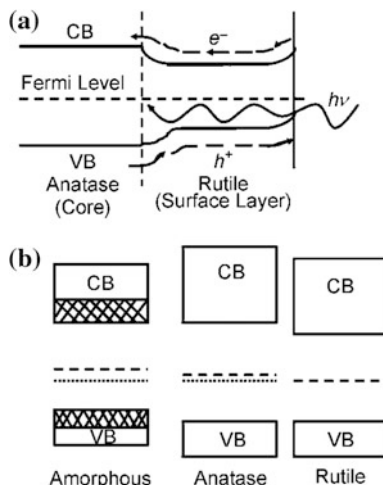


Fig. 17.19 **a** UV-vis DRS and **b** degradation curves of 2, 4-DCP over undoped TiO₂ and different Cr-TiO₂ samples under visible light irradiation: **a** undoped TiO₂, **b** 0.5 % Cr-TiO₂, **c** 1 % Cr-TiO₂, **d** 2 % Cr-TiO₂, and **e** 5 % Cr-TiO₂. The *inset* is the absorption spectra of 2, 4-DCP over different samples after visible light irradiation for 8 h [127]

17.5 Mechanism of the Enhanced Photocatalytic Activities by the Mixed-Phase TiO₂ Photocatalysis

Generally, most of the photo-generated electrons and holes of the single-phase TiO₂ will recombine in the bulk, and only a small amount of them can migrate to the surface to take part in the oxidation-reduction reactions with absorbed molecules, leading to relatively low photocatalytic efficiency. In 1991, Bickley et al. [128] first proposed the mechanism of enhanced photocatalytic activity of P25 which is composed of 80 % anatase and 20 % rutile. They proved that the mixed-phase structure of P25 is an anatase structure coated with a layer of rutile film by TEM. By evaluating the photocatalytic activities, they found that the photocatalytic activity of the mixed-phase TiO₂ was better than the activity of any other single-phase TiO₂ photocatalyst. Because the anatase and rutile have different band gaps, a bending band is generated on the interface of the two phases, as shown in Fig. 17.20. Under light irradiation, the photo-generated electrons migrate from rutile phase to anatase phase, while the holes migrate from anatase phase to rutile phase, so that the electrons and holes can be separated effectively, causing the mixed-phase P25 to have a high photocatalytic activity. Meanwhile, they speculated that the mechanism of the enhanced photocatalytic activity of mixed-phase TiO₂ is actually more complex and requires further study. Subsequently, researchers have studied the mobility direction of the photo-generated carriers, the molecular dynamics characterizations of the mixed-phase interface [129], the band structure of mixed-phase TiO₂ [130, 131], the suitable phase proportion in the mixed-phase TiO₂ and so on.

Fig. 17.20 Schematic band structure diagrams of P25 [128]



In 1995, Datye et al. [132], refuted the view of Bickley et al. [128] on the mixed-phase crystal structure in P25 by X-ray diffraction (XRD) and high-resolution transmission electron microscopy (HRTEM) techniques. Their characterization results showed that the mixed-phase P25 consists of individual single crystal particles of the anatase and the rutile phases of titania rather than particles of anatase covered by a layer of rutile. This view was confirmed by Zhang et al. [133] and Ohno et al. [134]. But they did not pursue further research on the mechanism of the enhanced photocatalytic performance resulted from the mixing of two phases.

In 2002, Kawahara et al. [135] designed a film model to study the migration direction of the photo-generated electrons and holes in anatase/rutile mixed-phase crystal under light irradiation. In their experiments, the mixed-phase crystal TiO_2 film was immersed into an AgNO_3 solution, and it was exposed to light over a period of time under an argon atmosphere. Experimental results showed that there were a large number of Ag particles on the surface of the rutile phase. It is possible that Ag^+ caused a reduction reaction on the rutile surface, indicating that the photo-generated electrons migrated from the conduction band of anatase to rutile (Fig. 17.21).

In 2003, Sun et al. [137, 138] deposited Pt on P25 for the photocatalytic degradation of phenol. However, they found that loading P25 with Pt didn't help to increase phenol decomposition and total carbon removal rates. In addition, they found that Pt appeared on the surface of anatase. Therefore, they proposed a charge separation mechanism for the mixed-phase P25 as "contact of the two phases causing the bending of the conduction band resulting in the holes in anatase transferring to rutile, while the electrons cannot transfer from anatase to rutile", as shown in Fig. 17.22. As a result, holes were concentrated in rutile and electrons were left in the anatase particles, which indicated that oxidation happens mostly on rutile and reduction mostly on anatase, thus the Pt deposition on anatase cannot increase its photo efficiency for phenol oxidation in water.

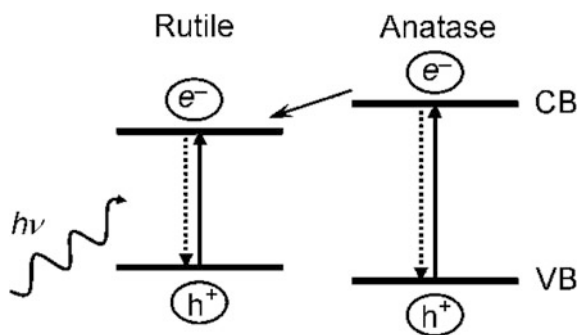


Fig. 17.21 A proposed schematic illustrations showing the migration of the electron transfer from anatase to rutile [135, 136]

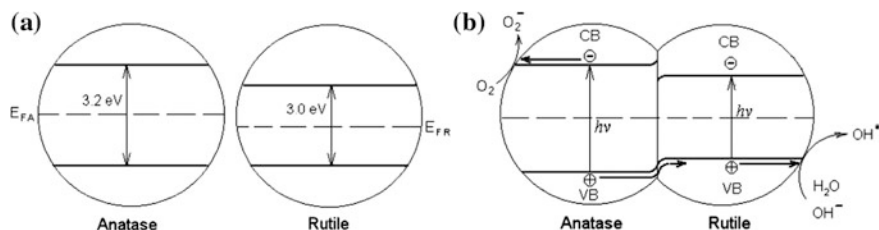


Fig. 17.22 **a** Band structure before anatase and rutile contact each other; **b** Band structure and the separation of photo-generated charge carriers after anatase and rutile contact each other [137]

In the same year, Hurum et al. [51] characterized the charge separation process of P25 by electron paramagnetic resonance (EPR). When P25 mixed-phase crystal was excited by visible light, due to the wide band gap of anatase, no electrons and holes were generated. However, in electron paramagnetic resonance spectroscopy, the electrons trapping sites were found on anatase surface, and it could be concluded that the electrons in rutile generated by visible light excitation transferred to the lower energy anatase lattice trapping sites, leading to a more stable charge separation and enhancing the photocatalytic activity of P25 (Fig. 17.23). Thereafter, this mechanism was confirmed by Liu et al. [139] through the photocatalytic activity experiments of anatase/rutile mixed-phase TiO₂ nanotubes. EPR spectroscopy is a new characterization method developed in recent years, which is a spectroscopic method to detect the unpaired electrons in the sample. Since then, researchers have began to characterize the migration of photo-generated electrons and holes in the mixed-phase TiO₂ using the EPR technique [140].

In 2006, Wang et al. [141] proposed an “antenna mechanism” to explain the mechanism of enhanced photocatalytic activity in mixed crystal phases. The specific mechanism is as follows: in the photocatalytic degradation process, TiO₂ absorbs light to produce photo-generated electrons and holes, which then take part

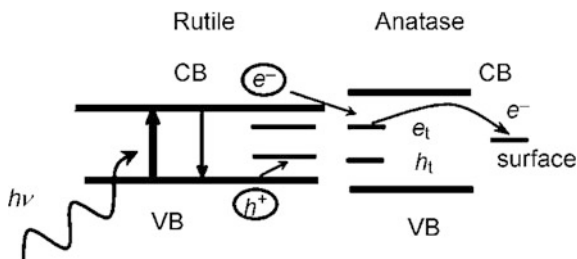


Fig. 17.23 A proposed schematic illustrations showing the charge separation using EPR [51]

in the oxidation-reduction reactions with the target molecules adsorbing on the catalyst's surface; However, the depth in liquid which light can reach is limited, so that the deeper the liquid molecules, the less light they receive. Therefore, the particles irradiated by light can absorb photo energy, this energy can then be transferred from to particles deeper in the liquid, finally oxidation-reduction reaction takes place on the particles which light cannot reach, enhancing the photocatalytic activity. In the process of photo energy being transferred to the photocatalyst deeper in the liquid, the long-chain particles act as an antenna system transferring the photon energy from the location of light absorption to the location of reaction, thus this is called an "antenna mechanism" (Fig. 17.24).

However, the research objects of the above-mentioned theoretical models are all P25 which has a fixed phase ratio of anatase and rutile. Because mixed-phase TiO_2 with different phase ratios have different photocatalytic activities, none of these existing models can adequately explain the physical phenomenon responsible for the existence of an optimum phase content in mixed-phase nanocrystalline TiO_2 corresponding to the maximum photocatalytic activity. Therefore, further research is needed to investigate the photocatalytic mechanism of mixed-phase TiO_2 with different phase ratios.

In 2008, Zachariah et al. [104] proposed a new mechanism model, which takes the factors of crystallite size distribution and phase composition into account. This mechanism overcomes the limitations of the existing models proposed earlier in the literature in terms of explaining the phenomenon of an existing optimal phase ratio in the mixed-phase TiO_2 for the best photocatalytic activity. The mechanism is based on the charge separation mechanism proposed by Sun et al. [137]. In the mechanism, when the size of nano TiO_2 is identical (Fig. 17.25a), the band gap of

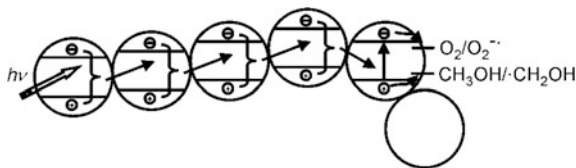
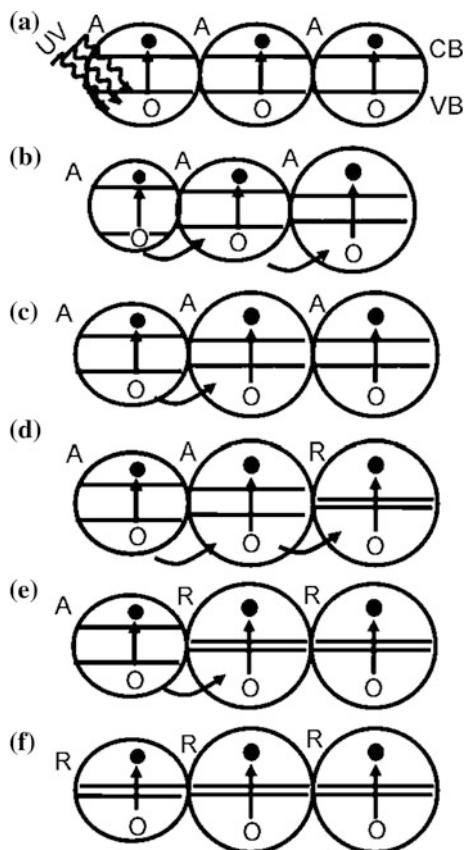


Fig. 17.24 Antenna effect by network structure leading to enhanced photocatalytic activity [141]

Fig. 17.25 The newly proposed mechanism model based on the band-gap variation in the connected nanocrystallites as a function of the size distribution and the phase involved [104].

A, R, CB, and VB represent anatase-TiO₂, rutile-TiO₂, the conduction band, and the valence band, respectively [104]



adjacent anatase crystallites is the same, so there is no driving force for the migration of photo-generated holes, resulting in high recombination rate of photo-generated charges and low photocatalytic activity. If the sizes of anatase crystallite are different and are below the critical size (Fig. 17.25b, c), the band gaps of the connected crystallites would depend on their size. Thus, photo-generated holes in one crystallite can easily escape into the other leading to effective charge separation. When a small amount of rutile was mixed with anatase (Fig. 17.25d), due to the different band gap, photo-generated holes will transfer from anatase to rutile, so that the photocatalytic activity is further enhanced. However, if the rutile content exceeds a certain amount, the migration of photo-generated holes in mixed-phase crystal TiO₂ is limited (Fig. 17.25e), and the photocatalytic activity begins to decline. Given that the crystallite sizes of pure rutile are above the critical size (Fig. 17.25f), their band gap values are the same, then the migration of photo-generated holes does not occur, and the photocatalytic activity is minimum. The key point of this mechanism is that whether the TiO₂ is mixed-phase crystalline or

single crystalline, as long as the band gap values are different, the photo-generated electron-hole pairs will be separated, thereby improving the photocatalytic activity.

In summary, the above mechanisms for the enhanced photocatalytic performance of mixed-phase TiO_2 all involve the migration behavior of photo-generated electrons and holes. In these mechanisms, the mixed-phase crystal structure enables the effective separation of photo-generated electrons and holes, while the specific paths of migration are still controversial. The determining factors for photocatalytic activity of mixed-phase crystal TiO_2 include grain size, structural forms (core-shell structure, cladding structure, random composite, etc.), phase composition. These factors lead to the emergence of bending band, trapping sites and other phenomena in the mixed-phase crystal. Therefore, under different experimental conditions, the obtained mixed-phase TiO_2 differ in photocatalytic mechanism, so further research is still required.

17.6 Conclusion and Outlook

Research in the synthetic methods and applications in photocatalysis of mixed-phase TiO_2 have made great progress since the discovery of photocatalytic water splitting and degradation of organic pollutants by TiO_2 nanomaterials in the 1970s. The superior photocatalytic performance of mixed-phase TiO_2 to single-phase TiO_2 has been well received and the mechanism of the enhanced photocatalytic activities by the mixed-phase TiO_2 photocatalysis has also been widely studied. However, there are still many problems remaining to be resolved. For example, (1) Particulate mixed-phase TiO_2 nanomaterials are prone to agglomeration, which will greatly decrease the photocatalytic activity and hinder the further development of the materials; (2) Although the relatively narrow band gap of rutile extends the absorption of the mixed-phase TiO_2 to some of the visible light range, it is still insufficient to utilize all visible light from sun light, limiting its application in photocatalysis; (3) Although researchers all agree that the mixed-phase crystal structure enables the effective separation of photogenerated electrons and holes, thereby enhancing the photocatalytic activity, there is much controversy about the mobility direction of photo-generated electrons and holes. Further research is essential to confirm the mechanism of the enhanced photocatalytic activity of mixed-phase TiO_2 . Therefore, developing mixed-phase TiO_2 nanomaterials and combining other advantageous structures (such as hierarchical structure), enhancing the visible light utilization, and exploring the mechanism by newly developed characterization techniques will continue to be the challenges and hot research topics in this area.

Although mixed-phase crystal TiO_2 has been studied for decades, it is still a popular topic due to its excellent photocatalytic activity. Nowadays, with increasingly serious energy and environmental problems, the applications of mixed-phase TiO_2 is anticipated to attract considerable attention.

References

1. A. Fujishima, K. Honda, *Nature* **238**, 37–38 (1972)
2. J.H. Carey, J. Lawrence, T. Hm, *Bull. Environ. Contam. Toxicol.* **16**, 697–701 (1977)
3. S.N. Habisreutinger, S.L. Mende, J.K. Stolarczyk, *Angewandte Chemie* **52**, 7372–7408 (2013)
4. H. Ke, K.C.D. Robson, P.G. Johansson, C.P. Berlinguette, G.J. Meyer, *J. Am. Chem. Soc.* **134**, 8352–8355 (2012)
5. Q. Guo, C. Xu, Z. Ren, W. Yang, Z. Ma, D. Dai, H. Fan, T.K. Minton, X. Yang, *J. Am. Chem. Soc.* **134**, 2827–2830 (2012)
6. B. Tian, F. Chen, J. Zhang, M. Anpo, *J. Colloid Interface Sci.* **303**, 142–148 (2006)
7. M. Tomkiewicz, G. Dagan, Z. Zhu, *Res. Chem. Intermed.* **20**, 701–710 (1994)
8. S. Zhu, G. Xie, X. Yang, Z. Cui, *Mater. Res. Bull.* **48**, 1961–1966 (2013)
9. T. Beuvier, M. Richard-Plouet, G.M. Mancini-Le, T. Brousse, O. Crosnier, L. Brohan, *Inorg. Chem.* **49**, 8457–8464 (2010)
10. X. Xin, M. Scheiner, M. Ye, Z. Lin, *Langmuir* **27**, 14594–14598 (2011)
11. E. Hosono, S. Fujihara, H. Imai, I. Honma, I. Masaki, H. Zhou, *ACS Nano* **1**, 273–278 (2007)
12. A.K. Sinha, S. Jana, S. Pande, S. Saha, A. Pal, T. Pal, S. Sarkar, M. Pradhan, M. Basu, *Cryst. Eng. Comm.* **7**, 1210–1212 (2009)
13. Q.Q. Cheng, Y. Cao, L. Yang, P.P. Zhang, K. Wang, H.J. Wang, *Mater. Res. Bull.* **46**, 372–377 (2011)
14. Y. Wang, L. Zhang, K. Deng, X. Chen, Z. Zou, *J. Phys. Chem. C* **111**, 2709–2714 (2007)
15. J. Wei, J. Yao, X. Zhang, W. Zhu, H. Wang, M.J. Rhodes, *Mater. Lett.* **61**, 4610–4613 (2007)
16. Y. Hu, *Angew. Chem. Int. Ed.* **51**, 12410–12412 (2012)
17. J.K. Oh, J.K. Lee, H.S. Kim, S.B. Han, K.W. Park, *Chem. Mater.* **22**, 1114–1118 (2009)
18. M. Ye, H.Y. Liu, C. Lin, Z. Lin, *Small* **9**, 312–321 (2013)
19. S. Peng, S. Ding, J. Yuan, W.L. Xiong, D.H. Kim, *ACS Nano* **5**, 7617–7626 (2011)
20. E. Lioz, G. Peng, X. Zhaosheng, P. Qin, C. Aravind Kumar, L. Bin, M.K. Nazeeruddin, G.T. Michael, *J. Am. Chem. Soc.* (2012)
21. W. Guo, C. Xu, X. Wang, S. Wang, C. Pan, C. Lin, Z. Wang, *J. Am. Chem. Soc.* **134**, 4437–4441 (2012)
22. Y.Q. Wang, G. Lin, Y.G. Guo, L. Hong, X.Q. He, S. Tsukimoto, Y. Ikuhara, L.J. Wan, *J. Am. Chem. Soc.* **134**, 7874–7879 (2012)
23. S.S.K. Lee, P. Schmuki, *J. Am. Chem. Soc.* **134**, 11316–11318 (2012)
24. W.N. Wang, W.J. An, B. Ramalingam, S. Mukherjee, D.M. Niedzwiedzki, S. Gangopadhyay, P. Biswas, *J. Am. Chem. Soc.* **134**, 11276–11281 (2012)
25. S.A.K. Leghari, S. Sajjad, F. Chen, J. Zhang, *Chem. Eng. J.* **166**, 906–915 (2011)
26. F. Chen, W. Zou, W. Qu, J. Zhang, *Catal. Commun.* **10**, 1510–1513 (2009)
27. M. Xing, J. Zhang, F. Chen, *Appl. Catal. B* **89**, 563–569 (2009)
28. T.H. Kim, C. Gómez-Solís, E. Moctezuma, S.W. Lee, *Res. Chem. Intermed.* **40**, 1595–1605 (2014)
29. F. Zuo, L. Wang, T. Wu, Z. Zhang, D. Borchardt, P. Feng, *J. Am. Chem. Soc.* **132**, 11856–11857 (2010)
30. J.K. Yong, H.L. Mi, H.J. Kim, G. Lim, Y.S. Choi, N.G. Park, K. Kim, I.L. Wan, *Adv. Mater.* **21**, 3668–3673 (2009)
31. S. Liu, Q. Li, C. Hou, X. Feng, Z. Guan, *J. Alloy. Compd.* **575**, 128–136 (2013)
32. M. Cai, X. Pan, W. Liu, J. Sheng, X. Fang, C. Zhang, Z. Huo, H. Tian, S. Xiao, S. Dai, *J. Mater. Chem. A* **15**, 4885–4892 (2013)
33. M.S. Liang, C.C. Khaw, C.C. Liu, S.P. Chin, J. Wang, H. Li, *Ceram. Int.* **39**, 1519–1523 (2013)

34. H. Kim, H.H. Yun, G. Cho, D. Kim, N. Lim, M. Pyo, *Electrochim. Acta* **56**, 9476–9481 (2011)
35. C. Zhang, Y. Huang, S. Chen, H. Tian, L.E. Mo, L. Hu, Z. Huo, F. Kong, Y. Ma, S. Dai, *J. Phys. Chem. C* **116**, 19807–19813 (2012)
36. E.G. Bae, H. Kim, Y.H. Hwang, K.S. Sohn, M. Pyo, *J. Mater. Chem* **22**, 551–556 (2011)
37. A.A. Ashkarran, M. Ghavamipour, H. Hamidinezhad, H. Haddadi, *Res. Chem. Intermed.* (2014)
38. M. Haruta, B.S. Uphade, S. Tsubota, A. Miyamoto, *Res. Chem. Intermed.* **24**, 329–336 (1998)
39. B. Tian, J. Zhang, T. Tong, F. Chen, *Appl. Catal. B* **79**, 394–401 (2008)
40. Y. Wu, H. Liu, J. Zhang, F. Chen, *J. Phys. Chem. C* **113**, 14689–14695 (2009)
41. W. Wang, J. Zhang, F. Chen, M. Anpo, D. He, *Res. Chem. Intermed.* **36**, 163–172 (2010)
42. W. Yan, C. Feng, Z. Min, J. Yang, Z. Zhang, *Appl. Catal. B* **100**, 84–90 (2010)
43. C. Feng, Y. Wang, J. Zhang, L. Yu, D. Li, J. Yang, Z. Zhang, *Appl. Catal. B Environ.* **113–114**, 61–71 (2012)
44. A. Charanpahari, S.S. Umare, S.P. Gokhale, V. Sudarsan, B. Sreedhar, R. Sasikala, *Appl. Catal. A* **443**, 96–102 (2012)
45. B. Tian, C. Li, F. Gu, H. Jiang, *Catal. Commun.* **10**, 925–929 (2009)
46. P. Zhang, C. Shao, X. Li, M. Zhang, X. Zhang, Y. Sun, Y. Liu, *J. Hazard. Mater.* **237–238**, 331–338 (2012)
47. Z.H. Zhang, Y. Yuan, L.H. Liang, Y.X. Cheng, G.Y. Shi, L.T. Jin, *J. Hazard. Mater.* **158**, 517–522 (2008)
48. G. Chattopadhyaya, D.G. Macdonald, N.N. Bakhshi, J.S.S. Mohammadzadeh, A.K. Dalai, *Catal. Lett.* **108**, 1–5 (2006)
49. M. Yang, Y. Men, S. Li, G. Chen, *Appl. Catal. A* **433**, 26–34 (2012)
50. S. Ren, R. Bechstein, L. S03, R.T. Vang, M. Sillassen, B. Esbjörnsson, A. Palmqvist, F. Besenbacher, *J. Phys. Chem. C* **115**, 24287–24292 (2011)
51. D.C. Hurum, A.G. Agrios, K.A. Gray, T. Rajh, M.C. Thurnauer, *J. Phys. Chem. B* **107**, 4545–4549 (2003)
52. R. Scotti, I.R. Bellobono, C. Canevali, C. Cannas, M. Catti, M. D’Arienzo, A. Musinu, S. Polizzi, M. Sommariva, A. Testino, *Chem. Mater.* **20**, 4051–4061 (2008)
53. V. Puddu, H. Choi, D.D. Dionysiou, G.L. Puma, *Appl. Catal. B* **94**, 211–218 (2010)
54. R. Zheng, X. Meng, F. Tang, *Appl. Surf. Sci.* **255**, 5989–5994 (2009)
55. Y. Jiao, F. Chen, B. Zhao, H. Yang, J. Zhang, *Colloids Surf., A* **402**, 66–71 (2012)
56. A.D. Paola, G. Cufalo, M. Addamo, M. Bellardita, R. Camprostrini, M. Ischia, R. Ceccato, L. Palmisano, *Colloids Surf., A* **317**, 366–376 (2008)
57. W. Li, C. Liu, Y. Zhou, Y. Bai, X. Feng, Z. Yang, L. Lu, X. Lu, K.Y. Chan, *J. Phys. Chem. C* **112**, 20539–20545 (2008)
58. D. Hanaor, C. Sorrell, *J. Mater. Sci.* **46**, 855–874 (2011)
59. J.A. Gamboa, D.M. Pasquevich, *J. Am. Ceram. Soc.* **75**, 2934–2938 (1992)
60. X.Z. Ding, Y.Z. He, *J. Mater. Sci. Lett.* **15**, 320–322 (1996)
61. J. Muscat, V. Swamy, N.M. Harrison, *Phys. Rev. B* **65**, 392–397 (2002)
62. T. Arlt, M. Bermejo, M.A. Blanco, L. Gerward, J.Z. Jiang, J. Staun Olsen, J.M. Recio, *Phys. Rev. B* **61**, 14414–14419 (2000)
63. R. Ren, Z. Yang, L.L. Shaw, *J. Mater. Sci.* **35**, 6015–6026 (2000)
64. N.A. Dubrovinskaya, L.S. Dubrovinsky, R. Ahuja, V.B. Prokopenko, V. Dmitriev, H.P. Weber, J.M. Osorio-Guillen, B. Johansson, *Phys. Rev. Lett.* **87**, 455–475 (2001)
65. U. Diebold, *Surf. Sci. Rep.* **48**, 53–229 (2003)
66. M. Yan, F. Chen, J. Zhang, M. Anpo, *J. Phys. Chem. B*, **109** (2005)
67. Y. Li, T.J. White, S.H. Lim, *J. Solid State Chem.* **177**, 1372–1381 (2004)
68. A.R. Khataee, M.B. Kasiri, *J. Mol. Catal. A: Chem.* **328**, 8–26 (2010)
69. N. Daneshvar, M.H. Rasoulifard, A.R. Khataee, F. Hosseinzadeh, *J. Hazard. Mater.* **143**, 95–101 (2007)
70. M.P. Moret, R. Zallen, D.P. Vijay, S.B. Desu, *Thin Solid Films* **366**, 8–10(13) (2000)

71. R.D. Shannon, J.A. Pask, *J. Am. Ceram. Soc.* **48**, 391–398 (1965)
72. M. Batzill, E.H. Morales, U. Diebold, *Phys. Rev. Lett.* **96** (2006)
73. G. Oskam, A. Nellore, R.L. Penn, P.C. Searson, *J. Phys. Chem. B* **107**, 1734–1738 (2003)
74. H. Zhang, J.F. Banfield, *J. Mater. Chem.* **9**, 2073–2076 (1998)
75. J.G. Li, T. Ishigaki, X. Sun, *J. Phys. Chem. C* **111**, 4969–4976 (2007)
76. R.I. Bickley, T. Gonzalez-Carreno, J.S. Lees, L. Palmisano, R.J.D. Tilley, *J. Solid State Chem.* **92**, 178–190 (1991)
77. S.-K. Lee, P.K.J. Robertson, A. Mills, D. McStay, *J. Photochem. Photobiol. A* **122**, 69–71 (1999)
78. S.J. Tsai, S. Cheng, *Catal. Today* **33**, 227–237 (1997)
79. T. Ohno, K. Sarukawa, M. Matsumura, *J. Phys. Chem. B* **105**, 2417–2420 (2001)
80. T. Ozawa, M. Iwasaki, H. Tada, T. Akita, K. Tanaka, S. Ito, *J. Colloid Interface Sci.* **281**, 510–513 (2005)
81. H. Xu, L. Zhang, *J. Phys. Chem. C* **113**, 1785–1790 (2009)
82. R.R. Bacsá, J. Kiwi, *Appl. Catal. B* **16**, 19–29 (1998)
83. K.Y. Jung, S.B. Park, H.D. Jang, *Catal. Commun.* **5**, 491–497 (2004)
84. S. Leghari, S. Sajjad, F. Chen, J. Zhang, *Chem. Eng. J.* **3**, 906–915 (2011)
85. J. Zhu, W. Zheng, B. He, J. Zhang, M. Anpo, *J. Mol. Catal. A: Chem.* **216**, 35–43 (2004)
86. Y. Wu, M. Xing, J. Zhang, *J. Hazard. Mater.* **192**, 368–373 (2011)
87. J. Ng, X. Wang, D.D. Sun, *Appl. Catal. B* **110**, 260–272 (2011)
88. J. Ovenstone, K. Yanagisawa, *Chem. Mater.* **11**, 2770–2774 (1999)
89. G. Li, S. Ciston, Z.V. Saponjic, L. Chen, N.M. Dimitrijevic, T. Rajh, K.A. Gray, *J. Catal.* **253**, 105–110 (2008)
90. M. Fehse, F. Fischer, C. Tessier, L. Stievano, L. Monconduit, *J. Power Sources* **231**, 23–28 (2013)
91. Y. Zhang, J. Chen, X. Li, *Catal. Lett.* **139**, 129–133(125) 2010
92. Q. Tay, X. Liu, Y. Tang, Z. Jiang, T.C. Sum, Z. Chen, *J. Phys. Chem. C* **117**, 14973–14982 (2013)
93. H. Zhang, J.F. Banfield, *J. Phys. Chem. B* **104**, 3481–3487 (2000)
94. X. Shen, B. Tian, J. Zhang, *Catal. Today* **201**, 151–158 (2013)
95. B. Zhao, F. Chen, Q. Huang, J. Zhang, *Chem. Commun.* **34**, 5115–5117 (2009)
96. L.M. Zhao, Z.J. Zhang, S.Y. Zhang, P. Cui, W. Shi, B. Zhao, P. Cheng, D.Z. Liao, S.P. Yan, *Cryst. Eng. Comm.* **3**, 907–913 (2011)
97. B. Zhao, F. Chen, Y. Jiao, J. Zhang, *J. Mater. Chem.* **20**, 7990–7997 (2010)
98. X. Shen, J. Zhang, B. Tian, M. Anpo, *J. Mater. Sci.* **47**, 5743–5751 (2012)
99. H. Cheng, J. Ma, Z. Zhao, L. Qi, *Chem. Mater.* **7**, 663–671 (1995)
100. G. Li, K.A. Gray, *Chem. Mater.* **19**, 1143–1146 (2007)
101. S. Lei, W. Duan, *J. Environ. Sci.* **20**, 1263–1267 (2008)
102. M. Wu, J. Long, A. Huang, Y. Luo, S. Feng, R. Xu, *Langmuir* **15**, 8822–8825 (1999)
103. X. Shen, J. Zhang, B. Tian, *J. Hazard. Mater.* **192**, 651–657 (2011)
104. A. Zachariah, K.V. Baiju, S. Shukla, K.S. Deepa, J. James, K.G.K. Warriar, *J. Phys. Chem. C* **112**, 11345–11356 (2008)
105. A. Bojinova, R. Kralchevska, I. Poullos, C. Dushkin, *Mater. Chem. Phys.* **106**, 187–192 (2007)
106. Z. Liu, X. Zhang, S. Nishimoto, J. Ming, D.A. Tryk, T. Murakami, A. Fujishima, *Langmuir* **23**, 10916–10919 (2007)
107. R.G. Nair, S. Paul, S.K. Samdarshi, *Sol. Energy Mater. Sol. Cells* **95**, 1901–1907 (2011)
108. P.I. Gouma, M.J. Mills, *J. Am. Ceram. Soc.* **84**, 619–622 (2001)
109. Z. Jing, M. Li, Z. Feng, J. Chen, C. Li, *J. Phys. Chem. B* **110**, 927–935 (2006)
110. Y.C. Hsu, H.C. Lin, C.H. Chen, Y.T. Liao, C.M. Yang, *J. Solid State Chem.* **183**, 1917–1924 (2010)
111. M. Ni, M.K.H. Leung, D.Y.C. Leung, K. Sumathy, *Renew. Sustain. Energy Rev.* **11**, 401–425 (2007)

112. I.R. Ocaña, A. Beltram, J.J.D. Jaén, G. Adami, T. Montini, P. Fornasiero, *Inorganica Chimica Acta* (2015)
113. Q. Xu, Y. Ma, J. Zhang, X. Wang, Z. Feng, C. Li, *J. Catal.* **278**, 329–335 (2011)
114. Y.K. Kho, A. Iwase, W.Y. Teoh, L. Mölder, A. Kudo, R. Amal, *J. Phys. Chem. C*, **114**, 2821–2829 (2010)
115. F. Xu, W. Xiao, B. Cheng, J. Yu, *Int. J. Hydrogen Energy* **39**, 15394–15402 (2014)
116. O. Rosseler, M.V. Shankar, M.K.-L. Du, L. Schmidlin, N. Keller, V. Keller, *J. Catal.* **269**, 179–190 (2010)
117. G. Marci, M. Addamo, V. Augugliaro, S. Coluccia, E. García-López, V. Loddo, G. Martra, L. Palmisano, M. Schiavello, *J. Photochem. Photobiology A: Chem.* **160**, 105–114 (2003)
118. L. Liu, H. Zhao, J.M. Andino, Y. Li, *ACS Catal.* **2**, 1817–1828 (2012)
119. H. Zhao, L. Liu, J.M. Andino, Y. Li, *J. Mater. Chem. A* **1**, 8209–8216 (2013)
120. S.N. Frank, A.J. Bard, *J. Am. Chem. Soc.* **99**, 303–304 (2002)
121. H. Xu, G. Li, G. Zhu, K. Zhu, S. Jin, *Catal. Commun.* **62**, 52–56 (2015)
122. Z. Luo, A. S. Poyraz, C. H. Kuo, R. Miao, Y. Meng, S. Y. Chen, T. Jiang, C. Wenos, S. L. Suib, Z. Luo, *Chem. Mater.* **27** (2014)
123. B. Zhao, Y. Jiao, F. Chen, H. Yang, J. Zhang, *J. Mol. Catal. A: Chem.* **348**, 114–119 (2011)
124. Y. Liao, W. Que, Q. Jia, Y. He, J. Zhang, P. Zhong, *J. Mater. Chem.* **22**, 7937–7944 (2012)
125. E. Grabowska, J. Reszczyńska, A. Zaleska, *Water Res.* **46**, 5453–5471 (2012)
126. Z. Ding, G. Lu, P. Greenfield, *J. Phys. Chem. B* **104**, 4815–4820 (2000)
127. B. Tian, C. Li, J. Zhang, *Chem. Eng. J.* **191**, 402–409 (2012)
128. R.I. Bickley, T. Gonzalez-Carreno, J.S. Lees, L. Palmisano, R.J.D. Tilley, *J. Solid State Chem. (United States)*, **92:1**, 178–190 (1991)
129. N.A. Deskins, S. Kerisit, K.M. Rosso, M. Dupuis, *J. Phys. Chem. C* **111**, 9290–9298 (2007)
130. P. Deák, B. Aradi, T. Frauenheim, *J. Phys. Chem. C* **115**, 3443–3446 (2011)
131. D.O. Scanlon, C.W. Dunnill, J. Buckeridge, S.A. Shevlin, A.J. Logsdail, S.M. Woodley, C.R.A. Catlow, M.J. Powell, R.G. Palgrave, I.P. Parkin, *Nat. Mater.* **12**, 798–801 (2013)
132. A.K. Datye, G. Riegel, J.R. Bolton, M. Huang, M.R. Prairie, *J. Solid State Chem.* **115**, 236–239 (1995)
133. Z. Zhang, C.C. Wang, R. Zakaria, J.Y. Ying, *J. Phys. Chem. B* **102**, 10871–10878 (1998)
134. T. Ohno, K. Sarukawa, K. Tokieda, M. Matsumura, *J. Catal.* **203**, 82–86 (2001)
135. T. Kawahara, Y. Konishi, H. Tada, N. Tohge, J. Nishii, S. Ito, *Angew. Chem. Int. Ed. Engl.* **41**, 2811–2813 (2002)
136. G. Li, K.A. Gray, *Chem. Phys.* **339**, 173–187 (2007)
137. B. Sun, A.V. Vorontsov, P.G. Smirniotis, *Langmuir* **19**, 3151–3156 (2003)
138. S. Bo, P.G. Smirniotis, *Catal. Today* **88**, 49–59 (2003)
139. B. Liu, L. Peng, *J. Alloy. Compd.* **571**, 145–152 (2013)
140. G. Li, L. Chen, M.E. Graham, K.A. Gray, *J. Mol. Catal. A: Chem.* **275**, 30–35 (2007)
141. C.-Y. Wang, R. Pagel, J.K. Dohrmann, D.W. Bahnemann, C. R. Chim. **9**, 761–773 (2006)

Chapter 18

Electrochemical Impedance Study on Poly(Alkylenedioxy)Thiophene Nanostructures: Solvent and Potential Effect

A. Sezai Sarac and Ash Gencturk

Abstract Conducting polymers can be doped and dedoped rapidly to a high charge density, hence, they can be applied as active materials for supercapacitors. Higher energy densities can be achieved because charging occurs through very thin thicknesses from the nano to microscale range. Taking into account the costs and compatibility of the materials, the modification of carbon fiber by electrocoating of poly(3,4-alkylenedioxythiophene)s for microsupercapacitor applications seems to be a very attractive method. [3,4-(2,2-dimethylpropylenedioxy)thiophene] was electrodeposited cyclic voltammetrically onto the carbon fiber micro electrode. An electrochemical impedance spectroscopic study was performed at applied potential, in different electrolytes and solvents and evaluated with our previous findings by reviewing.

18.1 Introduction

Owing to its good electrochemical properties and its low oxidation potential, its high conductivity [1], unusual thermal stability in the oxidized state, poly(3,4 ethylene-dioxythiophene) (PEDOT) has been widely investigated [1, 2]. The electron donating oxygen atoms in 3- and 4- position not only reduce the oxidation potential of the aromatic ring but also prevent α,β -coupling during the polymerization. The ethylene bridges minimize steric distortion effects resulting in a high stereoregularity of the polymer chain so that good-conjugation is guaranteed. Although in special fields of application PEDOT-type materials with improved properties were needed.

A. Sezai Sarac (✉) · A. Gencturk
Department of Chemistry, Polymer Science and Technology,
Istanbul Technical University, Maslak, 34469 Istanbul, Turkey
e-mail: sarac@itu.edu.tr
URL: <http://web.itu.edu.tr/~sarac/>

The conductivity and stability of PEDOT makes it useful as a conducting electrode in electrochemical supercapacitors [3].

As a consequence different types of poly(3,4-alkylenedioxythiophene)s structurally related to PEDOT have been developed [4–6]. In comparison with PEDOT the structure of 3,4-(2,2-dimethylpropylenedioxy)thiophene (PProDOT-Me₂) contains propylene bridges with quaternary central carbon and tetrahedral symmetry which should allow to stabilize [7] high levels of electrical charge. Poly [3,4-(2,2-dimethylpropylenedioxy)thiophene] (PProDOT-Me₂) is a very promising candidate for use in electrochromic devices switching between slightly different colors than PEDOT [8] with higher coloration efficiency [9], higher electrochromic contrast ratio [10] and very fast switching times. Using PProDOT-Me₂ several prototypes of devices [11–17] have been fabricated.

The tetrahedral substitution pattern of PProDOT-Me₂ causes the alkyl groups to be positioned above and below the plane of the Π conjugated chain, can allow high doping levels and inhibit π stacking [7], and the quaternary central carbon of the propylene bridge of 3,4-(2,2-dimethylpropylenedioxy)thiophene should cause structural conditions that can favour high charge capacitance in well ordered polymer films. Electrochromic properties of PProDOT-Me₂ film was studied on the potentiostatically deposited ITO coated glass slides previously. PProDOT-Me₂ shows good contrast in the visible region. This higher contrast is due to its higher percentage of transmittance in the oxidised form. At an applied potential of -600 mV, the polymer is in its fully neutral form; the band gap of the polymer E_g was calculated as 1.85 eV from onset of π to π^* transition [18].

Carbon based materials are well known for the electrochemical applications. As an electrode material, carbon has several advantages compared to other electrode materials, i.e. good polarisation of the electrode, high surface area, processability, low cost, accessibility, thermal and chemical stability. Porous carbon, offers large surface area and good polarisation due to its porosity, which makes it one of the advantages electrode materials for supercapacitors applications [19–22].

Carbon fibers have unique properties such as high strength, high modulus and low density. Recently, the surface properties of carbon fibers (CF) and modification of carbon fibre surfaces by electropolymerising with thiophene derivatives and their copolymers onto CF have been investigated in detail [23–25]. Ethylcarbazole-bis-EDOT-ethylcarbazole comonomer was electrochemically coated onto carbon fibre microelectrodes (CFME) and stable films were obtained at high potentials [26].

Micron size carbon fiber electrodes (CFMEs) show superior performance in potentiodynamic studies due to their size and cylindrical shape. They can be used once (disposable) a new surface area will be present at each measurement compared to conventional electrodes.

Generally nano- and micro-scale electrocoatings of conjugated polymers on carbon fiber easily allow the characterization of the deposited films by electrochemical, spectroscopic, and morphological techniques [27–31]. The comparison of

different polymers coated under identical electrochemical conditions gives reliable information [32–35], for the evaluation of the materials in terms of application. The high stability and electroactivity was obtained for PEDOT on CFME which is important for the view of practical applications.

For high capacitance the carbon electrodes should display a high specific surface area that can easily be accessed by the electrolyte; carbon porosity and surface chemistry are thus of primary importance. The carbons should display a high pore volume and pores at least wider than the size of the counterions. Although pore walls should be thicker than the screening length of the electric field in the solid part of the double layer without limiting the charge accommodation on the two sides of the carbon walls [36, 37]. Functional groups are also important because they affect the electronic properties of the carbon itself and its surface polarizability and can attract–repel the ions with different effects on carbon wettability.

The electrochemical impedance spectroscopy (EIS) is one of the most effective tools to obtain information about double layer capacitance, resistance, determination of the rate of the charge transfer and charge transport processes, solution resistance, etc. [19, 38]. It was used to explain behaviour of polymer coated electrodes by established theories using two models which are known as uniform and porous medium.

Electrochemical capacitors (supercapacitors, ultracapacitors) have been intensively investigated since they are expected to be useful for high-power, in particular, pulse-power applications. Several materials for this application have been reviewed [39, 40]. Supercapacitors [41] can be classified into two types: double layer supercapacitors, which store the energy within the electrochemical double layer at the electrode/electrolyte interface and redox supercapacitors. For redox supercapacitors, among several types of electrode materials, the use of conducting polymers has received much attention due to their use in both aqueous and organic electrolytes and their wide potential working range [42].

Supercapacitors are becoming increasingly important as an intermittent power supply as well as for load levelling and energy storage in fields ranging from distributed energy grids to domestic and transport applications [43–45].

Modified carbon fiber electrodes can be good candidates for storing the charges by two mechanisms: double layer effect, and pseudocapacitance, which can store more energy than double layer effects.

The paper presents an electrochemical deposition, characterization, impedance and morphologic study of PProDOT-Me₂ which is electrocoated on carbon fiber microelectrode. The electrochemical characteristics of PProDOT-Me₂ on carbon fiber microelectrode (CFME) in an electrolyte solution of tetrabutylammonium hexafluorophosphate (Bu₄NPF₆) in acetonitrile (ACN) shows that it can be used as an active electrode material for electrochemical supercapacitor applications.

18.2 Experimental Details

18.2.1 Chemicals

Acetonitrile (ACN), dimethylformamide (DMF) and propylenecarbonate (PC) were used as received from Riedel-de Haen Chemical without further purification. Tetrabutylammonium hexafluorophosphate (Bu_4NPF_6), tetrabutylammonium tetrafluoroborate (Bu_4NBF_4), tetraethylammonium hexafluorophosphate (Et_4NPF_6), tetraethylammonium tetrafluoroborate (Et_4NBF_4) were used from Fluka (%98) Chemicals. SGL SIGRAFIL C 320 B (A high strength and high modulus of elasticity coupled with high electrical conductivity carbon fibers) (SGL Carbon Group) containing single filaments in a roving were used as working electrodes. ProDOT-Me₂ monomer was prepared in three synthetic steps from commercially available starting materials using Mitsunobu conditions [46] followed by alkaline hydrolysis [47], and decarboxylation [48] of the resulting precursor.

18.2.2 Preparation of Carbon Fiber Microelectrode (CFME)

A carbon fiber was used as working electrode (3 cm of the CFME with an average diameter of $\sim 7 \mu\text{m}$) attached to a copper wire. About 1 cm of the CFMEs were dipped into the solution to keep the electrode area constant and the rest of the electrode was covered with the Teflon type. The CFMEs were cleaned with acetone and then dried up with an air-dryer before the experiments.

18.2.3 Apparatus and Procedure

Cyclic voltammograms (CV) of the polymer coated CFMEs and electrochemical impedance spectroscopy (with supplied Power Sine software package) were performed on a Princeton Applied Research (PAR) Parstat 2263 potentiostat, which is a self-contained unit that combines potentiostatic circuitry with phase-sensitive detection. The potentiostat was connected to a Faraday cage (BAS Cell Stand C3).

Electropolymerisation was performed using a three electrode system with a CFME as a working electrode, a platinum (Pt) wire as a counter and a silver (Ag) wire as a pseudoreference electrode. The cell has a volume of 5 mL solution, arranged in a way at a distance of, 1 cm from each other. The pseudoreference electrode was calibrated externally using a 5 mM solution of ferrocene/ferrocenium (Fc/Fc^+) redox couple in the electrolyte and potentials were reported against the Ag electrode.

After the electropolymerization process, coated CFME was washed with ACN and EIS measurements were performed in monomer free electrolyte solution using a three electrode system. Electrochemical polymerizations (electrocoatings) were performed in acetonitrile (ACN) containing 0.1 M Bu_4NPF_6 at 100 mV/s and electrochemical impedance measurements were performed in DMF, ACN, and PC containing 0.1 M electrolyte at applied potential. An electrochemical impedance software (PowerSine) was used to carry out impedance measurements by scanning in the 100 kHz–10 mHz frequency range with an applied AC signal amplitude of 10 mV and DC potentials were applied before EIS measurements. The impedance spectra were analysed using an AC impedance data analysis software (Z SimpWin V3.10).

Thin films of homopolymers, electrocoated onto carbon fiber surface were analyzed by FT-IR reflectance spectrophotometry (Perkin-Elmer, Spectrum One, with Universal ATR attachment with a diamond and ZnSe crystal C70951).

The morphological features of the coated carbon fibre electrodes were observed using a Hitachi S-2700 scanning electron microscopy (Nissei Sangyo GmbH, Rathingen, Germany), which was connected to an energy dispersive X-ray (EDX) microanalyser (Kevex type delta V, Foster City, CA, USA). The excitation energy was 10 keV at a beam current of 0.5 nA.

18.3 Results and Discussion

18.3.1 Electropolymerization of ProDOT-Me₂ on CFME

Oxidative electropolymerization of 5 mM ProDOT-Me₂ was performed potentiodynamically between 0.0 and 1.6 V versus Ag in ACN with 0.1 M Bu_4NPF_6 as the electrolyte. Polymer film was electrodeposited by CV on a CFME (area 0.02 cm²) at 100 mV s⁻¹, it was scanned in monomer free solution between 20 and 400 mV s⁻¹ in 0.1 M $\text{Bu}_4\text{NPF}_6/\text{ACN}$. The typical CV of the polymeric film exhibits a well defined and reversible redox process as shown in Fig. 18.1.

The processes involved thus give rise in CV to almost mirror image voltammograms that are the characteristics of capacitor discharge and recharge behavior, and also to response currents that increase with scan rate. The current of anodic peak potential, i_{pa} depends linearly on the scan rate (up to 40th cycle). The segments of the film consisting of the longer (and porous in some cases) polymer chains (filled more with dopant) are more compact than the short chain segments. The incorporation of ions into the compact parts of the film during oxidation is slow therefore the process can be controlled by diffusion. Multisweep cyclic voltammogram of 5 mM ProDOT-Me₂ in 0.1 M $\text{Bu}_4\text{NPF}_6/\text{ACN}$ on CFME show an increasing current density with each cycle ($E_{\text{pa1}} = 0.373$ V, $E_{\text{pa2}} = 0.524$ V, $E_{\text{pc1}} = 0.291$ V, $E_{\text{pc2}} = 0.484$ V at 20th cycle), resulting in the formation of thin film of conducting polymer on CFME at 100 mV/s and 20 cycle. The cyclic

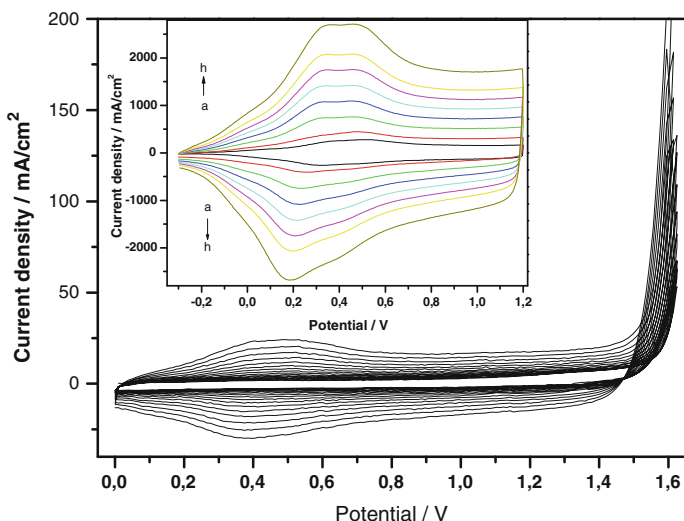


Fig. 18.1 Cyclic voltammetry of electrodeposition of monomer at 100 mV s^{-1} , 15 cycle, inset electrocoated CFME scanned in monomer free electrolyte at $20\text{--}400 \text{ mV s}^{-1}$

voltammogram of a polymer coated CFME in monomer free solution with different scan rates ($0.1 \text{ M Bu}_4\text{NPF}_6/\text{ACN}$) shows two oxidation peaks ($E_{\text{pa}1} = 0.360 \text{ V}$, $E_{\text{pa}2} = 0.483 \text{ V}$) and corresponding reduction peaks ($E_{\text{pc}1} = 0.184 \text{ V}$, $E_{\text{pc}2} = 0.389 \text{ V}$ (weak)) (inset Fig. 18.1). The current response of the film was directly proportional to the scan rate indicating a diffusion limited process, showing that the polymer film is electroactive from 0.0 to 1.6 V and sufficiently ensuring rapid film charging and discharging.

The capacitive behaviour of the CFME/film was first noticed from CV experiments in a monomer free solution (the inset in Fig. 18.1). The variation of specific capacitance C_{CV} calculated from CV with respect to potential and the scan rate of the CFME/ProDOT-Me₂/electrolyte system. Capacitance values from CV were calculated by applying the equation $C_{\text{CV}} = i/v \approx 2 \text{ mF}$ (where i : current density (A) v : scan rate (V/s)) and the current densities obtained from the oxidation and reduction peaks.

18.3.2 FTIR-ATR Characterisation of PProDOT-Me₂ Film on CFME

The electropolymerized thin film structure and doping with the respective anion of the supporting electrolyte was followed by FTIR-ATR. When the cycle number increased, the appearance of characteristic peaks of PProDOT-Me₂ was observed.

The peak intensities increased with increase in charge which confirms the film formation of PProDOT-Me₂ on CFME [49].

A pattern of five main bands (1470, 1287, 1155, doublet at 1017 and 826 cm⁻¹) were detected for a PProDOT(Me)₂ film coated on CFME during application of charge in the electrogrowth process (with 100 mV/s⁻¹ scan rate and 5–30 cycles). The band at 1470 cm⁻¹ (aromatic stretching of C=C bond) and a peak at 1287 cm⁻¹ (in plane deformation of C–H bond at β position) are known to be characteristic vibrational peaks of polythiophene. Vibrations at 1469–1470, 1295–1286 and 1160–1155 cm⁻¹ originate from the stretching of C–C and C=C bonds in the thiophene ring.

Further vibrations from the C–S bond in the thiophene ring can be seen at 833–825 cm⁻¹ assigned to C–S stretching. Vibrations at 1017–1048 cm⁻¹ are assigned to stretching in the alkylendioxy group [50].

18.3.3 Electrochemical Impedance Spectroscopy

Conducting polymer systems offer opportunities for storing electrochemical energy largely through their redox capacitance. The essential features of the electrochemical behavior of the conducting polymers that make them suitable as electrochemical capacitor materials, the continuous range of states of oxidation that arise with increasing electrode potential and the reversibility of the faradaic processes corresponding to charge withdrawal and reinjection.

The impedance is a complex quantity defined as $Z(\omega) = v(\omega)/I(\omega)$, where $v(\omega)$ is the transform of an alternating potential signal applied to the system and $I(\omega)$ is the transform of the alternating current signal that results. The impedance is a function of the frequency of the signal. It is most generally represented as $Z(\omega) = |Z|e^{i\theta}$. Its magnitude or modulus is $|Z|$, and its phase angle is θ . Both $|Z|$ and θ may also be functions of the frequency. As the imaginary contribution increases, the phase angle increases. For a capacitor, the impedance is purely imaginary, the phase angle is 90° and the current is out of phase with the voltage by 90°.

However, in a resistor, the impedance is real, the current and the voltage are in phase, and the phase angle is zero. A large phase angle indicates that the impedance is predominantly capacitive, while for small angles, it is resistive.

A working electrode/solution interface responds to various perturbations (i.e. potential steps and potential scans) which are usually of large amplitude, and they generally drive the working electrode far from equilibrium. Another approach is to perturb the cell by an alternating (usually sinusoidal) signal of small amplitude (nominally a few millivolts peak to peak, in our case 10 mV was applied) and observe the manner in which the system follows the perturbation at steady state.

An ideal double layer capacitance behavior of an electrode material in electrochemical supercapacitors is expressed in the form of a rectangular shape of the voltammetry characteristics, and Bode phase angles close to 90°. Polymer coated electrodes used in this study show an ideal capacitive line by the application of

electrochemical impedance spectroscopic measurements, indicating fast charge transfer at the carbon fiber/polymer and polymer/solution interfaces.

The low frequency capacitance values from impedance spectroscopy were obtained from the slope of a plot of the imaginary component Z_{IM} of the impedance at low frequencies versus the inverse of the reciprocal frequency f using the following equation [51]

$$C_{LF} = (2\pi f Z_{IM})^{-1} \quad (18.1)$$

18.3.4 Potential Effect on EIS of PProDOT-Me₂/CFME Coated Electrode

Electrochemical impedance measurements have been performed at different scan rates in the range of 0–400 mV/s and different scan number in the range of 5–40 cycle to find the optimum conditions of coating and also different applied potentials in ACN solution in the range of –0.1–1.3 V with a potential step of 0.1 V to cover polymerization range of the ProDOT-Me₂ monomer. The specific capacitance values of PProDOT-Me₂ film which were calculated from (Equation 18.1) using impedance data for frequencies of 10 and 16 mHz with applied DC potentials between 0.2 and 1.3 V in monomer free electrolyte solutions [18].

Low frequency capacitance increases with increase of diameter of polymer films (average values are calculated from the SEM pictures). Bode phase angle versus frequency data for PProDOT-Me₂/CFME prepared with two different charge densities showed high capacitance. Low frequency capacitance values of PProDOT-Me₂/CFME (C_{LF}) (for the scan rate of 100 mV/s during electrogrowth) were calculated for the scan numbers of 5, 20 and 40 cycles are 0.29, 1.39 and 12.05 mF/cm², respectively [18].

Although they are not exactly corresponding to C_{LF} , these results are comparable with the literature values obtained for PEDOT/Pt (in 0.1 M LiClO₄/PC) as 5 mF (surface area: 0.03 cm²) [52], and for PEDOT/Pt (in 0.1 M KCl/H₂O) as 2.11 mF (surface area: 0.07 cm²) [53].

In this study, electropolymerization was performed at 100 mV/s and 20 cycle Bu₄NPF₆/ACN and EIS measurements were performed at Bu₄NPF₆/PC, and Bu₄NPF₆/DMF electrolyte solution and also in Bu₄NPF₆, Bu₄NBF₄, NaClO₄, LiClO₄/ACN. Capacitive behaviour of PProDOT-Me₂ film in PC and DMF solutions were exhibited similar trend with ACN in electrochemical impedance measurement.

Figure 18.2 shows Nyquist plots, Magnitude of Z , and Bode Phase plots of coated CFME. Bode phase angle in which the frequency dependence of the system is clearer compared to Nyquist plots. We can separate the plot into three frequency regions namely low, medium and high frequency regions. Low frequency region starts with 10 mHz and finishes with 1 Hz, in this region (except 1.3 V) polymer

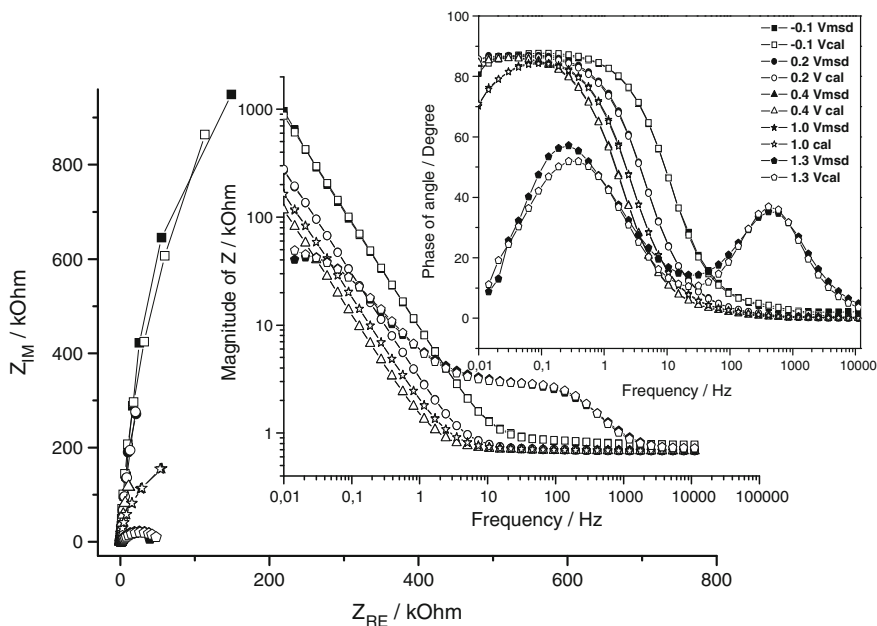


Fig. 18.2 Nyquist plots, (*middle inset* Magnitude of Z , *right inset* Bode Phase,) EIS measurements were performed in $\text{Bu}_4\text{NPF}_6/\text{DMF}$ at applied potentials (ProDOT-Me₂ film deposited at 100 mV/s, 20 cycle in $\text{Bu}_4\text{NPF}_6/\text{ACN}$)

films show a maximum peak at 100 mHz with a phase angle of (-87°). Between 1 and 100 Hz, phase angles sharply decrease. Two peaks were observed at applied potential at 1.3 V from the bode phase plot in $\text{Bu}_4\text{NPF}_6/\text{DMF}$, one of them at 100 mHz (-55°) and the other at 1000 Hz (-35°), on the other hand, in our previous paper only one peak was observed in $\text{Bu}_4\text{NPF}_6/\text{ACN}$. From the magnitude of Z plot, magnitude of impedance exhibits a large drop in the range of 0.01–10 Hz. From these observations we can conclude that the most capacitive behavior was obtained for the potential of 0.4 V.

18.3.5 Electrolyte and Solvent Effects on ProDOT-Me₂ Coated CFMEs

Figure 18.3 shows variation of low frequency capacitance values of PProDOT-Me₂ film in different electrolytes in ACN. CFME/PProDOT-Me₂ film has capacitance values higher in Bu_4NPF_6 compare to other electrolyte solutions. Cyclic voltammetric values have exhibited a peak at ~ 0.4 V which is the oxidation peak, similar trend was also observed for the case of capacitances obtained from EIS measurements.

Figure 18.4 shows variation of low frequency capacitance values of PProDOT-Me₂ film on CFME in different solutions with different polarity.

Fig. 18.3 Variation of low frequency capacitance (at 10 mHz) of the PProDOT-Me₂ film which was electrodeposited in 5 mM ProDOT-Me₂ at 100 mV/s, 20 cycle in 0.1 M Bu₄NPF₆/ACN (EIS was performed in Bu₄NPF₆, Et₄NPF₆, Bu₄NBF₄, Et₄NBF₄/ACN)

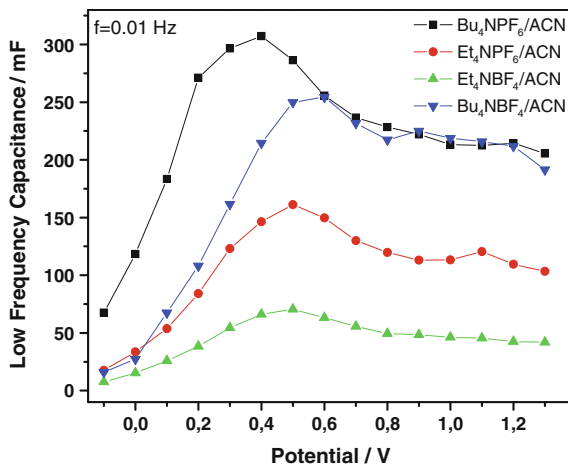
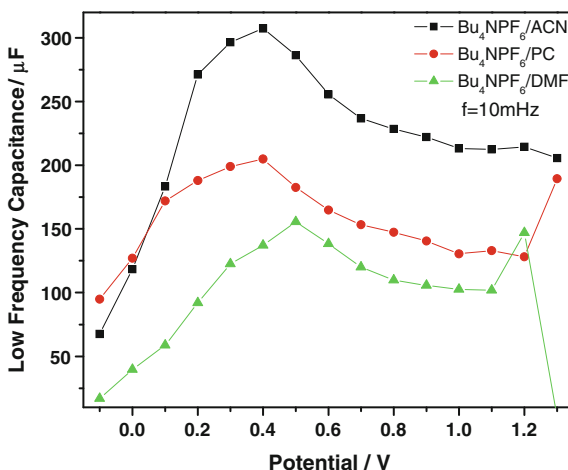


Fig. 18.4 Variation of low frequency capacitance (at 10 mHz) of the PProDOT-Me₂ film which was electrodeposited in 5 mM ProDOT-Me₂ at 100 mV/s, 20 cycle in 0.1 M Bu₄NPF₆/ACN (EIS was performed in Bu₄NPF₆/ACN, Bu₄NPF₆/PC, Bu₄NPF₆/DMF)



CFME/PProDOT-Me₂ film has capacitance values higher in ACN compared to PC and DMF. Shape of the capacitance curves at various potentials in ACN, PC and DMF exhibits similar fashion. The peak which is observed at around 0.4 V in ACN, PC and DMF solvents exhibits higher specific capacitances by corresponding the oxidation peak of the polymer (Fig. 18.1).

Figure 18.5 shows the dependence of polarity index and double layer capacitance of the modified electrode. Double layer capacitance (C_{DL}) and low frequency capacitance (C_{LF}) of the modified electrodes decrease, with the increasing of the polarity index (P') (ACN: 5.8, PC: 6.1, DMF: 6.4) [54]. High polarity result a negative effects on double layer capacitance.

Figure 18.6 shows the dependence of polarity index and double layer capacitance of the electrode and polarization resistance. Double layer capacitance (C_{DL}) of

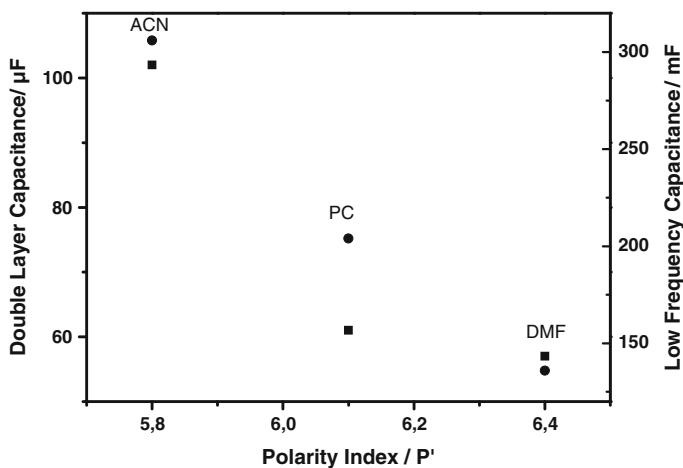


Fig. 18.5 Dependence of polarity index and double layer capacitance of the electrode and low frequency capacitance (*Filled circle* Low frequency capacitance)

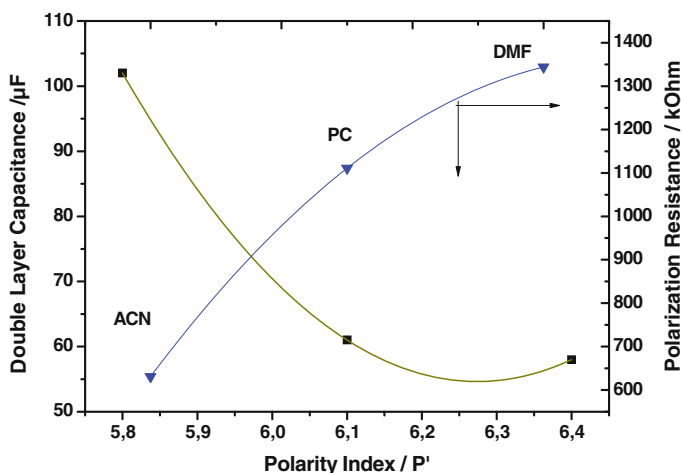


Fig. 18.6 Dependence of polarity index and double layer capacitance of the electrode and polarization resistance

the modified electrodes decrease and polarization resistance increases, with the increasing of the polarity index (P').

Potential and electrolyte effect on capacitance for other derivatives of Poly (alkylenedioxy)thiophenes like 3,4-(2,2-dibenzylpropylenedioxy)thiophene and 3,4-(2,2-dibutylpropylenedioxy)thiophene was described in detail in our previous papers [55, 56]. Bu_4NPF_6 caused the highest capacitance from the conductive salts investigation.

18.3.6 Electrical Equivalent Circuit Modeling

The electrochemical parameters of the CFME/PProDOT-Me₂/electrolyte system were evaluated using the ZSimpWin (version 3.10) software from Princeton Applied Research.

Two electrical equivalent circuits were proposed in the simulation of the impedance behaviour of the film from the experimentally obtained impedance data in previous paper for the condition of applied of the potential effect [18].

The model (Fig. 18.7) was built using series components; the first one is the bulk solution resistance of the polymer and the electrolyte R_s , the second one the parallel combination of the double layer capacitance C_{dl} and R_1 is the polarization resistance of the electrolyte. A series connection to R_1 was made up using Constant Phase Element (CPE) in parallel with R_2 and W (R_2 is the charge transfer and W is the Warburg impedance of the polymer). χ^2 is the function defined as the sum of the squares of the residuals. The chi squared (χ^2) was minimised below 10^{-4} .

Proposed equivalent circuit model was successfully applied to the experimental data which was performed in monomer free solution in different electrolytes and solvents to explain the electrical interfacial properties between the carbon fiber microelectrode, the polymer film and electrolyte in this potential region. Simulation results are given in Tables 18.1, 18.2 and 18.3 shows the fitting of the results of the

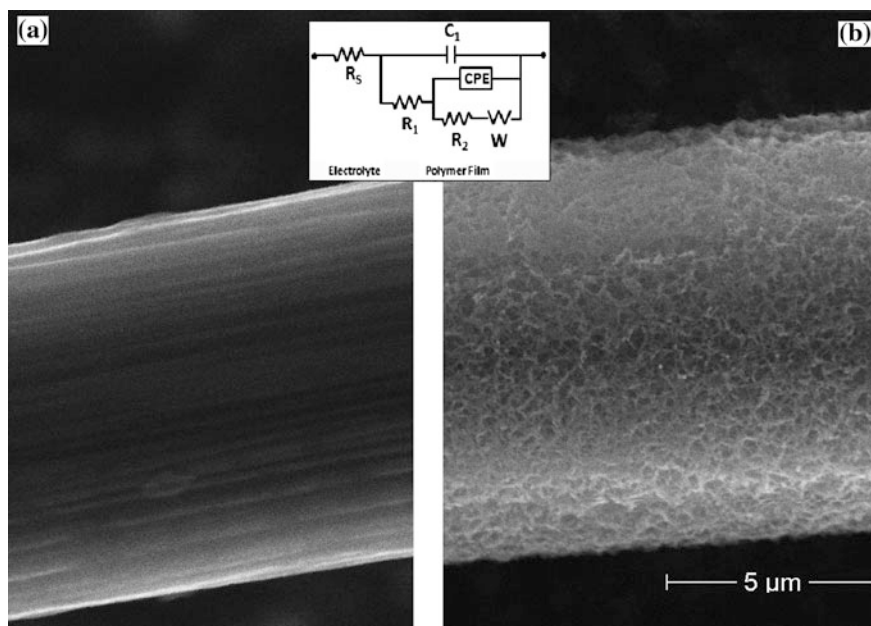


Fig. 18.7 SEM of CFME and ProDOT-(Me)₂/CFME **a** uncoated **b** coated at 100 mV/s and equivalent electrical circuit used in simulation

Table 18.1 Dependence of the parameters calculated for PProDOT-Me₂ film, at different applied potential

	E = 0.2 V	E = 0.4 V	E = 0.7 V	E = 1.0 V	E = 1.3 V
R_s/Ω	685.8	679.8	681.3	683.2	718.9
$C_{dl}/\mu F$	26.2	57.7	49.9	37.3	0.33
R_1/Ω	87.8	89.6	894.1	73.7	2157.0
$Q/Y_o/\mu S s^{-n}$	26.3	69.3	58.8	49.5	46.9
$R_2/k\Omega$	627.1	1300.0	904.3	524.5	49.8
$W/Y_o/\mu S s^{-n}$	0.87	6.03	5.42	119.2	7.051×10^6
Chi squared	8.91×10^{-5}	3.40×10^{-5}	4.46×10^{-5}	6.78×10^{-5}	1.15×10^{-4}

[(5 mM ProDOT-Me₂ was deposited at 100 mV/s, 20 cycle in 0.1 M Bu₄NPF₆/ACN), ($n \sim 0.96$ –0.84)]

EIS measurements were performed in 0.1 M Bu₄ NPF₆/DMF

Table 18.2 Dependence of the parameters calculated from the model (EIS performed in different electrolyte solutions)

	Bu ₄ NPF ₆	Et ₄ NPF ₆	Et ₄ NBF ₄	Bu ₄ NBF ₄
R_s/Ω	1852.0	856.3	858.0	350.6
$C_{dl}/\mu F$	102.6	55.9	29.62	120.8
$R_1/k\Omega$	47.2	0.71	0.11	0.16
$Q/Y_o/\mu S s^{-n}$	48.24	77.83	31.68	70.32
$R_2/k\Omega$	631.8	1473.0	6238.0	432.4
$W/Y_o/\mu S s^{-n}$	5.43	3.18	1.58	368.0
Chi squared	4.14×10^{-5}	9.86×10^{-5}	4.98×10^{-5}	8.47×10^{-5}

[(5 mM ProDOT-Me₂ deposited electrochemically at 100 mV/s, 20 cycle in 0.1 M Bu₄NPF₆/ACN), ($n \sim 0.96$ –0.93)]

Table 18.3 Dependence of the parameters calculated from the Model (EIS performed in different electrolyte solutions)

	ACN	PC	DMF
R_s/Ω	1852	1211	680
$C_{dl}/\mu F$	102.6	61.05	57.72
$R_1/k\Omega$	47.1	0.072	0.089
$Q/Y_o/\mu S s^{-n}$	48.2	127.6	69.34
$R_2/k\Omega$	632	1111	1344
$W/Y_o/\mu S s^{-n}$	5.43	6.28	6.04
Chi squared	4.14×10^{-5}	6.13×10^{-5}	3.40×10^{-5}

[(5 mM ProDOT-Me₂ deposited electrochemically at 100 mV/s, 20 cycle in 0.1 M Bu₄NPF₆/ACN) ($n \sim 0.96$)]

equivalent circuit from 0.2 to 1.3 V for the proposed Model, calculated chi squared values show that it is smaller than 1×10^{-4} , indicating very good agreement between experimental data and simulation results in the 0.2–1.3 V range (Table 18.1). R_s was found to be almost constant and it is independent of E_{dc}

($R_s \sim 680 \Omega$, for $E_{dc} = 0.2\text{--}1.3 \text{ V}$). As it was indicated in literature [19] the solution resistance R_s is not very much dependent on film thickness and E_{dc} . These experimental results allow describing R_s as the solution resistance.

18.3.7 Morphology of Nanoporous and Compact Conductive Polymer Coatings on Carbon Fiber

SEM pictures of uncoated carbon fiber and PProDOT-Me₂ coated carbon fiber morphology is given in Fig. 18.7. The uncoated CFME surface was smooth with its striations, and coated CFME surface exhibits highly porous structure at 100 mV/s. A large amount of electrolyte solution can gain access to the interior of the film through these pores. This open structure will allow fast transport of ions in the supporting electrolyte solution, and as a result, the effective diffusion length of the ions in the conducting polymer is much shorter than the thickness of the film.

Consequently, this process created the porous morphology, the pores were filled with the material being deposited (including dopant anion) decreasing the roughness, leading to an increase in impedance.

18.4 Conclusion

Simulation results of CFME/PProDOT-Me₂ system showed that electrical circuit model was successfully applied to the experimental data. From the effect of potentials on behaviour of capacitance in Bu₄NPF₆/DMF, polymer film shows higher capacitance values at 0.4 V compared to other applied potentials. Shape of the capacitance curves at various potentials in ACN, PC and DMF exhibits similar fashion but, PProDOT-Me₂ film has highest capacitance values for Bu₄NPF₆ in ACN compared to different electrolytes and highest capacitance was obtained in acetonitrile (ACN) compared to PC and DMF. Double layer capacitance (C_{DL}) and low frequency capacitance (C_{LF}) of the modified electrodes decrease while polarization resistance increases, with the increasing of the polarity index (P').

References

1. G. Heywang, F. Jonas, Adv. Mater. **4**, 116 (1992)
2. F. Jonas, L. Schrader, Synth. Met. **41**, 831–836 (1991)
3. K.S. Ryu, Y.G. Lee, Y.S. Hong, Y.J. Park, X. Wu, K.M. Kim, M.G. Kang, N.G. Park, S.H. Chang, Electrochim. Acta **50**, 843–847 (2004)
4. D.M. Welsh, L.J. Kloeppner, L. Madrigal, M.R. Pinto, B.C. Thompson, K.S. Schanze, K.A. Abboud, D. Powell, J.R. Reynolds, Macromolecules **35**, 6517 (2002)

5. A. Cirpan, A.A. Argun, C.R.G. Grenier, B.D. Reeves, J.R. Reynolds, *J. Mater. Chem.* **13**, 2422 (2003)
6. S.P. Mishra, R. Sahoo, A.V. Ambade, A.Q. Contractor, A. Kumar, *J. Mater. Chem.* **14**, 1896 (2004)
7. C.L. Gaupp, D.M. Welsh, J.R. Reynolds, *Macromol. Rapid Commun.* **23**, 885 (2002)
8. B.C. Thompson, P. Schottland, K. Zong, J.R. Reynolds, *Chem. Mater.* **12**, 1563 (2000)
9. C.L. Gaupp, D.M. Welsh, R.D. Rauh, J.R. Reynolds, *Chem. Mater.* **14**, 3964 (2002)
10. D.M. Welsh, A. Kumar, E.W. Meijer, J.R. Reynolds, *Adv. Mater.* **11**, 1379 (1999)
11. I. Schwendeman, J. Hwang, D.M. Welsh, D.B. Tanner, J.R. Reynolds, *Adv. Mater.* **13**, 634 (2001)
12. C. Xu, L. Liu, S. Legenski, M. Le Guilly, M. Taya, A. Weidner, *Proc. SPIE-Int. Soc. Opt. Eng.* **4692**, 442 (2002)
13. C. Xu, L. Liu, S. Legenski, M. Le Guilly, M. Taya, A. Weidner, *Proc. SPIE-Int. Soc. Opt. Eng.* **5051**, 404 (2003)
14. C. Xu, L. Liu, S.E. Legenski, D. Ning, M. Taya, *J. Mater. Res.* **19**, 2072 (2004)
15. P.-H. Aubert, A.A. Argun, A. Cirpan, D.B. Tanner, J.R. Reynolds, *Chem. Mater.* **16**, 2386 (2004)
16. A.A. Argun, M. Berard, P.-H. Aubert, J.R. Reynolds, *Adv. Mater.* **17**, 422 (2005)
17. A.A. Argun, J.R. Reynolds, *J. Mater. Chem.* **15**, 1793 (2005)
18. A.S. Sarac, B. Schulz, A. Gencturk, H.D. Gilsing, *Surf. Eng.* **24**, 5 (2008)
19. E. Barsoukov, J.R. Macdonald, *Impedance spectroscopy: theory, experiment, and applications*, 2nd edn. (Wiley Interscience Publications, NJ, 2005)
20. W.C. Chen, T.C. Wen, H.S. Teng, *Electrochim. Acta* **48**, 641–649 (2003)
21. T.C. Weng, H.S. Teng, *J. Electrochem. Soc.* **148**, A368–A373 (2001)
22. A. Lewandowski, M. Zajder, E. Frackowiak, F. Beguin, *Electrochim. Acta* **46**, 2777–2780 (2001)
23. A.S. Sarac, J. Springer, *Surf. Coat. Technol.* **160**, 227–238 (2002)
24. A.S. Sarac, U. Evans, M. Serantoni, J. Clohessy, V.J. Cunnane, *Surf. Coat. Technol.* **182**, 7–13 (2004)
25. A.S. Sarac, G. Sonmez, F.C. Cebeci, *J. Appl. Electrochem.* **33**, 295–301 (2003)
26. E. Sezer, A.H.S. Sarac, E.A. Parlak, *J. Appl. Electrochem.* **33**, 1233–1237 (2003)
27. A.S. Sarac, J. Springer, *Surface Coating Tech.* **160**, 227 (2002)
28. A.S. Sarac, U. Evans, M. Serantoni, V.J. Cunnane, *Carbon* **41**, 2725 (2003)
29. A.S. Sarac, U. Evans, M. Serantoni, J. Clohessy, V.J. Cunnane, *Surf. Coat. Tech.* **182**, 7 (2004)
30. A.S. Sarac, S.A.M.T. O'Neil, M. Serantoni, J. Henry, V.J. Cunnane, J.B. McMonagle, *Appl. Surf. Sci.* **222**, 148 (2004)
31. M. Serantoni, A.S. Sarac, D. Sutton, *Surf. Coat. Tech.* **194**, 36 (2005)
32. M. Jamal, A.S. Sarac, E. Magner, *Sens. Acts. B* **97**, 59 (2004)
33. A.S. Sarac, M. Serantoni, S.A.M.T. O'Neil, V.J. Cunnane, *J. Nanosci. Nanotechnol.* **5**, 1677 (2005)
34. A.S. Sarac, A. Bismarck, J. Springer, *J. Mater. Sci.* **41**, 389 (2006)
35. A.S. Sarac, *Microelec. Eng.* **83**, 1534 (2006)
36. C.O. Ania, J. Pernak, F. Stefaniak, E. Raymundo-Pinero, F. Beguin, *Carbon* **44**, 3126 (2006)
37. O. Barbieri, M. Hahn, A. Herzog, R. Köt, *Carbon* **43**, 1303 (2005)
38. A.J. Bard, L.R. Faulkner, *Electrochemical methods: fundamentals and applications*, 2nd edn. (John Wiley, New York, 2001)
39. S. Sarangapani, B.V. Tilak, C.P. Chen, *J. Electrochem. Soc.* **143**, 3791 (1996)
40. A.F. Burke, T.C. Murphy, In: D.H. Doughty, B. Vyas, T. Takamura, J.R. Huff (eds.), *Materials for Electrochemical Energy Storage and Conversion: Batteries, Capacitors and Fuel Cells*, Materials Research Society, Pittsburgh, 1995, p. 375
41. E. Frackowiak, F. Beguin, *Carbon* **39**, 937–950 (2001)
42. J.P. Ferraris, M.M. Eissa, I.D. Brotherston, D.C. Loveday, *Chem. Mater.* **10**, 3528–3535 (1998)

43. A.S. Arico`, P. Bruce, B. Scrosati, J.-M. Tarascon, W. Van Schalkwijk, *Nature Mater.* **4**, 366 (2005)
44. R. Kötzt, M. Carlen, *Electrochim. Acta* **45**, 2483 (2000)
45. A.G. Pandolfo, A.F. Hollenkamp, *J. Power Sources* **157**, 11 (2006)
46. K. Zong, L. Madrigal, L. Groenendaal, J.R. Reynolds, *Chem. Commun.* 2498 (2002)
47. B. Sankaran, J.R. Reynolds, *Macromolecules* **30**, 2582 (1997)
48. G. Rauchschalbe, F. Jonas, EP 1142888, (10.10.2001), CA 135 (2001) 289187e
49. A.S. Sarac, H.D. Gilsing, A. Gencturk, B. Schulz, *Progr. Organic Coating* **60**, 281–236 (2007)
50. S. Trasatti, P. Kurzweil, *Platinum Met. Rev.* **38**, 46 (1994)
51. P. Fiordiponti, G. Pistoia, *Electrochim. Acta* **34**, 215–221 (1989)
52. V. Noel, H. Randriamahazaka, C. Chevrot, J. Electroanal. Chem. **558**, 41 (2003)
53. J. Bobacka, A. Lewenstam, A. Ivaska, *J. Electroanal. Chem.* **489**, 17 (2000)
54. R.C. Weast, *CRC Handbook of Chemistry and Physics*, 1st edn. (1988)
55. A.S. Sarac, S.E. Ozgul, A. Gencturk, B. Schulz, H.-D. Gilsing, H. Faltz, *Prog. Organic Coatings* **69**, 527–533 (2010)
56. M.C. Turhan, A.S. Sarac, A. Gencturk, H.-D. Gilsing, H. Faltz, B. Schulz, *Synthetic Metals* **162**, 511–515 (2012)

Chapter 19

Application of Nanoporous Zeolites for the Removal of Ammonium from Wastewaters: A Review

Mustafa Turan

Abstract This contribution deals with low-cost nanoporous zeolite for the removal of ammonium from wastewaters. Adsorption processes are widely used to remove certain classes of pollutants from wastewaters, especially those that are not easily biodegradable. Natural and synthetic zeolites are widely used as adsorbents for gas or liquid purification. Natural zeolite is abundant in the world and is usually regarded as low-cost material. In this review, an extensive list of nanoporous zeolites used for ammonium removal in the literature has been compiled. The review evaluates (i) ammonium adsorption capacities and other parameters in batch systems; (ii) ammonium removal efficiency from wastewaters in fixed-bed systems; and (iii) a combination of biological treatment and adsorption on zeolite for the removal of ammonium from wastewaters. It is important to note that the adsorption capacities of the zeolites presented in this paper vary depending on the characteristics of zeolite, the extent of chemical modifications and the concentration of adsorbate.

19.1 Introduction

The ammonium sources, which are municipal, agricultural and industrial, contribute to the accelerated eutrophication of lakes and rivers, dissolved oxygen depletion and fish toxicity in receiving water. Complete removal of ammonium from process or waste effluents is required due to its extreme toxicity to most fish species. A variety of biological and physicochemical methods and technologies have been proposed for the removal of ammonium from the environment and industrial water systems [1–3].

M. Turan (✉)

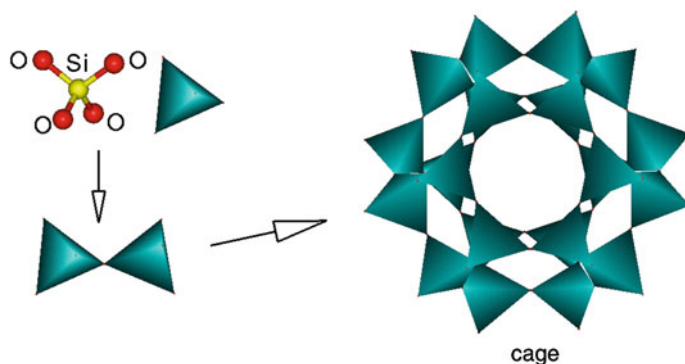
Department of Environmental Engineering, İstanbul Technical University,
34469 İstanbul, Turkey
e-mail: turanm@itu.edu.tr

Commonly, nitrogen removal is achieved by a biological nitrification–denitrification process, where ammonium (NH_4^+) is transformed first to nitrite, then to nitrate, and finally to nitrogen gas [4]. However, since biological methods do not respond well to shock loads of ammonia, unacceptable peaks over the discharging levels may frequently appear in the effluent NH_4^+ concentrations [5]. Besides, the treatment of ammonium nitrogen from wastewaters of low organic content by a biological process usually needs to be supplemented with a carbon source, which may add to the treatment cost. In such a situation, ion-exchange offers an alternative method for the removal of NH_4^+ .

Sanitary landfilling is one of the main disposal options for municipal waste throughout the world. Leachate generated from landfill sites is characterized with high concentrations of organic matter, inorganic salts, heavy metals and ammonium–nitrogen. The discharging of leachate may pose a great threat to the environment. Among the leachate constituents, ammonium is the most significant pollutant in the long term, and its concentration can be found up to several thousand mg/L [6–8]. The release of high level of ammonium without proper treatment decreases the performance of conventional biological treatment systems and accelerates eutrophication, dissolved oxygen depletion and toxicity of living organisms in surface waters [9]. Therefore, the removal of high ammonium concentration from leachate has become a critical issue in recent years. Various treatment technologies have been applied to remove ammonium from leachate, such as biological processes [10, 11], electro-oxidation [12, 13], air stripping [8, 9], chemical precipitation [14, 15], adsorption and ion exchange [6, 16]. Among these treatment methods, ion exchange using zeolites is very attractive due to its low cost and relatively simple application.

The adsorption process provides an attractive alternative for the removal of ammonium from wastewaters while low-cost and readily available adsorbents have been used for various applications. In the literature, aluminosilicates such as clays, zeolites, bentonite and sepiolite are the most commonly used low-cost adsorbents for environmental applications [5, 6, 16]. Zeolites are microporous (pore size <2 nm) crystalline aluminosilicates that are widely used in separation, filtration, adsorption and catalysis [17]. The catalytic performance of zeolites is generally attributed to the existence of a network of micropores with uniform size and shape. However, micropores can also be detrimental to catalytic reactions by limiting the diffusion of reagents and/or products throughout the crystals. One possibility to minimize diffusion limitations is to use zeolite nanocrystals, with a size typically smaller than $0.5\ \mu\text{m}$ [18, 19]. ZSM-5 with very regular hollow architectures can be easily obtained from calcined silicalite-1 nanocrystals [20].

Nanoporous zeolites have negative charges that arise due to isomorphous substitution of Al^{3+} for Si^{4+} and its typical unit cell formula is given either as $\text{Na}_6[(\text{AlO}_2)_6(\text{SiO}_2)_{30}] \cdot 24\text{H}_2\text{O}$ or $(\text{Na}_2, \text{K}_2, \text{Ca}, \text{Mg})_3[(\text{AlO}_2)_6(\text{SiO}_2)_{30}] \cdot 24\text{H}_2\text{O}$ [21]. The framework structure may contain linked cages, cavities or channels which are of the right size to allow small molecules to enter (Fig. 19.1). Three dimensional crystal structure of zeolite contains two dimensional channels which embody some ion exchangeable cations such as Na, K, Ca and Mg. These exchangeable cations

**Fig. 19.1** Zeolite structure**Table 19.1** Chemical composition of natural zeolites for comparison

Constituent (%)	Turkish zeolite [5]	Australian zeolite [41]	Chinese zeolite [42]	Croatian zeolite [43]	Turkish zeolite [28]	Serbia zeolite [33]
SiO ₂	74.4	65.02	65.72	64.93	70.27	66.36
Al ₂ O ₃	11.5	12.76	13.5	13.39	12.90	13.81
CaO	2.00	2.88	3.10	2.00	2.05	3.65
K ₂ O	5.00	1.85	3.14	1.30	1.49	0.99
Fe ₂ O ₃	1.10	1.82	1.30	2.07	1.38	1.69
MgO	0.50	0.95	0.63	1.08	1.71	1.03
Na ₂ O	0.60	0.50	1.16	2.40	3.21	1.03
TiO ₂	0.10	0.34	0.212	–	0.19	–
P ₂ O ₅	0.02	0.10	0.040	–	0.04	–
MnO ₂	0.001	0.04	0.050	–	0.02	–
Loss on ignition	5.85	11.61	11.12	9.63	6.57	13.40

give rise to the ion-exchange properties of the material [22]. Such sorptive properties have been utilized for a variety of purposes such as removal of ammonium [5, 23–28], heavy metal [29–33] and dye [34–39].

Various natural zeolites around the world have shown varying ion-exchange capacities for cations such as ammonium and heavy metal ions. Some zeolites also show adsorption of anions and organics from aqueous solution. Modification of natural zeolites can be done by several methods such as acid treatment, ion exchange and surfactant functionalisation, making the modified zeolites achieving higher adsorption capacity for organics and anions [40]. Table 19.1 presents the chemical analysis of natural zeolites for comparison.

This review (i) presents batch adsorption studies, where adsorption equilibrium, kinetic models and thermodynamic parameters are considered; (ii) evaluates

column adsorption studies with adsorption column design parameters and bed depth service time (BDST) model; (iii) examines modification of natural zeolites and synthesized zeolites as well as regeneration process; (iv) describes ammonium removal efficiency and other parameters for zeolites in the batch and fixed-bed column systems; (v) investigates ammonium removal efficiency for wastewater treatment systems combined with zeolites.

19.2 Methodology

19.2.1 Batch Adsorption Studies

19.2.1.1 Equilibrium Isotherms

Sorption studies were performed by batch technique to obtain rate and equilibrium data. The removal efficiency (%) of adsorbent (zeolite) and the adsorption capacity of solution (ammonium) (q_e , mg/g) was calculated, respectively, as follows;

$$\text{Removal efficiency (\%)} = (C_i - C_e)/C_i \times 100 \quad (19.1)$$

$$q_e = (C_i - C_e)V/m \quad (19.2)$$

where C_i and C_e are the initial and equilibrium solution concentrations (mg/L), respectively, V the volume of the solutions (L) and m the weight of adsorbent used (g). The widely used Langmuir isotherm [44] has found successful application in many real sorption processes and the linear form of Langmuir model is expressed as;

$$C_e/q_e = 1/kb + C_e/b \quad (19.3)$$

where C_e is the concentration of adsorbate in equilibrium solution (mg/L) and the value of b (mg/g) and k (L/mg) are maximum adsorption capacity of adsorbent and the adsorption energy coefficient, respectively. The well-known Freundlich isotherm [45] used for isothermal adsorption is a special case for heterogeneous surface energy, in which the energy term of the Langmuir equation varies as a function of surface coverage strictly due to the variation of the sorption. The linear form of Freundlich model is given as;

$$\log q_e = \log K_f + (1/n) \log C_e \quad (19.4)$$

where K_f is roughly an indicator of the adsorption capacity and n is the adsorption intensity. The magnitude of the exponent $1/n$ gives an indication of the favorability of adsorption. Values, $n > 1$ represent favorable adsorption condition.

19.2.1.2 Kinetics and Thermodynamics

In order to investigate the mechanism of adsorption kinetics, the pseudo-first-order adsorption, the pseudo-second-order adsorption and the intraparticle diffusion model were used to test dynamical experimental data. The first-order rate expression of Lagergren [46] is expressed as follows;

$$\log(q_e - q) = \log q_e - k_1 t / 2.303 \quad (19.5)$$

where q_e and q are the amounts of dye adsorbed on adsorbent (mg/g) at equilibrium and at time t , respectively and k_1 is the rate constant of first-order adsorption (1/min). In many cases the first-order equation of Lagergren does not fit well to the whole range of contact time and is generally applicable over the initial stage of the adsorption processes [47]. The second-order kinetic model [48] is defined as;

$$t/q = 1/(k_2 q_e^2) - t/q_e \quad (19.6)$$

and

$$h = k_2 q_e^2 \quad (19.7)$$

where k_2 (g/mg min) is the rate constant of second-order adsorption and h is the initial adsorption rate (mg/g min). The intraparticle diffusion equation can be described as;

$$q = k_i t^{0.5} \quad (19.8)$$

where k_i is intraparticle diffusion rate constant (mg/g min^{-0.5}). The pseudo-second-order rate constant of solute adsorption is expressed as a function of temperature by the Arrhenius type relationship;

$$\ln k = \ln A - E_a / RT \quad (19.9)$$

where E_a is the Arrhenius activation energy of sorption, representing the minimum energy that reactants must have for the reaction to proceed, A is the Arrhenius factor, R is the gas constant and is equal to 8.314 J/(mol K) and T is the solution temperature.

The thermodynamic parameters such as change in free energy (ΔG°), enthalpy (ΔH°) and entropy (ΔS°) were determined using the following equations [49, 50];

$$K_C = C_e / C_{Ae} \quad (19.10)$$

$$\Delta G^\circ = -RT \ln K_C \quad (19.11)$$

$$\log K_C = \Delta S^\circ / (2.303R) - \Delta H^\circ / (2.303RT) \quad (19.12)$$

where K_C is the equilibrium constant, C_{Ae} is the amount of adsorbate (g) adsorbed on the adsorbent per dm^3 of the solution at equilibrium, C_e is the equilibrium concentration (g/dm^3) of the dye in the solution, T is the solution temperature (K) and R is the gas constant. Generally, the absolute magnitude of the change in free energy for physisorption is between -20 and 0 kJ/mol; chemisorption has a range of -80 to -400 kJ/mol [51]. The negative values of the standard enthalpy change (ΔH°) indicate that the interaction of the adsorbate with adsorbent is exothermic in nature.

19.2.2 Column Adsorption Studies

19.2.2.1 Column Design Parameters

Adsorption performance of the fixed-bed can be evaluated while bed volumes (BV) at breakthrough point ($C/C_0 = 0.1$) is maximum. The breakthrough curves were constructed by plotting the normalized effluent concentration (C/C_0) versus time (t) and/or bed volumes (BV). The bed volumes, BV, and the empty bed contact time, EBCT, at the fixed-bed column are defined as follows;

$$BV = V_F/V_R = Q_t/V_R \quad (19.13)$$

$$EBCT = V_R/Q \quad (19.14)$$

where V_F is the total volume of wastewater treated during the adsorption process at time t (L), V_R is the fixed-bed volume of zeolite (L), C_0 is the influent concentration (mg/L), C is the effluent concentration at time t (mg/L), Q is the feed flow rate in the fixed-bed (L/min) and t is the service time (min). The formation and the movement of the adsorption zone can be evaluated numerically [52, 53]. The time required for the adsorption zone to become established and move completely out of the bed at exhaust time is;

$$t_{Exh} = V_{Exh}/Q = BV_{Exh}EBCT \quad (19.15)$$

The rate at which the adsorption zone (U_z) is moving up or down through the bed is;

$$U_z = h_z/t_z = h/(t_{Exh} - t_f) \quad (19.16)$$

From (19.16), the height of the adsorption zone (h_z) is obtained;

$$h_z = ht_z/(t_{Exh} - t_f) \quad (19.17)$$

where V_{Exh} is the total volume of wastewater treated in the zeolite column at exhaust time (L), h_z is the height of adsorption zone (cm), h is the total bed height (cm) and t_f is the time required for the adsorption zone to initially form (min). The t_f value can be found as follows;

$$t_f = (1 - F)t_z \quad (19.18)$$

At breakthrough, the fraction (F) of adsorbent present in the adsorption zone still possessing the ability to remove solute is;

$$F = S_z/S_{\text{max}} = \int_{V_b}^{V_{\text{Exh}}} (C_0 - C)dV/C_0(V_{\text{Exh}} - V_b) \quad (19.19)$$

where V_b is the total volume of the wastewater treated to the breakthrough point (L), S_z is the amount of solute that has been removed by the adsorption zone from breakthrough to exhaustion; and S_{max} is the amount of solute removed by the adsorption zone when completely exhausted. The percentage of the total column saturated at breakthrough point is

$$\% \text{ saturation} = [1 + (F - 1)h_z/h] \times 100 \quad (19.20)$$

19.2.2.2 Bed Depth Service Time (BDST) Model

The objective of fixed-bed operations is to reduce the concentration in the effluent so that it does not exceed a specific breakthrough concentration (C_b). The original work on the BDST model was carried out by Bohart and Adams, who proposed a relationship between bed depth, Z , and the time taken for breakthrough to occur [54]. The service time, t , is related to the process conditions and operating parameters;

$$\ln(C_0/C_b - 1) = \ln(e^{k_{\text{ad}}N_0Z/V} - 1) - k_{\text{ad}}C_0t \quad (19.21)$$

Hutchins [55] proposed a linear relationship between the bed depth and service time;

$$t = N_0Z/(C_0V) - (1/k_{\text{ad}}C_0)\ln(C_0/C_b - 1) \quad (19.22)$$

where C_b is the solute concentration at breakthrough point (mg/L), k_{ad} is the adsorption rate constant (L/(mg h)), N_0 is the adsorption capacity (mg of solute/L of adsorbent), Z is the bed depth of column (cm), V is the linear flow velocity of feed to bed (cm/h) and t is the service time of column (h). The critical bed depth, Z_0 , is the minimum column height necessary to produce an effluent concentration C_b . By letting $t = 0$, Z_0 is obtained from (19.22)

$$Z_0 = (V/k_{ad}N_0)\ln(C_0/C_b - 1) \quad (19.23)$$

Equation (19.22) enables the service time, t , of an adsorption bed to be determined by a specified bed depth, Z , of adsorbent. t and Z are correlated with the process parameters and the influent concentration of solute, solution flow-rate and the adsorption capacity [56] (19.22) can be expressed in the following form:

$$t = mZ + n \quad (19.24a)$$

where

$$m = \text{slope} = N_0/C_0V \quad (19.24b)$$

$$n = \text{intercept} = -(1/k_{ad}C_0)\ln(C_0/C_b - 1) \quad (19.24c)$$

19.2.3 Modification of Zeolite

Several modification methods have been tried to enhance the ammonium adsorption capacity of zeolite, including hydrothermal pretreatment, NaOH, HCl or NaCl solution treatment, integrated calcination, and others.

Sodium-natural zeolite (SNZ): The optimum modification conditions for Yemeni natural zeolite on the effectiveness of ammonium ion (NH_4^+) removal was investigated [57]. The natural zeolite (NZ) powder (200–230 mesh) was washed with distilled water to remove undesirable materials and dried at a temperature of 100 °C for 24 h and then modified with sodium chloride. The suspension was stirred in 500 ml conical flasks using a magnetic stirrer water bath at a rate of 120 rpm and 90 °C for 2 h. Subsequently, the suspension was filtered and washed with distilled water. The wet modified material was dried at 100 °C in an oven for 24 h. In a similar study [58], natural zeolite was also modified with NaCl to investigate the mechanisms for ammonium adsorption onto zeolite.

Silicate-Carbon Modified Zeolite (SCMZ): The Silicate-Carbon Modified Zeolite (SCMZ) was used for removing ammonium from drinking water [59]. The modification process applied in the preparation of SCMZ consisted of three steps: (1) The clinoptilolite powder (74 μm) was repeatedly washed with tap water and then dried at 100 °C for 4 h. (2) The zeolite was dried and 1000 g of the dry zeolite was dispersed into 1 L sodium chloride solution. The sample was repeatedly washed with tap water and then dried at 100 °C for 4 h. Finally, the dried sample was ground and sieved to 74 μm again and designated as sample A. (3) Sample A, Na_2SiO_3 and powdered activated carbon were mixed at weight/weight ratio of 100:9:2, and mixed evenly; then, 10 % tap water was added and stirred again. The mixture was shaped into a cylinder by an extrusion method and then dried at 100 °C for 2 h and calcined in a muffled furnace at 500 °C for 2 h and the SCMZ filter was obtained.

Calcium formed zeolite (CaY): A modified clinoptilolite zeolite- Ca^{2+} -formed (CaY) was prepared for the removal of ammonium ions from aqueous solutions [60]. The chosen clinoptilolite was ground and sieved into 0.425–0.970 mm particle size and then washed twice with distilled water (volume ratio of liquid/solid 3:1). NaY was gained by treating the samples three times with saturated NaCl solution at boiling point for 2 h and then changed into NH_4^+ -formed clinoptilolite (NH_4Y) by washing it with 1 mol/l NH_4Cl solution for 20 min at room temperature. Finally, CaY was prepared through washing the NH_4Y with $\text{Ca}(\text{OH})_2$ solution at boiling point for 2–3 h.

Integrated modification process: a natural calcium-rich zeolite modified by an integrated process was used for ammonium ion uptake (AIU) from drinking water [42]. This process consisted of pretreatment (grinding and sieving), sodium salt modification and calcination. The introduction of sodium ion enhanced the heat resistance of the raw material from 150 to 400 °C and Na^+ ion treatment followed by calcination effectively improved pore and surface properties of zeolite.

19.2.4 Synthesized Zeolites

Due to its uniform micropore structure and high surface area, synthetic zeolite usually has a higher adsorption capability compared with natural zeolite. That's why many researchers have investigated the synthesis of zeolite from low-cost raw materials like coal fly ash, red mud, fly ash, halloysite etc. in order to obtain low-cost and effective synthetic zeolites.

Zeolite synthesized from halloysite mineral: The adsorption behavior of ammonium ions (NH_4^+) from aqueous solution onto synthesized NaA zeolite was investigated [61]. NaA zeolite was prepared using alkaline fusion followed by hydrothermal treatment method. 2 g of halloysite powder was placed in a Ni crucible and fused with 2.6 g NaOH powder at 673 K for 2 h to promote their reactivity. The fused mass obtained was cooled and mixed with deionized water (38 mL) by magnetic stirring for 30 min until the reaction gel was homogenized. The aluminosilicate gel was aged for 2 h at 313 K in sealed Teflon reactor to rearrange the reactant for forming nuclei. After that, the mixture was crystallized at 363 K for 5 h under static condition. Solid powder was filtered-off, washed and dried at 383 K for 12 h and the product was obtained. In a similar study [62], the removal of ammonium from wastewater by zeolite synthesized from halloysite was also investigated.

Zeolite synthesized from fly ash by a fusion method: A fusion method, involving alkaline fusion followed by hydrothermal treatment, was adopted for the synthesis of zeolite [63, 64]. 9 g of fly ash were mixed with NaOH powder to obtain a homogeneous mixture. The ratio of fly ash to NaOH was 1:1.3 (w/w). The mixture was then heated in a nickel crucible in air at 600 °C for 90 min. The fusion products were ground and poured into a plastic bottle, followed by the addition of distilled water to form a mixture that contains 17.25 % of the fusion product (w/w). The

mixture was stirred intensely at 80 °C for 1 h to form an aluminosilicate gel and then poured into a stainless alloy autoclave and kept in oven at 100 °C for 9 h. After hydrothermal treatment, the precipitated samples were extracted and washed with distilled water until the pH of the samples was about 9. The samples were dried at 100 °C for 12 h and sieved through a 100 mesh (<150 μm). In another study [65], the simultaneous removal of ammonium and phosphate by zeolite synthesized from fly ash (ZFA) and saturated with different cations was investigated.

Integrated synthesis of zeolites 4A and Na-P1 using coal fly ash: Fly ash from Brazilian coal combustion was used for green integrated synthesis of zeolite 4A and

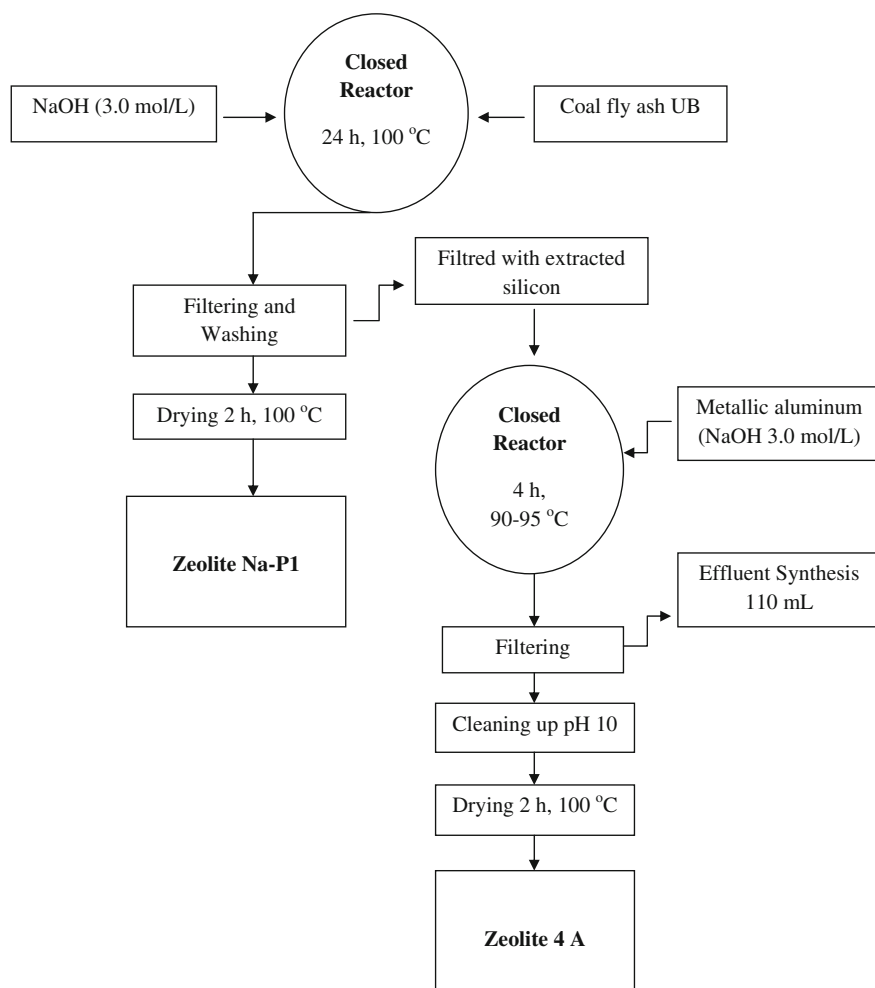


Fig. 19.2 Flow diagram of the zeolites synthesis 4A and Na-P1 by integrated process 1 [66]

Na–P1 [66]. Integrated synthesis of zeolites 4A and Na–P1 was carried out using two different pathways, denominated processes 1 and 2. In process 1 (Fig. 19.2) ash was activated by NaOH solution (3.0 mol/L and L/S ratio of 6 L/kg) using the conventional hydrothermal process in a closed borosilicate reactor at 100 °C for 24 h. The resulting product was separated by filtration and washed several times with deionized water to remove excess sodium and lower the alkaline content, then dried (100 °C for 2 h) and stored in a desiccator. The molar ratio was adjusted to 1.0 (for zeolite 4A formation) by adding an alkaline solution (3.0 mol/L) containing Al, previously dissolved from aluminum metal powder. The resulting solution was transferred to a borosilicate glass reactor, stirred manually (2 min at 25 °C) and immediately heated at 90 °C for 1.5 h, followed by another 2.5 h at 95 °C.

Zeolite synthesized from red mud: The removal of ammonium from aqueous using zeolite synthesized from red mud was studied [67]. Red mud was initially sieved and the particles smaller than 100 mesh were kept for zeolite synthesis. The hydrothermal method was adopted for the synthesis of zeolite. Extra silicon ($\text{Na}_2\text{SiO}_3 \cdot 9\text{H}_2\text{O}$, Nona hydrate of sodium silicate) was added to make the total mole ratio $\text{Si}/\text{Al} = 4.5$, and a NaOH solution was added to make ratio $\text{Na}/\text{Si} = 1.3$. The mixture was then heated in a homogeneous reactor for an 8 h aging time, and a 16 h crystallization time at a temperature of 120 °C. After applying the above treatment, the precipitated samples were extracted and washed with distilled water until the pH of the samples was about 9. Finally, the samples were dried at 100 °C for 12 h for further studies. In order to reduce the negative influence of high Fe_2O_3 content (40.96 %) of red mud to zeolite synthesis, high alkalinity conditions were used for zeolite synthesis in this study.

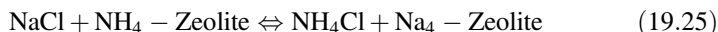
19.2.5 Regeneration

The ammonium-removal efficiency of zeolite decreases after operating for a long duration. Thus, zeolite should be regenerated before using it again. There are various regeneration methods such as heating regeneration, acid regeneration, NaCl regeneration and $\text{Ca}(\text{OH})_2$ regeneration (Table 19.2). The breakthrough zeolite

Table 19.2 Methods of zeolite regeneration [59]

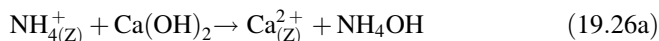
Regenerate methods	Step of regeneration
Heating regeneration	The breakthrough zeolite was calcined in a muffle furnace at 200 °C for 24 h
Acid regeneration	The breakthrough zeolite was treated with 1.0 mol/L HCl solution for 24 h, repeatedly washed with tap water, and dried at 45 °C for 4 h
NaCl regeneration	The breakthrough zeolite was treated with 2.0 mol/L NaCl solution for 24 h, repeatedly washed with tap water and dried at 100 °C for 4 h
$\text{Ca}(\text{OH})_2$ regeneration	Clinoptilolite saturated with ammonium ions was regenerated by 0.2 mol/l $\text{Ca}(\text{OH})_2$ emulsion at 100–110 °C

regenerated using NaCl maintained its initial adsorption capacity. The following reaction occurred in the process of NaCl regeneration:



In the process of regeneration, Na^+ exchanged the NH_4^+ that was adsorbed on zeolite, and the ammonium-removing property of zeolite was restored [59].

A consideration of practical interest with regard to the application of a new exchanger is its desorption and regeneration ability. The $\text{Ca}(\text{OH})_2$ regeneration happened as follows:



This operation went on with generating free ammonia, so some ammoniated by-products (such as NH_4Cl) can be gained. The mass of the zeolite during the regeneration and the adsorption capacity of the regeneration modified zeolite for ammonium ions after regeneration were measured, respectively. The column was regenerated with $\text{Ca}(\text{OH})_2$ emulsion at pH 11–12 and at flow velocity of 5 ml/min [60].

19.3 Nanoporous Zeolites for Ammonium Removal from Wastewaters

19.3.1 Ammonium Adsorption Capacities and Other Parameters for Zeolites in the Batch Systems

An overview of nanoporous zeolites for the removal of ammonium is presented in this article along with their adsorption capacities and other parameters and the experimental conditions used in batch systems (Table 19.3). In 2004, the ion uptake of ammonium ion onto New Zealand clinoptilolite and mordenite are compared [27]. The results showed that the mordenite exhibited higher overall uptake concentrations at equilibrium compared to clinoptilolite having solution concentrations greater than 80 mg/L. The ion exchange capacity of New Zealand clinoptilolite for NH_4^+ ion was found as 6.59 mg/g.

In 2006, the ion exchange kinetics and equilibrium isotherms of ammonium ion on the natural Turkish zeolite (clinoptilolite) was studied [5]. Ammonium adsorption capacity increased from 1 to 4.75 mg/g with the increase of initial NH_4^+ concentration (Fig. 19.3). For lower initial concentration of NH_4^+ , equilibrium time was lower than that of the higher concentration due to the increased competition for the active sites with the increase of NH_4^+ concentration. This is consistent with the ion-exchange surface becoming increasingly saturated with ammonium ion [68].

Table 19.3 Ammonium adsorption capacities and other parameters for natural and modified zeolites

Adsorbent	Wastewater	d_p (mm)/ S_p (m ² /g)	q_m (mg/g)	C_i (mg/L)	t (min)	pH	References
Zeolite (Iranian)	Contaminated surfacewater		49.7	10–100	5–90	2–10	[75]
Synthesized zeo. NaA-H	Synthetic wastewater		44.3	10–300	120	7	[61]
Acid treat. zeo. (Azerbaijan)	Synthetic wastewater	0.0005–0.001	43.47	10–200	2–120	4–9	[84]
Synthesized zeo. Y-H	Synthetic wastewater	0.001–2/79.62	38.64	10–300	60	6–7	[62]
Synthesized zeolite FA	Domestic wastewater	<0.15	37.45	10–300	40	8	[63]
Zeolite (Jordanian)	Synthetic wastewater	0.5–1	37.175	15–50	2880	5.5–9	[90]
Zeolite (Zeolith, Deutch.e.K.)	Synthetic wastewater	0.013	36	10–200	60	6–8.5	[76]
Synthesized zeolite (ZFA)	Simulated wastewater	80 mesh	35.31	2.5–1200	1440	4.98–6	[65]
Zeolite (USA, TR, EU)	Synthetic wastewater	0.355–2.50	22.8–32.4	5–500	250	7.5	[89]
Synthesized zeo. FA Na-PI	Swine wastewater	0.003–0.005	31–4	1205–125	30	7.11;10.7	[66]
Zeolite (New Zealand)	Simulated wastewater		29	25–200	5760		[69]
Zeolite NZ, BZ, BAZ (Chinese)	Synthetic wastewater	1–15/6.64	18.3–28.9	16.5	5400		[79]
Synthe. zeo. FA LC-Z, HC-Z	Synthetic wastewater	<0.15	28.65–6.12	10–300	0–360	8	[64]
Modified zeolite MWZ (Chinese)	Synthetic wastewater	0.083	23.83	0–200	360	6.5	[85]
Acid treated zeolite (Turkey)	Municipal wastewater	1–2	23.70	8.8–40.3	5–120	4–7	[28]
Zeolite (Croatian)	Synthetic wastewater	0.1–0.2	21.6	25.7–700	250	3–8	[70]
Modified zeolite (Japan)	Synthetic wastewater	0.001–25/20–115	12.5–19.5	400	1440		[82]
Zeolite/modified zeo. (Bulgarian)	Synthetic wastewater	0.2–1/26	7.85–18.4	50–250	60	2.5–9	[83]
Zeolite (Iranian)	Synthetic wastewater	0.5–2.3/22	17.8	10–220	180	5.5–7.6	[88]
Synthesized zeolite RM	Synthetic wastewater	0.007; 0.01	17.5	5–500	30	2–11	[67]
Zeolite/modif. Zeo. (Chinese)	Synthetic wastewater	1.43;0.80/14; 60.8	11.0–17.3	10–4000	15–1440	3.4–11.1	[58]
Zeolite/modified zeo. (Chinese)	Synthetic wastewater	2;1/10.53–30.1	4.04–17.24	25	120	7.1	[42]
Zeolite (Turkey)	Synthetic wastewater	0.3–0.6	15.4	5–150	720	6.5	[72]
Zeolite (Iranian)	Synthetic wastewater	0.6–0.8	11.31–13.27	90–3610	10–300		[78]

(continued)

Table 19.3 (continued)

Adsorbent	Wastewater	d_p (mm)/ S_p (m ² /g)	q_m (mg/g)	C_i (mg/L)	t (min)	pH	References
Zeolite (Romania)	Synthetic wastewater	0.2–0.4/12–35.7	13.15–11.85	70	1440	7	[91]
Zeolite (Turkey)	Synthetic wastewater	0.3–0.6	11.25	20–100	5–50	4–10	[71]
Zeolite (Iranian)	Synthetic wastewater	0.6–0.8	11.31–13.27	90–3610	10–300		[78]
Zeolite (Romania)	Synthetic wastewater	0.2–0.4/12–35.7	13.15–11.85	70	1440	7	[91]
Zeolite (Turkey)	Synthetic wastewater	0.3–0.6	11.25	20–100	5–50	4–10	[71]
Zeolite (Turkey)	Municipal wastewater	0.5–1	10.4	20	1440	7.3	[94]
Zeolite (Turkey)	Synthetic wastewater	0.075	9.64	20–400	1440	8	[39]
Synthesized zeo.-Na-PI	Acid mine drainage	0.038–0.25	9.6		30	3.3	[86]
Zeolite (Chinese)	Synthetic wastewater	0.074–0.3	9.41	50–300	300	4–9	[81]
Modified zeo.-Ca Y (Chinese)	Simulated wastewater	0.425–0.970	3.03–7.67	300	120	4–9	[60]
Zeolite (Turkey)	Contaminat. Surf. water	1–2 and 2–2.8	7.4	1–1400	60	6.5	[26]
Zeo. (New Zealand)	Simulated wastewater	0.5–0.71	6.59	10–200	4320	7.5	[27]
Zeolite (Australian)	Greywater		6.30	5–50	480	2–10	[41]
Zeolite (Chinese)	Landfill leachate	1–2	5.87–6.02	2000–2300	480	3–9.9	[92]
Zeolite (Turkey)	Synthetic wastewater	1–1.4/11.8	5.95	25–150	60	7	[5]
Zeolite (New Zealand)	Wastewater	0.25/2.83/30–40	5.74–5.72	15–750	1020	5.4–8.2	[87]
Zeolite (USA)	Synthetic wastewater	2–4/521	3.18–4.5	0–500	1440		[93]
Zeolite (Croatian)	Leakage waters WD	0.5–2.0	4.2	800	15–360	7.3	[43]
Zeolite (Hungarian)	Synthetic wastewater	0–0.125/1500	3.79	50–500	30–180	6–7	[77]
Zeolite/modif. zeolite (Yemeni)	Synthetic wastewater	0.008; 64/36; 41	2.83; 3.68	10–250	20–120	7	[57]

H halloysite; *FA* fly ash; *RM* red mud; *Na-ZFA* zeolite synthesized from fly ash; *LC-Z* low-calcium zeolite; *HC-Z* high calcium zeolite; *Na-PI* zeolite synthesized from fly ash; *Zeolite Y* zeolite synthesized from halloysite; *NZ* natural fresh zeolite; *BZ* bio-zeolite made from natural fresh zeolite; *BAZ* bio-zeolite made from ammonium-bearing zeolite; *MWZ* microwave-treated zeolite; *WD* waste dump

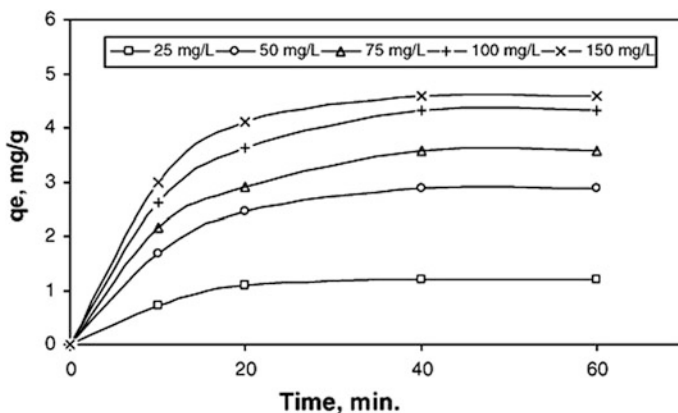


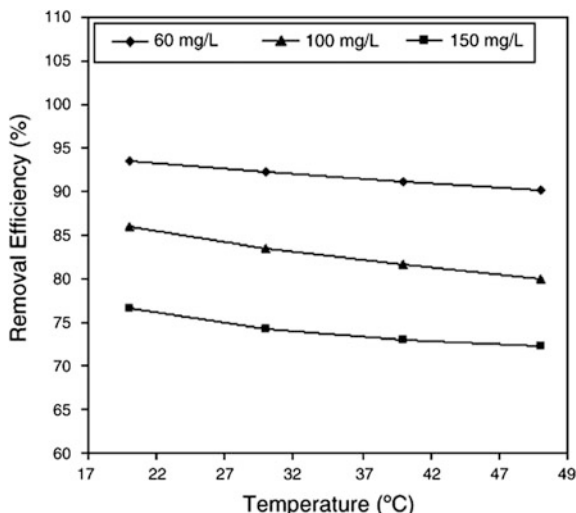
Fig. 19.3 Effect of contact time and initial concentration on the uptake of NH_4^+ by natural zeolite ($T = 40^\circ\text{C}$, solid concentration = 10 g/L, agitation rate, 200 rpm) [5]

NH_4^+ -N removal increased in the range of 8.8–40.3 mg NH_4^+ -N/L and also decreased after 40.3 mg NH_4^+ -N/L. This is attributed to the weak binding between ion exchanger material (zeolite) and NH_4^+ ion and can be expected from zeolite having micropores and macropores [69]. The maximum amount of exchanged NH_4^+ using natural Turkish clinoptilolite was 5.95 mg/g at 25°C .

The ability of a Croatian clinoptilolite- and montmorillonite-rich tuffs to take up ammonium from aqueous solutions was investigated [70]. The clinoptilolite efficiency for ammonium removal was far superior (by 45–50 %) to the removal efficiency of montmorillonite. Maximum loadings in ammonium exchange in clinoptilolite tuff, montmorillonite tuff, alkaline modified montmorillonite tuff and acid modified montmorillonite tuff, as calculated by the Langmuir equation, are 1.20, 0.54, 0.64, and 0.21 mmol/g NH_4^+ , respectively. Doğan et al. [71] investigated the parameters affecting on the equilibrium of ammonium exchange by zeolite and the applicability of linear and nonlinear regression for the estimation of kinetics and isotherm models. In a similar study, Gunay [72] evaluated the experimental data of ammonium exchange by natural Bigadic clinoptilolite using nonlinear regression analysis. Thermodynamic investigation indicated that ammonium removal by clinoptilolite was favorable at lower temperatures and exothermic in nature.

The effect of shaking time on the removal of NH_4^+ ion using natural zeolite was investigated [39]. Removal efficiency of NH_4^+ ion increased with increasing shaking time and 75 % of NH_4^+ ion removal was completed within 15 min. The removal efficiency increased by raising the pH from 4 to 8 and reached a maximum value at pH 8 and then decreased. The decrease in the NH_4^+ removal efficiency after pH 8 may be due to two factors: (i) partial dissolution of the natural zeolite [73]; (ii) the conversion probability of NH_4^+ into NH_3 [74]. Moussavi et al. [75] examined the applicability of natural zeolite for simultaneous removal of ammonia and humic acid, two of the most encountered contaminants in the surface waters.

Fig. 19.4 The effect of temperature on NH_4^+ ion removal using the natural zeolite [39]



The maximum experimental adsorption capacities of ammonia and humic acid as binary components were 49.7 and 10.5 mg/g, respectively.

The effect of temperature on ammonium adsorption was studied at 25, 40 and 55 °C [75] and at 21, 30, 40, and 50 °C [39]. The results indicate that the amount of NH_4^+ exchanged onto zeolite increases with a decrease in temperature as seen in Fig. 19.4, which may be due to a tendency for the ammonium molecules to escape from the solid phase to the bulk phase with an increase in temperature of the solution. This result also indicated that the adsorption process of NH_4^+ onto natural zeolite was exothermic in nature. In 2009, removal of both ammonium and phosphorus nutrients by natural zeolite has been studied in lab scale by using a mechanically stirred batch system [76]. Results indicated that the removal efficiency and the uptake of ammonium by zeolite are both dependent on the initial ammonium and zeolite concentrations, indicating rise in uptake of ammonium by zeolite for higher ratios. Widiastuti et al. [41] also investigated the effectiveness of removing ammonium ion and the theoretical aspects of adsorption as well as desorption–regeneration studies. Natural zeolite shows good performance with up to 97 % of ammonium removal. Gibbs free energy change (ΔG°) at 25 °C is negative (−19.52 kJ/mol) indicating the spontaneous nature of the adsorption process. The enthalpy change (ΔH°) is positive (30.96 kJ/mol) indicating endothermic adsorption process. The entropy change (ΔS°) at 25 °C is also positive (0.169 kJ/mol K) indicating increasing randomness at the solid-solution interface during adsorption.

In another study, the adsorption capacity of clinoptilolite zeolite towards ammonium ions was evaluated [77]. The highest adsorption capacity of clinoptilolite towards ammonium ions was found 3.79 mg/g for the initial concentration of 300 mg/L. The highest removal level reached by clinoptilolite for ammonium was 99.74 % for the initial concentration of ammonia of 50 mg/L. Malekian et al. [78] studied to remove ammonium (NH_4^+) from aqueous solutions with different Na^+

concentrations by natural Iranian zeolite. The maximum NH_4^+ exchange capacity by the zeolite at equilibrium was measured at 0.03 M Na^+ concentration, which was 11.31 and 13.27 mgNH_4^+/g for millimeter- and nanometer-sized zeolites, respectively. Wen et al. [79] reported the sorption kinetics of ammonium onto three types of zeolite, where he found that ion exchange increased with decreasing particle size of the zeolite. The capacity sequence was found to be $\text{NZ} > \text{BZ} > \text{BAZ}$ while larger particle size (8–15 mm) was used (Table 19.3). Compared to the virgin zeolite, the ion exchange rate in the bio-zeolite was reduced about 25–30 % and the rate-controlling step for ion exchange shifted from pore diffusion in the virgin zeolite to film diffusion in the bio-zeolite [80].

In 2010, the removal of ammonium ion from aqueous solutions using natural Chinese (Chende) zeolite was studied [81]. The effect of other cations on the removal of ammonium followed the order of preference $\text{Na}^+ > \text{K}^+ > \text{Ca}^{2+} > \text{Mg}^{2+}$ at identical mass concentrations and the effect of the presence of individual anions followed the order of preference carbonate > chloride > sulfate > phosphate at identical mass concentrations of ammonium ions. The maximum amount of NH_4^+ adsorbed by the zeolite at equilibrium (M) was 9.41 mg/g in the individual presence of phosphate and 5.05 mg/g in the individual presence of carbonate.

Chemical modifications of zeolite were also conducted to improve its removal performance and adsorption capacity for ammonium ions. In 2009, it was studied to modify the natural clinoptilolite zeolite available locally (Akita Prefecture, Japan) for its ammonium ions retention capacity [82]. The natural clinoptilolite was modified chemically and mechanically with changing time duration of sodium hydroxide treatment and ball to powder mass ratio in wet ball milling, respectively. The maximum NH_4^+ retention capacities of natural clinoptilolite (NZeo), clinoptilolite treated with NaOH solution for 72 h (Zeo-72) and wet-milled clinoptilolite (WM-50) according to Langmuir model were 0.89, 1.15 and 1.39 mmol/g , respectively. Vassileva and Voikova [83] investigated ammonium ions removal from aqueous solutions by natural and NaCl pretreated Bulgarian clinoptilolite. The maximum adsorption capacities for ammonium ions shown by natural and pretreated clinoptilolites (CL Na and CL Na(t)) were 7.85, 12.29 and 18.40 mg/g , respectively.

In a another study, Lin et al. [58] investigated the mechanisms of ammonium adsorption onto natural zeolite and NaCl-modified counterpart under different ammonium levels (10–4000 mg-N/L) and initial pH 3.4–11.1. Alshameri et al. [57] investigated the potential use of Yemeni (Al-Ahyuq) natural zeolite on the effectiveness of ammonium ion removal in a batch system. The highest ion exchange capacity was obtained at 35 °C and pH of 8. In 2014, Khosravi et al. [84] studied the removal of ammonium ion from aqueous solutions using the natural Western Azerbaijan zeolite. NaCl treatment of the adsorbent showed an improvement of ammonium ion removal from aqueous solutions over other treatments. Maximum uptake for ammonium ion was 43.47 mg/g .

The development of physical and chemical properties of a natural calcium-rich zeolite modified by an integrated process, as well as the relation between the development and ammonium ion uptake (AIU), was studied [42]. By the optimum modification process, NH_4^+ ion exchange capacity was largely enhanced and the

maximum AIU (17.24 mg/g) of the modified zeolite was 4.3 times that of the raw zeolite (4.04 mg/g). In 2008, natural Chinese zeolite pretreated by sodium chloride solution under microwave irradiation was used to remove ammonium from aqueous solution [85]. The microwave-treated zeolite had the highest ammonium adsorption capacity, followed by the thermal-treated zeolite and the fresh zeolite.

Many researchers have reported zeolite synthesis using halloysite mineral, fly ash and red mud etc. for ammonium ion removal. In 2006, Wu et al. [65] studied Na-ZFA (zeolite synthesized from fly ash) converted into Ca-, Mg-, Al-, and Fe-ZFA by salt treatment and the simultaneous removal of ammonium and phosphate by ZFA saturated with different cations. It was shown that Al^{3+} -ZFA had the highest removal efficiencies (80–98 %) for ammonium. It was concluded that ZFA could be used in simultaneous removal of NH_4^+ and phosphate at low concentrations with pre saturation by an appropriate cation such as Al^{3+} through salt treatment. Zhao et al. [62] investigated ammonium adsorption properties of the synthesized zeolite Y using batch experiments. Due to its low cost, high adsorption capacity and fast adsorption rate, the zeolite Y synthesized from halloysite has the potential to be utilized for the cost-effective removal of ammonium from wastewater. The maximum adsorption capacity is 38.64 mg/g at 25 °C. In another similar research, adsorption behavior of ammonium ions (NH_4^+) from aqueous solution onto NaA zeolite was investigated [61]. A maximum adsorption capacity of 44.3 mg/g of NH_4^+ was achieved.

Zhang et al. [63] investigated the removal of ammonium from aqueous solutions using zeolite synthesized from fly ash by a fusion method, which combines alkaline fusion followed by hydrothermal treatment. The maximum ammonium uptake capacity obtained is 24.3 mg/g. The regenerated zeolite has almost the same ammonium uptake capacity as the original zeolite. In another similar study [64], zeolites were synthesized from low-calcium (LC-Z) and high-calcium (HC-Z) fly ashes, respectively. The obtained maximum ammonium uptake capacities are 23.8 mg/g for LC-Z and 3.17 mg/g for HC-Z in the synthetic solution. The Ca^{2+} leaching and the lower zeolite content in HC-Z account for its lower uptake capacity. In 2015 [66], Brazilian coal fly ashes were used for integrated synthesis of zeolites 4A and Na-P1 by two different routes and under moderate operating conditions (temperature and pressure). Treatment of swine wastewater with zeolite Na-P1 resulted in a high removal capacity for total ammoniacal nitrogen (31 mg/g).

In another study [86], the zeolite synthesized in the optimized process (77 % Na-P1) was used to remove contaminants from acid mine drainage (AMD). Treatment of this wastewater with 10 g/L of synthetic zeolite for 30 min presented high removal of arsenic, nickel, calcium, copper, iron and manganese ions and the partial yet significant removal of ammonium, magnesium, potassium and zinc ions. Zhao et al. [67] also studied the removal of ammonium from aqueous solutions using zeolite synthesized from red mud. The CEC increased from 81.9 to 111 mmol/100 g during the synthesis process. The maximum ammonium adsorption capacity obtained is 17.5 mg/g.

19.3.2 Ammonium Removal from Wastewaters in the Batch and Column Systems

The removal of ammonium ions from wastewater using natural zeolite was carried out in both batch and fixed bed (column) systems in the literature. In 2005, Sarioğlu [28] studied the ammonium ion adsorption onto Doğantepe (Turkey) zeolite. Increasing initial ammonium nitrogen concentration from 5.0 to 12.0 mg/L, increased the exchange capacity from 0.70 to 1.08 mg NH_4^+ -N/g (Table 19.3). The flowrate and pH values, at which the highest adsorption capacities were obtained, were found to be 0.5 ml/min (0.87 mg NH_4^+ -N/g zeolite) and 4 ml/min (mg NH_4^+ -N/L), respectively (Table 19.4). The cation exchange capacity of Dogantepe zeolite was found to be 164.62 meq/100 g.

In 2003, the ammonia removal from contaminated drinking water using clinoptilolite zeolite was studied in bottle adsorption tests and optimum adsorption conditions were established [26]. In the fixed bed studies, the breakthrough curves showed that increasing the influent concentration (10–20 mg/L) increased the effluent concentration. Twice-regenerated clinoptilolite showed the best performance compared with the others. The regeneration of clinoptilolite with 30 g/L NaCl solution at pH 11.5 was found to have activated the clinoptilolite and increased the ion exchange performance.

Nguyen and Tanner [87] evaluated the capacity of two natural New Zealand zeolites (clinoptilolite and mordenite) to remove NH_4^+ from a range of wastewaters under both batch and flow through conditions. The maximum NH_4^+ removal capacity of the New Zealand zeolites studied (5.7–8.3 g NH_4 -N/kg) is within the range (5–15 g NH_4 -N/kg) reported for overseas zeolites. Ashrafzadeh et al. [88] studied the capability of Iranian natural clinoptilolite for ammonia removal from aqueous solutions. In the batch system, an adsorption capacity about 17.8 mg NH_4^+ /g at initial ammonia concentration of 50 mg/L was obtained. The adsorption capacity of clinoptilolite in the continuous mode was about 15.16 and 15.36 mg NH_4^+ /g for the original and regenerated types of clinoptilolites, respectively.

In 2008, removal of low-concentration ammonium by eight commercial natural zeolites was studied [89]. In the batch tests, the ammonium adsorption capacity (up to 32.4 mg NH_4^+ /g) decreased with increasing particle size (355–2500 μm). In the column tests, the adsorption capacity of chabazite was 48.3 mg NH_4^+ /g. The zeolites were rapidly regenerated by an alkaline sodium chloride solution fed to the columns. Mashal et al. [90] investigated the performance of natural Jordanian zeolite tuff to remove ammonia from aqueous solutions. Breakthrough time increased by increasing the bed depth as well as decreasing zeolite particle size, solution flow-rate, initial NH_4^+ concentration and pH. In 2011, zeolitic volcanic tuffs (ZVT) originated from Transylvania, Romania were used to remove ammonium ions from model solution [91]. The ionic exchange process was realized in batch and fixed bed column conditions. It was found that ionic exchange efficiency varies according to their zeolite content.

Table 19.4 Ammonium removal efficiency and other parameters for zeolites in fixed-bed systems

Adsorbent	Wastewater	d_p (mm)	C_i (mg/L)	Q (mL/min)	Removal (%)	$BV_b(BV)/t_b$ (h)	References
Zeolite (Turkey)	Contaminated drinkingwater	1–2 and 2–2.8	10–20	70.2	90	275 and 140	[26]
Zeolite (Turkey)	Municipal wastewater	1–2	5–12	4, 2, 1 and 0.5	80	/4, 11, 25 and 56	[28]
Zeolite (Croatian)	Leakage water	0.5–2	820	25	100	4.5	[43]
Zeolite (Turkey)	Landfill leachate	0.6–1.2	100–400	6.8–21	90	29.11–22.47	[16]
Zeolite (Turkey)	Landfill leachate	0.6–1.2	200	21, 11 and 6.8	90	/7, 30 and 65	[6]
Zeolite (Chinese)	Landfill leachate	1–2	2057	4–9	69.4–31.6	/3	[92]
Zeolite (Chinese)	Synthetic wastewater	0.25–0.70	50	2	100	/22	[95]
Zeo. (New Zealand)	Synthetic wastewater	0.25–2.83	5–15	0.47; 15.9	100	60	[87]
Modif. zeo. CaY (Chinese)	Simulated wastewater	0.425–0.970	250	10–70	100	78	[60]
Zeolite (Iranian)	Synthetic wastewater	0.25–2	49	54 BV/h	100	200	[88]
Zeo. (USA, TR, EU)	Synthetic wastewater	0.5–0.7	4.5	5	100	1400	[89]
SCMZ (Chinese)	Contaminated drinkingwater	0.074	2.5–10	8–12 m/h	90	/47	[59]
Zeolite (USA)	Synthetic wastewater	0.6–2.8	300	2–3	90	/100	[93]
Zeo. (Ukraine)/NaA	Simulated wastewater	0.71–1	27–580	28.8–29 and 31.7–31 BV/h	89.8–97.9	25–217 and 35–272	[96]
Zeo. (Jordanian)	Synthetic wastewater	0.5–0.71	15–50	100–250	90	/1.5;0.4	[90]
Zeolite (Turkey)	Municipal wastewater	0.5–1; 1–2–2–4	20	0.3 m/h	100	/54, 66, and 78	[94]
Zeolite (Ukraine)	Contaminated drinkingwater	0.3–0.6;0.6–1.5	2	5 m/h	70–89; 54–94		[97]
Zeo. (Azarbaijan)	Synthetic wastewater	0.5–2.0	1–50	1–3	99.43		[99]
Zeolite (Romania)	Synthetic wastewater	0.2–0.4	70	3.3	97–100	/0.5–3.7	[91]

SCMZ silicate-carbon modified zeolite (SCMZ); NaA zeolite synthesized from fly ash; t_b breakthrough time; BV_b bed volumes (BV) at breakthrough point

Various researches have been conducted to remove ammonium from landfill leachates. In 2008, Karadağ et al. [6] studied the removal of ammonium ion (NH_4^+) from sanitary landfill leachate (Istanbul, Turkey) using zeolite (clinoptilolite) in both batch and column studies. The competitive effect was observed to be more effective in lower ammonium concentration and higher total cation concentrations. Conditioning and regeneration using NaCl improved removal efficiency of ammonium from leachate. In a similar study, Ye et al. [92] investigated the removal of ammonium from landfill leachate using natural zeolite. In the batch studies, the ammonium removal efficiency increased with an increase in zeolite dosage from 25 to 150 g/L and an increase in stirring speed from 200 to 250 r/min. The optimal pH for the removal of ammonium was found as 7.1. In the column studies, the total ammonium removal percentage during 180 min operation time decreased with the flow rate though the ion-exchange capacity varied to a very small extent with the flow rate ranging from 4 to 9 mL/min.

Ji et al. [60] investigated the removal of ammonium ions from aqueous solution using a modified clinoptilolite- Ca^{2+} -formed clinoptilolite (CaY) prepared from natural clinoptilolite. The batch results show that the CaY has a high selectivity for NH_4^+ and the exchange decreases with increasing temperature. The column results indicated that the effluent of simulated wastewater treated with CaY could meet the integrated wastewater discharge standard of China. Cyrus and Reddy [93] investigated the efficiency of zeolite in the removal of NH_4^+ -N from swine wastewater and desorption of the sorbed NH_4^+ -N from the zeolite saturated with N. Column studies indicated that, columns with flow rates of 2 and 3 mL/min, reached the breakthrough point within about 100 h. In 2013, a three-column filtration system was designed and operated for the investigation of ammonium and suspended solids removal from wastewater by using natural zeolite [94]. The optimum results were obtained as 100 and 38 % ammonium and suspended solids removals, respectively for 0.5 mm diameter during an operation period of 78 h.

19.3.3 Ammonium Removal from Wastewaters in the Fixed Bed Systems

Various researches have been done on zeolite fixed bed column in recent years. Wang et al. [95] studied the removal of ammonia nitrogen with a low concentration (50 mg/L) from aqueous solution in a zeolite column. They found that the wastewater with 50 mg/L ammonia nitrogen took 7 h to flow through a modified zeolite (10 g) column having diameters of 24–64 meshes at a flow rate of 2 mL/min. The saturation extent of adsorption was up to 7.95 mg/g and the saturation adsorption time was 22 h (Table 19.4). Malovany et al. [96] studied the concentration of ammonium from municipal wastewater using ion exchange process. In total 23 runs of saturation and regeneration, synthetic wastewaters of different kinds and pretreated municipal wastewater were used.

Experimental investigation of ammonium ion removal from drinking water was carried out using natural zeolite (clinoptilolite) [97]. The efficiency fluctuated from 89 to 70 % (0.3–0.6 mm particle size) and from 94 to 54 % (0.6–1.5 mm particle size). In 2011, clinoptilolite zeolite was processed, first, by sodium chloride solution and then mixed with Na_2SiO_3 and powdered activated carbon. The mixture went through extrusion forming and calcination to prepare the Silicate-Carbon Modified Zeolite (SCMZ) that was used for removing ammonium from drinking water using column system [98]. The adsorption capacity of SCMZ at a solution pH of 7, filtration rate of 10 m/h and the initial ammonium concentration of 5 mg/L was 0.115 mg NH_4^+ -N/g, which was much higher than that of natural zeolite. SCMZ filtering media would exhibit high ammonium-removal efficiency when regenerated by three cycles with NaCl solution.

Farkas et al. [43] investigated the treatment of leakage waters from waste dumps using activated carbon and natural zeolite clinoptilolite. Leakage waters were characterized by a high content of ammonium (820 mg/L) and organic pollutants (1033 mg/L C). Addition of activated carbon, regardless of its mass, increased the exchange of ammonium ions to only 5.7 mg NH_4^+ /g zeolite. In the model solution having the same ammonium concentration as the real solution, 17.70 mg NH_4^+ /g zeolite was exchanged. The zeolite fixed bed was used to remove ammonium from landfill leachate [16]. Higher effluent volumes and removal rates were produced by lower ammonium concentrations and increased expansion ratios in the fluidized-bed column reduced the treatment efficiency. Combining ammonia stripping and an upflow fixed-bed system is feasible in reducing higher ammonium concentrations to satisfactory levels. In 2012, Khosravi et al. [99] studied the removal of ammonium ion from wastewater using natural Western Azerbaijan zeolite. The maximum removal was achieved at initial ammonium concentration of 30 mg/L, flow rate of 1 mL/min and pH 6.

19.3.4 Ammonium Removal for Wastewater Treatment Systems Combined with Zeolites

The influence of combined systems with zeolites for wastewater treatment was investigated by a number of researchers. Turan et al. [100] studied a combined anaerobic fluidized bed and zeolite fixed bed system in sanitary landfill leachate treatment. While excellent ammonia removal (>90 %) was obtained with the untreated zeolite, the regenerated zeolites showed higher performance. This combined anaerobic and adsorption system is an effective tool to remove high COD and high ammonia in landfill leachate (Table 19.5).

Multifunctional permeable reactive barrier (multibarrier), combining microbial degradation and abiotic ion exchange processes for removal of ammonium from landfill leachate was studied [101]. Feeding the columns with leachate deprived from ammonium resulted in a complete release of the previously sorbed ammonium

Table 19.5 Ammonium removal efficiency for wastewater treatment systems combined with zeolites

Adsorbent	Wastewater	d _p (mm)	C _i (mg/L)	Q (mL/min)	Removal (%)	BV _b (BV)/t _b (h)	pH	References
AFB/Zeo. (Turkey)	Landfill leach./AFB	0.5	410	9.03	90	/480		[100]
PRB/Nitritf/Zeolite	Landfill leachate	1.0–2.0	244–260	0.07 ± 0.02	97.5–96.8	18; 13.5 and 41	7.6	[101]
PRB/Zeo. (Chinese)	Contaminated Groundwater	1–2	0.3–36.6	24 m/h	60	/7200	6.4–8	[102]
BAF/Zeo. (Chinese)	Textile wastewater	2.4–4	8	4	35.1			[103]
C/F-F/Powder zeo.4A	Landfill leachate	0.003	1644	3.94	91	/17	7.74	[104]
BA/Silver-zeolite	Synthetic wastewater	0.002–0.003	5.04		20–37	0.25–2		[105]
NBC/Zeo. (Chilean)	Synthetic wastewater	1	707.1		70	/240–288	7.2	[106]
AnMBRs/Zeo. USA	Wastewater/AnMBR permeate	8 × 40 mesh	31	4–8 BV/h	95	/48–24	7.2	[107]
PRB/Zeo. (Chinese)	Contaminated Groundwater	3–5	1488.7	0.5	70.8	/200	7.9	[108]
CW/Zeo. (Chinese)	Contaminated surfacewater		1.21–3.70	0.4 m ³ /h	83–94	/35.4	8.4	[109]
CW/Zeo. (Chinese)	Synthetic wastewater	0.27–0.45	100–1500		97.1–72.0	/1–72	6	[110]

AFB anaerobic fluidized bed; PRB permeable reactive barrier; BAF biological aerated filter; C/F-F coagulation/flocculation-Fenton; AnMBRs anaerobic membrane bioreactors; CW constructed wetland; BA bactericidal activity; NBC nitrifying biomass carrier

from the clinoptilolite due to exchange with metal cations present in the leachate. Li et al. [102] studied to remove $\text{NH}_4^+\text{-N}$ from contaminated groundwater in a pilot-scale zeolite permeable reactive barrier (PRB). After some 300 days of operation, about 40 % of the influent $\text{NH}_4^+\text{-N}$ was converted to nitrite and nitrate through nitrification; where zeolite contributed to ~60 % of the observed removal of $\text{NH}_4^+\text{-N}$ through adsorption.

Ammonium nitrogen removal characteristics of zeolite media in Biological Aerated Filter (BAF) process for treatment of textile wastewater was studied [103]. Several biofilters were conducted to compare the performance of natural zeolite with sand and granular activated carbon (GAC) as packing media. The removal of $\text{NH}_4^+\text{-N}$ by ion exchange, nitrification and biological synthesis was estimated as 35.1, 40.2 and 22.3 %, respectively. Turki et al. [104] examined the application of an integrated technique that consisted of coagulation/flocculation-Fenton-powder zeolite adsorption methods for the treatment of raw leachates from a landfill site of Bizerte (Northern Tunisia). The removal efficiency of ammonia was 91 %. In 2012, the antimicrobial activity of silver-zeolite against *Escherichia coli*, *Vibrio harveyi*, *Vibrio cholerae* and *Vibrio parahaemolyticus* was examined in liquid medium [105]. Silver-zeolite removed 20–37 % ammonia from aqueous solutions.

A batch nitrification process was studied using synthetic wastewater as substrate and Chilean natural zeolite as biomass carrier at ambient temperatures (20 °C) [106]. Ammonium N removals higher than 70 % were obtained in experimental sets II and III when zeolite was used. The third experimental set (III) revealed that the highest N consumption took place between days 11 and 16 of digestion. Deng et al. [107] found that ammonium exchange of natural zeolite could be an economical method of nitrogen removal from the permeates of anaerobic membrane bioreactors (AnMBRs). A significant bottleneck of zeolite processes would be the requirement of substituting exhausted zeolite with virgin one due to the reductions in ammonium exchange capacity and regeneration efficiency (RE).

The optimum composition of permeable reactive barrier (PRB) materials for remediating groundwater heavily contaminated by landfill leachate was investigated [108]. In column tests, various mixtures of zero-valent iron (ZVI), zeolite (Zeo) and activated carbon (AC) with 0.01–0.25, 3.0–5.0 and 0.7–1.0 mm grain sizes, respectively were used. The optimum weight ratio of the three PRB materials for removing the contaminants and maintaining adequate hydraulic conductivity was found to be 5:1:4.

A strip-like wetland was constructed in the riparian zone for investigation of ammonium nitrogen ($\text{NH}_3\text{-N}$) removal in the Peach River [109]. An oxygen-deficient condition with dissolved oxygen of 0.87–1.60 mg/L, was observed in the zeolite layer, which benefits anaerobic ammonium oxidation (anammox) bacteria survival. The intensified test driven by artificial pumping shows that average removal rates of $\text{NH}_3\text{-N}$ and TN are 41.6 mg/m³d and 63.2 mg/m³d, respectively. In a similar study [110], the kinetic adsorption of ammonium nitrogen ($\text{NH}_4^+\text{-N}$) by CW substrate materials such as blast furnace

slag (BFS), zeolite, ceramsite, vermiculite, gravel, paddy soil, red soil and turf, was investigated using batch experiments and kinetic adsorption isotherms. The maximum adsorption capacity of $\text{NH}_4^+\text{-N}$, estimated as 33.333 g/kg for zeolite.

19.4 Conclusions

In this review, the recent development of nanoporous zeolites in water and wastewater treatment has been studied worldwide for ammonium removal. The properties and modification of natural zeolite are discussed. Modification of natural zeolites can be done in several methods such as hydrothermal pretreatment, NaOH, HCl or NaCl solution treatment, making the modified zeolites achieve higher adsorption capacity for the removal of ammonium ions. Due to uniform micropore structure and high surface area, synthetic zeolite usually has a higher adsorption capability compared with natural zeolite. Therefore, in order to obtain low-cost and effective synthetic zeolite, many researchers have investigated the synthesis of zeolite from coal fly ash, red mud, fly ash, halloysite etc. An overview of nanoporous zeolites for the removal of ammonium is presented in this article along with their adsorption capacities and other parameters and the experimental conditions used in batch systems. Various researches on the zeolite fixed bed columns have been done in recent years. The influence of combined systems with zeolites for wastewater treatment was investigated by a number of researchers. Thus, the use of zeolite as a low-cost adsorbent may contribute to the sustainability of the environment. Undoubtedly nanoporous zeolites offer a lot of promising benefits for commercial purposes in the future.

References

1. M. Rozic, S. Cerjan-Stefanovic, S. Kurajica, V. Vancina, E. Hodzic, *Water Res.* **34**, 3675 (2000)
2. M.S. Celik, B. Ozdemir, M. Turan, I. Koyuncu, G. Atesok, H.Z. Sarikaya, *Water Sci. Technol. Wat. Supply* **1**, 81 (2001)
3. L. Lin, G.Y.S. Chan, B.L. Jiang, C.Y. Lan, *Waste Manage.* **27**, 1376 (2007)
4. A. Thornton, P. Pearce, S.A. Parsons, *Water Res.* **41**, 433 (2007)
5. D. Karadag, Y. Koc, M. Turan, B. Armagan, J. Hazard. Mater. **136**, 604 (2006)
6. D. Karadag, S. Tok, E. Akgul, M. Turan, M. Ozturk, A. Demir, J. Hazard. Mater. **153**, 60 (2008)
7. H. Zhang, X.G. Wu, X.W. Li, *Chem. Eng. J.* **210**, 188 (2012)
8. F.M. Ferraz, J. Povinelli, E.M. Vieira, *Environ. Technol.* **34**, 2317 (2013)
9. J.C. Campos, D. Moura, A.P. Costa, L. Yokoyama, F.V.D. Araujo, M.C. Cammarota, L. Cardillo, J. Environ. Sci. Health A **48**, 1105 (2013)
10. H. Alvarez-Vazquez, B. Jefferson, S.J. Judd, J. Chem. Technol. Biotechnol. **79**, 1043 (2004)
11. J.L. Gao, M. Chys, W. Audenaert, Y.L. He, S.W.H. Van Hulle, *Environ. Technol.* **35**, 1226 (2014)

12. A. Cabeza, A. Urtiaga, M.J. Rivero, I. Ortiz, J. Hazard. Mater. **144**, 715 (2007)
13. H. Zhang, Y.L. Li, X.G. Wu, Y.J. Zhang, D.B. Zhang, Waste Manage. **30**, 2096 (2010)
14. D. Kim, H.D. Ryu, M.S. Kim, J. Kim, S.I. Lee, J. Hazard. Mater. **146**, 81 (2007)
15. T. Zhang, L.L. Ding, H.Q. Ren, J. Hazard. Mater. **166**, 911 (2009)
16. D. Karadag, E. Akkaya, A. Demir, A. Saral, M. Turan, M. Ozturk, Ind. Eng. Chem. Res. **47**, 9552 (2008)
17. A. Corma, Chem. Rev. **95**, 559 (1995)
18. C.J.H. Jacobsen, C. Madsen, T.V.W. Janssens, H.J. Jakobsen, J. Skibsted, Micropor. Mesopor. Mater. **39**, 393 (2000)
19. L. Tosheva, V.P. Valtchev, Chem. Mater. **17**, 2494 (2005)
20. Y. Wang, A. Tuel, Micropor. Mesopor. Mater. **113**, 286 (2008)
21. G.M. Haggerty, R.S. Bowman, Environ. Sci. Technol. **28**, 452 (1994)
22. L.L. Ames, Am. Mineral. **45**, 689 (1960)
23. S.E. Jorgensen, Water Res. **10**, 213 (1975)
24. M. Gaspard, A. Neveu, G. Matin, Water Res. **17**, 279 (1983)
25. I. Rodriguez-Iznaga, A. Gomez, G. Rodriguez-Fuentes, A. Benitez-Aguilar, J. Serrano-Ballan, Micropor. Mesopor. Mater. **53**, 71 (2002)
26. M. Turan, M.S. Celik, J. Water Supply Res. Technol. Aqua **52**, 159 (2003)
27. L.R. Weatherley, N.D. Miladinovic, Water Res. **38**, 4305 (2004)
28. M. Sarioglu, Separ. Purif. Technol. **41**, 1 (2005)
29. V.J. Inglezakis, M.D. Loizidou, H.P. Grigoropoulou, Water Res. **36**, 2784 (2002)
30. E. Erdem, N. Karapinar, R. Donat, J. Colloid Interf. Sci. **280**, 309 (2004)
31. M. Turan, U. Mart, B. Yüksel, M.S. Çelik, Chemosphere **60**, 1487 (2005)
32. M.K. Doula, Water Res. **40**, 3167 (2006)
33. N.V. Medvidovic, J. Peric, M. Trgo et al., Micropor. Mesopor. Mater. **105**, 298 (2007)
34. B. Armagan, O. Ozdemir, M.S. Celik, M. Turan, J. Chem. Technol. Biotechnol. **78**, 725 (2003)
35. O. Ozdemir, B. Armagan, M. Turan, M.S. Celik, Dyes Pigments **62**, 49 (2004)
36. D. Karadag, M. Turan, E. Akgul, S. Tok, A. Faki, J. Chem. Eng. Data **52**, 1615 (2007)
37. D. Karadag, E. Akgul, S. Tok, F. Erturk, M.A. Kaya, M. Turan, J. Chem. Eng. Data **52**, 2436 (2007)
38. J. Torres-Perez, M. Solache-Rios, M.T. Olguin, Separ. Sci. Technol. **42**, 299 (2007)
39. K. Saltali, A. Sari, M. Aydın, J. Hazard. Mater. **141**, 258 (2007)
40. S. Wang, Y. Peng, Chem. Eng. J. **156**, 11 (2010)
41. N. Widiastuti, H.W. Wu, H.M. Ang, D.K. Zhang, Desalination **277**, 15 (2011)
42. Z. Liang, J. Ni, J. Hazard. Mater. **166**, 52 (2009)
43. A. Farkas, M. Rozic, Z. Barbaric-Mikocevic, J. Hazard. Mater. **117**, 25 (2005)
44. I. Langmuir, J. Am. Chem. Soc. **40**, 1361 (1918)
45. H. Freundlich, Z. Phys. Chem. **57**, 384 (1906)
46. S. Lagergren, Kungliga Svenska Vetenskapsakademiens. Handlingar **24**, 1 (1898)
47. G. McKay, Y.S. Ho, Water Res. **33**, 578 (1999)
48. G. McKay, Y.S. Ho, Process Biochem. **34**, 451 (1999)
49. G.C. Catena, F.V. Bright, Anal. Chem. **61**, 905 (1989)
50. C. Namasivayam, R.T. Yamuna, Environ. Pollut. **89**, 1 (1995)
51. H. Nolllet, M. Roels, P. Lutgen, P. Van der Meeren, W. Verstraete, Chemosphere **53**, 655 (2003)
52. L.D. Benefield, J.F. Judkins Jr, B.L. Weand, *Process Chemistry for Water and Wastewater Treatment* (Prentice Hall, Englewood Cliffs, 1982)
53. S. Kundu, A.K. Gupta, J. Colloid Interface Sci. **290**, 52 (2005)
54. G.S. Bohart, E.Q. Adams, J. Am. Chem. Soc. **42**, 523 (1920)
55. R.A. Hutchins, Chem. Eng. **80**, 133 (1973)
56. G.M. Walker, L.R. Weatherley, Water Res. **31**, 2093 (1997)
57. A. Alshameri, A. Ibrahim, A.M. Assabri, X. Lei, H. Wang, C. Yan, Powder Technol. **258**, 20 (2014)

58. L. Lin, Z. Lei, L. Wang, X. Liu, Y. Zhang, C. Wan, D.-J. Lee, J.H. Tay, *Separ. Purif. Technol.* **103**, 15 (2013)
59. M. Li, X. Zhu, F. Zhu, G. Ren, G. Cao, L. Song, *Desalination* **271**, 295 (2011)
60. Z.-Y. Ji, J.-S. Yuan, X.-G. Li, *J. Hazard. Mater.* **141**, 483 (2007)
61. Y. Zhao, B. Zhang, X. Zhang, J. Wang, J. Liu, R. Chen, *J. Hazard. Mater.* **178**, 658 (2010)
62. Y. Zhao, B. Zhang, Y. Zhang, J. Wang, J. Liu, R. Chen, *Separ. Sci. Technol.* **45**, 1066 (2010)
63. M. Zhang, H. Zhang, D. Xu, L. Han, D. Niu, B. Tian, J. Zhang, L. Zhang, W. Wu, *Desalination* **271**, 111 (2011)
64. M. Zhang, H. Zhang, D. Xu, L. Han, D. Niu, L. Zhang, W. Wu, B. Tian, *Desalination* **277**, 46 (2011)
65. D. Wu, B. Zhang, C. Li, Z. Zhang, H. Kong, *J. Colloid Interf. Sci.* **304**, 300 (2006)
66. A.M. Cardoso, M.B. Horn, L.S. Ferret, C.M.N. Azevedo, M. Pires, *J. Hazard. Mater.* **287**, 69 (2015)
67. Y. Zhao, Y. Niu, X. Hu, B. Xi, X. Peng, W. Liu, W. Guan, L. Wang, *Desalin. Water Treat.* **1**, 1 (2015)
68. N.P. Hankins, S. Plankarom, N. Hilal, *Separ. Sci. Technol.* **39**, 3639 (2004)
69. T.C. Jorgensen, L.R. Weatherley, *Water Res.* **37**, 1723 (2003)
70. M. Rozic, S. Cerjan-Stefanovic, L. Curkovic, *Croat. Chem. Acta* **75**, 255 (2002)
71. D. Karadag, Y. Koc, M. Turan, M. Ozturk, *J. Hazard. Mater.* **144**, 432 (2007)
72. A. Gunay, *J. Hazard. Mater.* **148**, 708 (2007)
73. J.B. Dixon, S.B. Weed (eds.), *Mineral Soil Environment* (SSSA Pub. Inc., Madison, WI, 1989), pp. 874–911
74. E. Maranon, M. Ulmanu, Y. Fernandez, I. Anger, L. Castrillon, *J. Hazard. Mater.* **137**, 1402 (2006)
75. G. Moussavi, S. Talebi, M. Farrokhi, R.M. Sabouti, *Chem. Eng. J.* **171**, 1159 (2011)
76. N. Karapinar, *J. Hazard. Mater.* **170**, 1186 (2009)
77. M. Zabochnicka-Świątek, K. Malińska, *Global NEST J.* **12**, 256 (2010)
78. R. Malekian, J. Abedi-Koupai, S.S. Eslamian, S.F. Mousavi, K.C. Abbaspour, M. Afyuni, *Appl. Clay Sci.* **51**, 323 (2011)
79. D. Wen, Y.-S. Ho, X. Tang, *J. Hazard. Mater.* **133**, 252 (2006)
80. O. Lahav, M. Green, *Water SA* **26**, 51 (2000)
81. H.M. Huang, X.M. Xiao, B. Yan, L.P. Yang, *J. Hazard. Mater.* **175**, 247 (2010)
82. V.K. Jha, S. Hayashi, *J. Hazard. Mater.* **169**, 29 (2009)
83. P. Vassileva, D. Voikova, *J. Hazard. Mater.* **170**, 948 (2009)
84. A. Khosravi, M. Esmhosseini, S. Khezri, *Res. Chem. Intermed.* **40**, 2905 (2014)
85. L. Lei, X. Li, X. Zhang, *Separ. Purif. Technol.* **58**, 359 (2008)
86. A.M. Cardoso, *Fuel* **139**, 59 (2015)
87. M.L. Nguyen, C.C. Tanner, *New Zealand J. Agric. Res.* **41**, 427 (1998)
88. S.N. Ashrafzadeh, Z. Khorasani, M. Gorjiara, *Separ. Sci. Technol.* **43**, 960 (2008)
89. J. Langwaldt, *Separ. Sci. Technol.* **43**, 2166 (2008)
90. A. Mashal, J. Abu-Dahrieh, A.A. Ahmed, L. Oyedele, N. Haimour, A. Al-Haj-Ali, D. Rooney, *World J. Sci. Technol. Sustain. Develop.* **11**, 144 (2014)
91. A. Maicaneanu, H. Bedelea, S. Burca, M. Stanca, *Separ. Sci. Technol.* **46**, 1621 (2011)
92. Z. Ye, J. Wang, L. Sun, D. Zhang, H. Zhang, *Environ. Technol.* **1**, 1 (2014)
93. J.S. Cyrus, G.B. Reddy, *J. Environ. Sci. Health A* **46**, 408 (2011)
94. D. Alkas, B. Beler, Baykal, C. Kinaci, *Desalin. Water Treat.* **51**, 3047 (2013)
95. H. Wang, H. Gui, W. Yang, D. Li, W. Tan, M. Yang, C.J. Barrow, *Desalin. Water Treat.* **52**, 753 (2014)
96. A. Malovanyy, H. Sakalova, Y. Yatchyshyn, E. Plaza, M. Malovanyy, *Desalination* **329**, 93 (2013)
97. A. Mažeikiene, M. Valentukevičiene, J. Jankauskas, *J. Environ. Eng. Landscape Manage.* **18**, 54 (2010)
98. M. Li, X. Zhu, F. Zhu, G. Ren, G. Cao, L. Song, *Desalination* **271**, 295 (2011)

99. A. Khosravi, M. Esmhosseini, J. Jalili, S. Khezri, J. Incl. Phenom. Macrocycl. Chem. **74**, 383 (2012)
100. M. Turan, H. Gulsen, M.S. Çelik, J. Environ. Eng. **131**, 815 (2005)
101. T. Van Nooten, L. Diels, L. Bastiaens, Sci. Technol. **44**, 3486 (2010)
102. S. Li, G. Huang, X. Kong, Y. Yang, F. Liu, G. Hou, H. Chen, Water Sci. Technol. **70**, 1540 (2014)
103. W.-S. Hang, H.-T. Tran, D.-H. Park, R.-H. Zhang, D.-H. Ahn, J. Ind. Eng. Chem. **15**, 524 (2009)
104. N. Turki, K. Elghniji, D. Belhaj, J. Bouzid, Desalin. Water Treat. **1**, 1 (2014)
105. K.K. Krishnani, Y. Zhang, L. Xiong, Y. Yan, R. Boopathy, A. Mulchandani, Bioresour. Technol. **117**, 86 (2012)
106. S.J. Montalvo, L.E. Guerrero, R. Borja, Int. J. Environ. Sci. Technol. **11**, 43 (2014)
107. Q. Deng, B.R. Dhar, E. Elbeshbishy, H.-S. Lee, Environ. Technol. **35**, 2008 (2014)
108. D. Zhou, Y. Li, Y. Zhang, C. Zhang, X. Li, Z. Chen, J. Huang, X. Li, G. Flores, M. Kamon, J. Contamin. Hydrol. **168**, 1 (2014)
109. Y. Pei, J. Wang, Z. Wang, B. Tian, J. Environ. Sci. Health A **46**, 715 (2011)
110. W.L. Zhu, L.H. Cui, Y. Ouyang, C.F. Long, X.D. Tang, Pedosphere **21**, 454 (2011)

Chapter 20

Synthesis and Biological Applications of Quantum Dots

Manolya Kukut Hatipoglu, Seda Kelestemur and Mustafa Culha

Abstract One of the fastest moving fields in nanotechnology is the application of quantum dots (QDs) in biology. These light-emitting materials are a new class of fluorescent labels in biomedical field. The superior properties such as resistance to photobleaching, a broad range of excitation and size-tunable light emission facilitate the applications of QDs in nanodiagnostics, imaging and targeted delivery. In addition, new synthesis approaches and modification strategies with several biodegradable polymers and biomolecules bring in water solubility features to QDs, which lead the way to the applications including cellular labelling, deep-tissue imaging and assay labelling. Here we discuss the synthesis, surface modification strategies and in vivo /in vitro biological applications of QDs.

20.1 Introduction

The development of biocompatible nanomaterials is one of the fastest moving fields of nanotechnology due to promising application potentials in nanomedicine. Quantum dots (QDs) are drawing attention among other nanomaterials such as fullerenes, carbon nanotubes, graphenes, noble metal nanoparticles, and metal oxide nanoparticles. The QDs are the fluorophores offering a unmatched promise in the nanomedicine toolbox [1]. The strong light absorbance, bright fluorescence, narrow and symmetric emission bands, broad range of excitation and high photostability are among unique optical properties of QDs [2]. These tiny light-emitting particles have several potential applications including biosensing, bioimaging, bio

M.K. Hatipoglu · S. Kelestemur · M. Culha (✉)
Genetics and Bioengineering Department, Yeditepe University,
Faculty of Engineering and Architecture, Kayisdagi, Istanbul, Turkey
e-mail: mculha@yeditepe.edu.tr; mculha2@gmail.com

S. Kelestemur
Istanbul Medipol University Regenerative and Restorative Medicine Research Center,
Kavacak, Istanbul, Turkey

assay, targeted drug delivery and new therapeutic agent development [3]. The diversity of the applications of QDs in nanomedicine field is attributed to their exceptional photophysical properties [4]. QDs consist of a semiconductor core with an outer shell of another semiconductor material with a larger spectral band gap [1]. They can be synthesized with diameters from a few nanometers to a few micrometers with controlled size distribution [4]. Synthesis method can be tailored to specific requirements with core, shell and coating characteristics that all affect the optical properties of QDs [5]. The size and composition of QDs with numerous surface functionalization possibilities make them amenable materials for medical applications [1]. The utilization of QDs in vitro and in vivo bioimaging for real time and long-term applications are offering exciting opportunities. The conjunction of QDs with biorecognition molecules makes them perfect fluorescent probes for detection, imaging and labelling in vivo and vitro applications. This chapter provides a short review of synthesis, surface modification strategies and biological imaging applications of QDs in vitro and in vivo.

20.2 Synthesis of CdSe and CdSe/ZnS QDs

The synthesis methods for CdSe quantum dots (QDs) have been improved over the past years. Although it is very challenging to control the size and shape of the QDs, and it is now possible with the developed synthesis methods. Nucleation of organometallic precursors in solvent is an efficient approach for synthesizing QDs. The inorganic precipitation was carried out at high temperature in glasses or by using chemical approaches at room temperature [6]. Micellular methods were performed to obtain monodisperse QDs through precipitation of Cd and Se precursors in micelle solutions [7, 8]. The quantum effect of nanocrystals depends on their size distributions and surface reconstructions. The micellular approaches are insufficient in elimination of the intrinsic traps during the synthesis and needed remarkable improvement to obtain narrow sized QDs [6].

Finally, the CdSe QDs with high optical quality and controlled size were synthesized by Bawendi and co-workers using tri-*n*-octylphosphine (TOP) and tri-*n*-octylphosphine oxide (TOPO) as stabilizing agents [9]. The organometallic reagents, dimethylcadmium, bis(trimethylsilyl) selenium and bis(tert-dimethylbutylsilyl) tellurium, were injected rapidly into the reaction solvent to control the nucleation step. The size selective precipitation allowed obtaining monodisperse QDs in different size ranges. Alivisatos and co-workers have developed a synthesis method [10] by increasing the injection temperature during the nucleation step and decreasing the temperature in the growth phase, which affected the size, shape and optical properties of QDs. Also, tributylphosphine was used instead of TOP, which the excess amount can be easily removed at high temperatures. The most important improvement on this method was obtaining nearly monodisperse nanocrystals without a need of a size selective precipitation [11]. In 1998, Peng, Wickham and Alivisatos focused on the study of size distribution of the QDs. Ostwald ripening is the theory that explains the diffusion of dissolved small crystals on

the surface of the large crystals, which are thermodynamically favored according to larger crystals due to their lower surface energy. Alivisatos and co-workers demonstrated the effect of addition monomer reagent in the reaction solvent to obtain narrow size distributed nanocrystals by using Gibbs-Thompson equation [12].

$$S_r = S_b \exp(2\sigma V_m / rRT) \quad (20.1)$$

S_r is the dispersibility of the nanocrystal and S_b is corresponding to the bulk solid; σ is specific surface energy; r is the radius of the nanocrystal; V_m is the molar volume of the materials; R is the gas constant; and T is the temperature.

$$f2\sigma V_m / rRT \ll 1, \quad (20.2)$$

The diffusion controlled growth rate of the particle with size r is

$$dr/dt = K(1/r + 1/\delta)(1/r^* - 1/r) \quad (20.3)$$

K is a constant proportional to the diffusion constant of the monomer. δ is the thickness of the diffusion layer. At a fixed concentration, r^* is the particle radius, which the nanocrystal is stable in suspension with proper monomer concentration. The small nanocrystals can grow faster than the larger particles when the nanocrystals size is larger than the r^* . In the study of Alivisatos et al. it was mentioned that the narrow sized nanocrystals can be obtained through adjusting the concentration of the monomer [11]. Although, nearly monodispersed QDs were synthesized but their quantum yield was low due to the organic compounds on their surfaces, which were used in the synthesis procedure. It was demonstrated that wrapping the surface with an inorganic compound is promising to increase the quantum yield [8]. Hines and Guyot-Sionnest first began working on obtaining monodisperse and highly luminescence quantum yielded QDs through coating the CdSe surface with ZnS, which have a large band gap area increasing the fluorescence properties [13]. The CdSe QDs were synthesized by using the method developed by Bawendi and co-workers. The Zn/S/TOP solution was injected in an optimized period of time at 300 °C. The CdSe/ZnS core shell structures were synthesized at different size ranges by Bawendi et al. [14]. The ZnS capping layer increased the quantum yield up to 50 % was a promising result paved the way for future studies. The main difficulty of the core shell synthesis procedure is achieving a perfect ZnS coating on the surface of CdSe. During the injection, Zn and S atoms tend to grown on ZnS layer composed on CdSe. Injection Cd atoms during the growth phase provide a perfect coverage of the QDs. It was demonstrated that the complete shell coating increase the quantum efficiency of nanocrystals effectively [15].

Important progress has been made in recent years for the synthesis of QDs with high optical properties. Green synthesis methods were developed to obtain non-toxic QDs with a simple method. The synthesis of QDs in aqueous media is more environmentally friendly approach according to organic phase synthesis. Liu and coworkers synthesized QDs at lower injection temperature by providing long

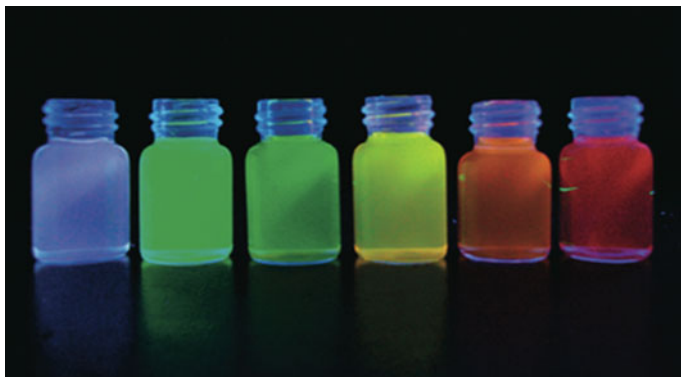


Fig. 20.1 Fluorescence of CdSe QDs synthesized at room temperature with different emission. Reprinted with permission from [14] Copyright (2008) American Chemical Society

reaction times for the nucleation and growth phase of nanocrystals as seen in Fig. 20.1 [16]. The thiol bounded QDs were synthesized with 22 % photoluminescence quantum yield used as an alternative method to organic phase synthesis [17]. Siy and co-workers obtained a narrowly distributed in size and high quality QDs synthesized low reaction temperature compared to the hot injection approaches [18]. Eventually Park and co-workers developed a new approach by using amino acid cysteine and derivatives to synthesize very small sized hydrophilic QDs by controlling size growth. A sol-gel method is effective for the synthesis of monodispersed and properly shaped nanocrystals [19]. The great interest for the green synthesis approaches [20] for QDs favored the development of safety sol-gel synthesis methods by optimizing the reaction parameters [21].

20.3 Design of CdSe and CdSe/ZnS QDs for Biological Applications

The CdSe and CdSe/ZnS QDs were synthesized in organic solvents as described above. The core nanocrystals possess high fluorescence efficiency and stability in organic media and coating the surface with ZnS layer increases the quantum yield and quality. However, the organic ligand capped nanocrystals cannot be used in biological applications due to their hydrophobic behavior. The QDs should be soluble in biological buffers while maintaining their physicochemical properties and size distributions. A surface modification is needed to provide hydrophilic behavior to the QDs. Also, the surface modification prevents the release of Cd^{+2} ions, which have a very toxic effect and limits the biological applications of the QDs [22]. There are several approaches for modification of CdSe or CdSe/ZnS QDs to be used in biological applications. There are several routes to design biocompatible nanocrystals including, ligand exchange, silica coating and conjugation with

amphiphilic molecules, poly ethylene glycol (PEG), phospholipid micelles and biomacromolecules.

20.3.1 *Ligand Exchange Process*

Ligand exchange process is used to exchange the hydrophobic ligands on the surface of QDs with thiol-ended ligands, which bound on the surface through chemisorption. Mercaptoacetic acid [23], mercaptopropanoic acid [24], cystamine [25], cysteine [26] are used with carboxyl($-\text{COOH}$), amine($-\text{NH}_2$) and hydroxyl ($-\text{OH}$) groups on their free edges allowed water solubility and further modifications. Although it was reported the ligand exchange process between the organic ligands on the surface of QDs and thiolalkyl molecules do not increase the size of the QDs, it reduces the quantum yield and the stability [27]. The decrease in the quantum efficiency depends on the redox energy levels of thiol molecules. When the valence band of the nanocrystal is lower than the redox energy of the thiol molecule, the holetrapping is thermodynamically favorable through the thiol molecule. The hole trapping inhibits the radiative recombination of exciton resulted with the decrease of the quantum yield [28].

In biological applications, the CdSe/ZnS QDs are routinely used, which the thiol groups bound onto the ZnS layer through weak disulfide bonds [29]. The thiol molecules bounded on the surface of QDs also form disulfide linkages and causes oxidation and aggregation. To increase the long-term stability, dihydrolipoic acid (DHLA) was used as a ligand exchange molecule. DHLA forms chelate with the metal atoms on the surface of QDs and provide stability only in the basic conditions [30]. Using DHLA and PEG together is another approach to provide usage of the QDs also in acidic conditions [31]. Dithiol-containing molecules such as tridentate [32] and multidentate ligands [33], and also multivalent thiol polymer ligands [34] are used to provide long-term stability [35]. In a recent study, poly acrylamide acid was used as a modifier in ligand exchange approach and provided solubility to the QDs however it quenched the fluorescence [36].

20.3.2 *Silanization Process*

Coating the surface of the CdSe and CdSe/ZnS QDs with a silica layer is another preferred approach. A silane-coupling agent can be used to generate a silica shell coating on the surface, completely. The silica coating provides stability to the QDs at neutral pH and under high salt conditions. Further, the modifications can also be pursued with the selected ligands and polymers. In an aqueous solution, water interacts with the silica layer through hydrogen bonding and this interaction provides a steric repulsion between the nanocrystals, which reduces agglomeration. As an advantage of the silanization, the fluorescence properties of nanocrystals are not

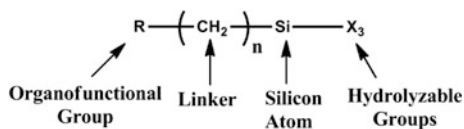


Fig. 20.2 Structure of silane coupling agent

quenched after silica coating process. The first route for silanization process is direct ligand exchange process where generally 3-mercaptopropyl-trimethoxysilane is used [37] to exchange with TOPO or the other hydrophobic ligands on the surface of nanocrystals [38]. The general formula for a typical silane coupling agent bears two classes of functionality as seen in Fig. 20.2. X is a hydrolyzable group, which is generally an alkoxy, acyloxy, halogen or amine. First, the hydrolysis takes place followed by the formation of silanol group which can condense with the other silanol groups to form siloxane linkages [23, 39].

The second route is Stober approach, a sol-gel process, used to coat the surface with a silica layer by using tetraethyl ortosilicate (TEOS) [40]. Stober method is easy to apply and high monodisperse QDs can be obtained [41]. The optimization of the silane coupling agent concentration is important to prevent the aggregation of QDs and if it cannot be achieved, the aggregates may be encapsulated in a silica shell, which increase the size of the structure.

20.3.3 Amphiphilic Molecules

The main purpose of coating the surface of QDs with an amphiphilic molecule is the interaction of the hydrophobic residue of the molecule with the hydrophobic ligands on the surface while the hydrophilic residue provides solubility in aqueous environment. The synthesis procedure is started in an organic solvent and followed by adding aqueous solution after evaporating the organic solvent [42].

Poly(acrylic acid) (PAA), a polymer with carboxylic acid groups, can be used as an amphiphilic molecule for the attachment of an aliphatic amine molecule such as N-octylamine (OA) and 5-amino-1-pentanol [43]. It was also shown that using 3-(dimethylaminopropyl) carbodiimide as a cross-linker for lysine modification on the PAA-OA increase the stability of functionalized CdSe QDs in aqueous environment [44].

Poly(maleic anhydride) copolymers are generally synthesized through copolymerization between maleic anhydride and olefins [45]. Designed maleic anhydride based copolymers such as poly(maleic anhydride-alt-1-tetradecene) [46], poly(maleicanhydride altoctadecene) [47], poly(maleic anhydride-alt-1-decene) [48], were used for the surface functionalization with the aim of employing in biological applications [49].

Block copolymers are used as surface modifiers through forming micellar structure around the QDs. The amphiphilic block copolymers assembled in proper

solvents and formed the micellar nanocrystal structures, which could be used in several biological applications [50]. Using block copolymers as surface modifiers have advantages; first the fluorescence properties of QDs are not quenched because the polymer is not directly interacting with the nanocrystal's core atoms. Second, they are stable in aqueous solutions due to the hydrophobic domains around the nanocrystals [51].

20.3.4 PEG and Phospholipid Micelle

PEG is a polymer used as a surface modifier for QDs, which is employed in several biological applications such as drug delivery [52], imaging [53] and in vivo studies [54]. Non-specific protein binding is prevented by encapsulation the nanocrystal with PEG and also the modified nanocrystals cannot be recognized by the reticuloendothelial system (RES), which decreases their circulation time in the organism [55].

PEG polymer can be used by modification with thiol, amine and carboxyl active groups. As the thiol groups bound on the nanocrystal's surface and PEG chains are responsible to provide solubility in all the acidic, basic and salty conditions of aqueous media, where active amine and carboxyl residues allow further modifications [56].

PEG is also a preferred material for designing amphiphilic block copolymers, where in a recent study, poly(ethylene glycol)-graft-chitosan was synthesized and used to modify CdSe QDs. The reaction scheme was shown in Fig. 20.3. The results

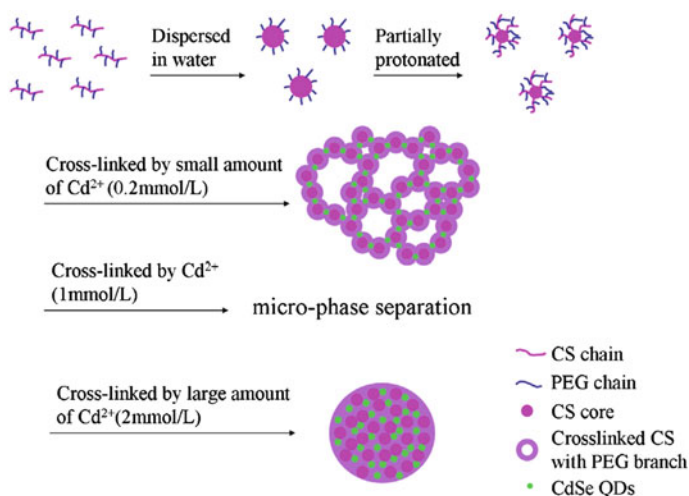


Fig. 20.3 Reaction scheme of modification of CdSe QDs with poly(ethylene glycol)-graft-chitosan [48]. Reprinted from Colloids and Surfaces B-Biointerfaces, Vol. 115, Jiang Z, Zhao C, Liu X, Synthesis of poly(ethylene glycol)-graft-chitosan and using as ligand for fabrication of water-soluble quantum dots, 260–266, Copyright (2014), with permission from Elsevier

were promising that the modification reduces the cellular toxicity of QDs and also allowed using in biological applications through providing water solubility [57].

CdSe and CdSe/ZnS QDs can be encapsulated in phospholipid micelles to use both in vivo and in vitro imaging [58]. The encapsulation approach does not alter the fluorescence and surface properties of the QDs and can be used in drug delivery. The water insoluble anticancer drugs can be delivered by this safe delivery system [59]. When the nanocrystals encapsulated with poly(ethylene glycol)-phospholipid micelle, PEG improve the solubility and circulation time of the QDs and also it can be modified with antibodies to use in targeted delivery [60].

20.3.5 *Biomacromolecules*

CdSe and CdSe/ZnS QDs are modified with biomacromolecules such as peptides, proteins, nucleic acids and polysaccharides to use in biological applications. First, the surface of the QDs is modified with several ligands, which allows for further modification with biomacromolecules through non-covalent interaction and covalent binding.

Non-covalent interactions of QDs with biomacromolecules occur through electrostatic interaction, hydrogen bonding and through metal-coordination. The water-soluble QDs can conjunct with the proteins through electrostatic interactions. The positively charged domains of the proteins can self-assemble on to the negatively charged surface modified QDs [61]. It was demonstrated by Lee et al. that the PEG modified water soluble QDs formed a complex with negatively charged deoxyribose nucleic acid (DNA) with an electrostatic interaction [62]. The strength of hydrogen bonding interaction between QDs and biomolecules depends on the pH and ionic strength of the aqueous media [63]. The association of proteins on the QDs was explained in a study that the hydrogen bonding between carboxyl groups of proteins and amine terminated silica coated QDs is thermodynamically favorable process [64].

The peptide and QDs conjugates are generally used for cellular delivery applications [65]. As another approach, the peptides self-assemble on the QDs' surface through expressed polyhistidine sequences via metal affinity coordination with zinc ions on the surface of CdSe/ZnS QDs [66]. A mixture of different peptides can cover the surface with the self-assembly approach, which have unique functionalities in a single peptide-QDs conjugate system [67]. Although non-covalent interaction is an easy process and not need a cross-linker agent for conjugation but it is difficult to control the binding percentages of modifier on the surface of the nanocrystals. The strength of the interaction is also not very stable in cellular environment, where the free biomolecules have affinity to bind on the surface of the QDs through ligand exchange process. The water soluble QDs containing functional groups ($-\text{NH}_2$, $-\text{COOH}$, $-\text{OH}$ and $-\text{SH}$) can be modified with biomolecules through covalent attachment to provide stability in aqueous media. In general, surface modified QDs with free carboxyl active groups can bind with active amine

groups of proteins, peptides and nucleic acid using a cross-linker 1-ethyl-3-(dimethylaminopropyl) carbodiimide hydrochloride (EDC) [68]. N-hydroxy succinimide (NHS), which has to be used for formation of a stable intermediate product to increase the yield of the reaction, while EDC cannot react with the amine groups rapidly. Biomacromolecules are covalently attached to the quantum dots through amide bond formation using EDC and NHS in two-step process [63, 69]. The reaction scheme is shown in Fig. 20.4.

Thiol groups have affinity for chemisorption onto the CdSe/ZnS QDs, which can allow the direct attachment of biomacromolecules without using a chemical agent. Through this approach the hydrophobic layer on the CdSe/ZnS QDs replace with the thiol containing biomacromolecules, which is called ligand exchange process. Thiol containing peptides, proteins and oligonucleotides are used to provide water-solubility and stability to the QDs [70]. In an aqueous synthesis approach, the QDs are synthesized in water including thiol containing biomacromolecules to obtain non-toxic and practicable QDs in bioimaging and biolabelling [71].

Bioorthogonal click chemistry has pronounced interest in recent years, which the functional groups of biomacromolecules are not affected from the reaction process [72]. The reaction is called as “click”, which is the of Huisgen 1,3-dipolar cycloaddition reaction between azides and terminal alkynes and Cu(I) are used as catalyst [73]. The reaction scheme is shown in Fig. 20.5. Although this approach used as a surface modification strategy for QDs, it quenches the fluorescence property [74]. The development of copper-free click chemistry provides opportunity for conjugation biomacromolecules on the surface of QDs without altering their properties [75].

Carbohydrates play a very important role in biological systems and can covalently attached to a terminal polymer chain of a diblock amphiphilic polymer through Huisgen-type click chemistry as demonstrated in a recent study [76]. The QDs were encapsulated into the glycoconjugated polymer system, which are used for their binding affinities to Concanavalin A with nontoxic behaviors as seen in Fig. 20.6 [76].

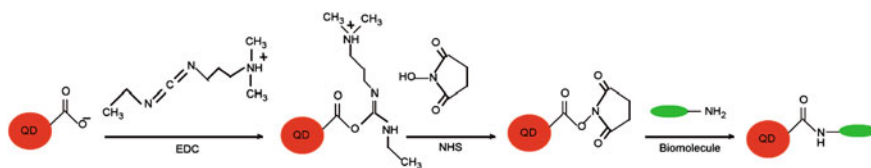


Fig. 20.4 Schematic diagram of the covalent attachment of biomolecule with QDs

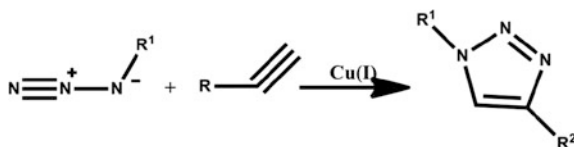
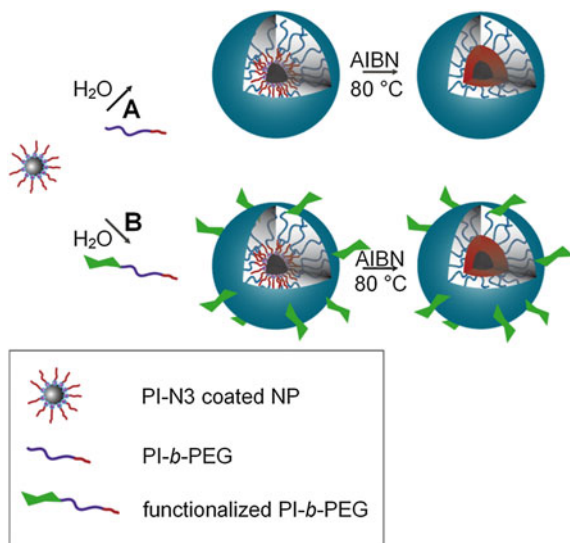


Fig. 20.5 Cu-catalyzed azide-alkyne cycloaddition

Fig. 20.6 Schematic diagram of poly(isoprene)-b-poly(ethylene glycol) diblock copolymer modification of QDs with unmodified (a) and glycoconjugated polymer (b) [64]



20.4 Biological Applications of QDs

The aim of using CdSe and CdSe/ZnS QDs in biological applications is to develop probes with high selectivity, versatility, capacity to penetrate cells and organelles [2, 77]. However, there are certain biological challenges to face such as biocompatibility, non-specific binding and toxicity. The reproducible fabrication is another difficulty to meet. A large amount of work on applications of QDs is devoted for bioconjugation and biological labelling for tracking and imaging [78]. The first application of QDs as label in biology was reported in 1998 [79]. QDs are amazingly appropriate for immunolabeling, cell motility assays, in situ hybridization and live cell markers [80, 81]. The engineered QDs (water-soluble) can be cross-linked to biomolecules such as proteins, peptides, oligonucleotides, and antibodies to target specific biomolecules [82]. The fact that the applications of QDs are limited depending on the administration routes and dispersion within the organism. They tend to aggregate in cytoplasm of living cells and commonly trapped in endosomes and lysosomes. There are varieties of tracking and detection methods of QDs including confocal microscopy, total internal reflection microscopy, wide-field epifluorescence microscopy and fluorometry [83, 84].

QDs have a diverse array of applications in imaging biological samples by a unique property that several QDs can be excited at the same wavelength of light [1, 85, 86]. This superiority acquires a reputation to QDs as optical bar codes of biological screening [87]. These labels can be specific to a particular analyte or a cell type since one wavelength could be directed [88]. QDs can be employed to obtain information location, saturation and distribution of marked cells or proteins [85]. This multiplexing potential is practicable on cancer diagnostics and research

to detect cancer biomarkers [85]. Detection of several proteins on Western blots with QDs opens up a new era in proteomics field [81]. In a study of Makrides et al. showed that streptavidin coated QDs would be used in detection of two proteins in a Western blot assay [89]. Additionally, their stability and resistance to metabolic pathways in live cells allow the usage of QDs in long term studies [90].

20.4.1 *In Vitro Targeting with Antibody Conjugation*

High photostability, resistance to metabolic degradation and lack of cytotoxicity facilitates the utilization of QDs for imaging thick tissues for a long period of time without any damage to the pattern [80].

QDs bioconjugates have been found as a powerful tool to recognize and track plasma membrane agents on living cells [82]. The conjunction of QDs with cancer specific agents such as peptides, ligands and antibodies makes them propitious for cancer detection and imaging. Figure 20.7 shows some of the QD-conjugate structures used in biological applications.

One of the reasons of the myriad of applications of QDs conjugates is the antibody-QD conjugate, which is the simplest way to label if an antibody exists for the target. It is the fastest and easiest route to follow. For instance, hairy-related 2 protein (Her2) receptor, which is a well-known cancer marker over-expressed in many breast cancers, was detected by Wu et al. by the development of streptavidin-coated QDs [90]. The assay employed humanized anti-Her2 antibody and a biotinylated goat anti-human IgG secondary antibody on SK-BR-3 breast cancer cells.

Another similar approach was used to detect intracellular targets including microtubules in mouse 3T3 fibroblasts, nuclear antigens, and F-actin filaments via

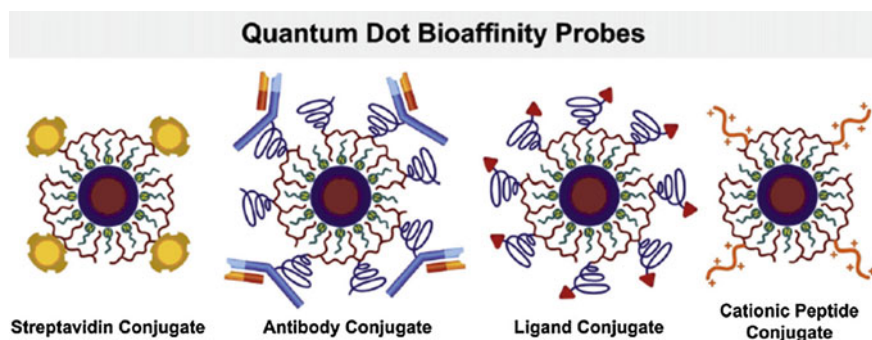


Fig. 20.7 QD-conjugate structures used in biological applications. Reprinted from *Advanced Drug Delivery Reviews*, Vol. 60, Andrew M. Smith, Hongwei Duan, Aaron M. Mohs, Shuming Nie, *Bioconjugated quantum dots for in vivo molecular and cellular imaging*, 1226–1240, Copyright (2008), with permission from Elsevier

streptavidin-coated QDs. A single wavelength was used for the detection on SK-BR-3 cancer cells [84]. MCF-7 and BT-474 breast cancer cells were selectively labelled as an extension of this work by Yezhelyev et al. [91]. They used conjugates of QDs with antibodies for Her2, epidermal growth factor receptor (EGFR), estrogen receptor (ER), progesterone receptor (PR) and mammalian target of rapamycin (m-TOR). The labelled cells were detected by visible and near infrared (NIR). Also, Kawashima et al. employed the same system, the conjugates of QDs with EGF and anti EGFR antibody, for targeting EGFR in human ovarian epidermoid carcinoma cells (A431) [91, 92]. Tada et al. reported that, KPL-4 breast cancer cells were selectively labeled with conjugate of QDs with Herceptin, similar to the previous studies [91, 93]. Also, Weng et al. implemented the labeling system to a multimodal approach by drug delivery of an antibody via immunoliposomes and imaging cells with QDs [94]. They used carbodiimide chemistry between carboxylic acid functionalized QDs and a primary amine group of liposome for coupling. Selective labeling of cancer cells, high-contrast fluorescence imaging and carrying anti-cancer drugs are just a few advantages of multimodal immunoliposomes. In a previous study of Zhang et al. mentioned that, conjugate of QDs with anti-type 1 insulin-like growth factor receptor (IGF1R) was used on up-regulated IGF1R in MCF-7 breast cancer cells for detection and imaging. The results claimed that the QD-anti-IGF1R conjugate is bearing promising candidates for imaging of breast cancer cells [95]. Figure 20.8 shows the fluorescence images of human pancreatic cancer cells (MiaPaCa) treated with InP/ZnS QD-anti-Claudin-4 antibody conjugates, which is a study of Yong et al. [96]. They claimed that conjugates of QDs with anti-Claudin-4 antibody and anti-prostate stem cell antigen (anti-PSCA) selectively detected by human pancreatic cells. The proteins Claudin-4 and PSCA are generally over-expressed in both primary and metastatic pancreatic

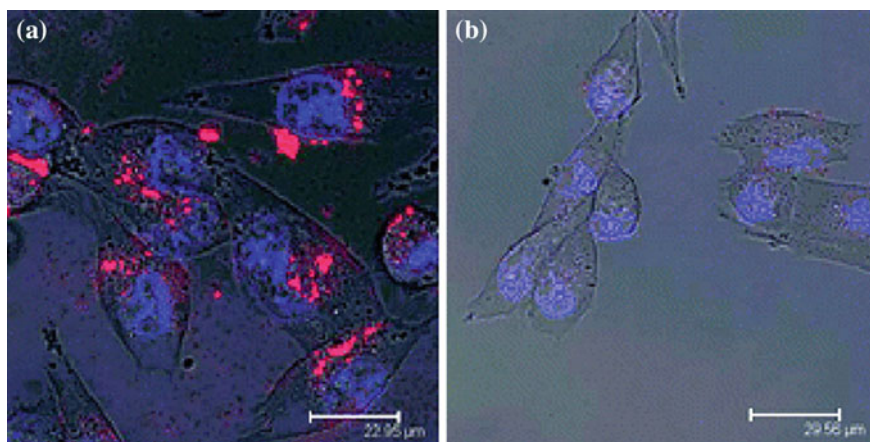


Fig. 20.8 Fluorescence images of human pancreatic cancer cells incubated with **a** InP QD-anti-Claudin-4 antibody conjugate and **b** InP QD without antibody. Reprinted with permission from **a** [94] and **b** [96]. Copyright (2008, 2009) American Chemical Society

cancer cells. Lidke et al. showed that, QD conjugates could be used to distinguish cultured human cancer cells [97]. Coupling CdSe-ZnS QDs with EGF, which is a very small protein specific to erbB/HER membrane receptor, enables the detection of this protein at the single-molecule level with a continuous observation of protein diffusion on the cellular membrane. Gao et al. used prostate specific membrane antigen (PSMA) coupled with QDs by the injection to the tail vines of mice for the detection of prostate cancer xenografts [98]. Detection of human prostate cancer cells (C4-2) proceeded through labelling an antibody for PSMA with QDs.

A conjugate of QDs was developed with three different peptides for molecular imaging of brain and lung endothelial cells and breast carcinoma cells. The studies were carried both in vivo and in vitro by Akerman et al. [99]. The cellular organelles including nucleus and mitochondria can also be monitored with QDs-peptide conjugates [84]. In another study, Larson et al. used water-soluble QDs in an imaging application that indicated the stability of QDs was over 9 months by fluorescence correlation spectroscopy [100]. In the latest study Gao et al. used QD-antibody probes by detecting the staining intensity [101, 102]. They used QDs based multicolor multicycle molecular profiling technology to target LaminA, Cox-4, Ki67 and HSP90 in single cells, which is reported by Gao et al. in 2013 [101, 102].

20.4.2 *In Vitro Targeting with Peptide Conjugation*

A single QD can be conjugated with several proteins and peptides simultaneously depending on its size (see Fig. 20.9) [103, 104]. Medintz et al. showed that approximately 15–20 maltose binding proteins (MBP, 44 kDa) can be conjugated to

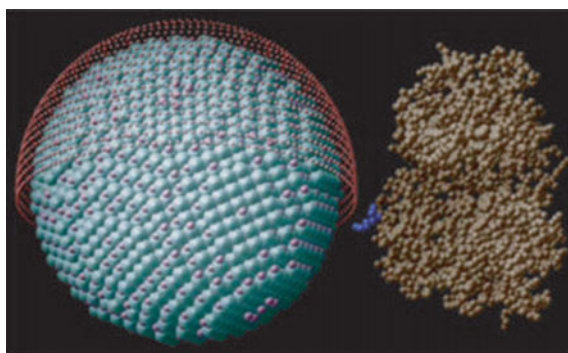


Fig. 20.9 Comparison of QD size to a MBP molecule. 555-nm-emitting CdSe/ZnS core/shell QD, diameter ~ 60 Å, surface-functionalized with dihydrolipoic acid (red shell ~ 9 – 11 Å) has a diameter ~ 78 – 82 Å. The diagram depicts the homogeneous orientation MBP assumes relative to the QD MBP a midsize protein ($M_r \sim 44$ kDa) has dimensions of $30 \times 40 \times 65$ Å [105]. Copyright (2004) National Academy of Sciences, U.S.A

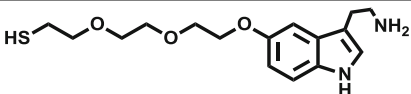
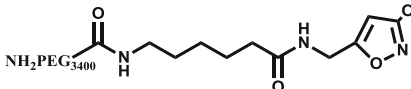
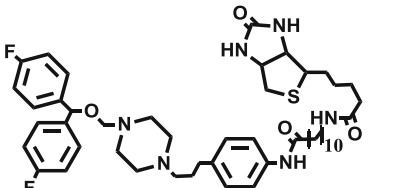
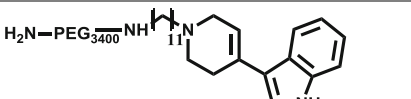
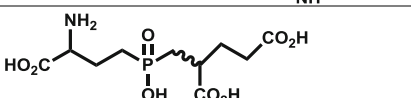
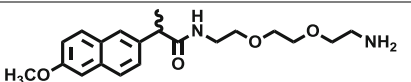
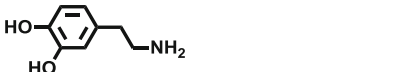
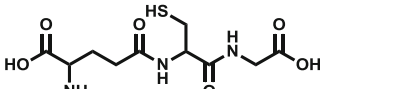
a 6 nm sized QD [103]. Conjunction of proteins and peptides to QDs can be divided into three categories as carbodiimide chemistry, direct binding of thiolated proteins to the QDs and adsorption or non-covalent interaction of engineered proteins. QDs-streptavidin conjugates can be counted as an example of carbodiimide chemistry with high quantum yield, which is also commercially available.

The direct attachment of proteins to the QDs occurs between cysteine residues and the surface of QDs [99, 106]. Non-specific adsorption of proteins to the QDs has handicaps including lack of homogeneity of proteins on the surface of QDs and inability to control protein/QD ratio. Homogeneity is an important issue that, heterogeneous orientation of proteins may result in protein-peptide/QD conjugates without functionality. Peptides are efficient carriers of QDs inside the living cells. There are typical examples of Tat peptides [107], polyarginine [108], cholera toxin B [109] and allatostatin [110] as QD carriers. Curtis and co-workers demonstrated that the conjugates of Tat peptides can be used as vehicles to carry QDs inside human fibroblast cells and nucleus of COS 1 cells [67]. An intercellular delivery of QDs was shown by Waggoner et al. by simply conjugating QDs with poly (L-arginine), where QDs was internalized in HeLa MG63 and 3T3 cells [108]. Another example of internalization is cholera toxin B coated QDs that the conjugates well dispersed in cytoplasm of various cell lines [109]. Additionally, a neuropeptide from *Drosophila melanogaster*, which is called as allatostatin, found as a potential carrier of QDs [110]. Utilization of peptide modified QDs is generally used to target cellular proteins including growth factor receptors, G protein-coupled receptors, integrins and ion channels [111]. Also, there are many examples of QD conjunctions with HIV TAT, PEP-1, polyarginine and SV40 T antigen providing intercellular delivery [98, 112]. Conjunction of arginine-glycine-aspartic acid (RGD, ~30–50) peptides with QDs used to observe avb3 integrins in mouse tumor via NIR [113]. In a recent study, anti-hemagglutinin (HA) conjugated QDs was used to investigate the kinetics of G protein coupled receptor (GPCRs) and endosomal trafficking pathways, where QDs were used as fluorescent probes [114].

20.4.3 In Vitro Targeting with Small Molecule Conjugation

High affinity and selectivity of ligands as small molecules to the cell surface receptors make them propitious candidates as targeting agents [111]. Some examples of small molecules are shown in Table 20.1. The ligands with high affinity lead the way to the protein of interest for QDs. Labeling the binding site of the protein by QDs bring many opportunities as monitoring the protein and helping to clarify the regulation of protein expression and cell surface mobility [111]. Ligand-conjugated QDs may also be utilized in drug discovery that each ligand can be conjugated to different sized QDs, which means different color for a multiplexed fluorescent assay [111]. The major difficulty in ligand-QDs conjugates is tailoring ligand according to the needs such as a linker arm to optimize binding, which requires a good command of organic chemistry.

Table 20.1 Small molecule ligands that have been conjugated to quantum dots

	PEGylated serotonin derivative [116]
	Muscimol derivative [117]
	Dopamine transporter antagonist [118]
	Serotonin transporter antagonist [119]
	GPI a tumor targeting ligand [120]
	A derivative of the NSAID naproxen [121]
	Dopamine [122]
	Glutathione [123]

According to a study in 2007, aptamer labeled QDs could selectively be delivered to prostate cancer cells [115]. The A10 RNA was used as an aptamer to label PSMA positive LNCaP prostate cells. The labeling system was found as successful to distinguish the difference between positive and negative LNCaP prostate cells. In a study of Kawashima et al. they used the ligand-QDs conjugates as a biological probe and found that CHO and A431 cells were efficiently labeled with EGF as a result of specific binding of EGF to EGFR [92]. The results of this study unleash the route for detection of various cancer cells since EGRF is an over-expressed protein in many cancer types.

20.4.4 *In Vivo Applications of QDs*

Many optical techniques have been employed for biomedical imaging using animal models but the poor transmission of visible light through biological tissue prevents its utilizations. Weissleder claimed that, NIR might be applicable for deep tissue imaging [124]. A few of organic dyes can be used in this prospect but photobleaching limits their use for the purpose [125]. Unlike the organic dyes, QDs are ideal materials for imaging in living tissue since their emission wavelengths can be tuned into the desired excitation region of the spectrum. The emission wavelength of the QDs can be adjusted by altering their composition and size for their excitation in the NIR region as photostable fluorophores [126]. The general administration of QDs or any other nanoparticles is systemic intravenous delivery to the blood stream. Depending on to administration route; the interaction of QDs with plasma components, biodistribution in organs and non-specific interaction with cells are important issues to deal with. The biodistribution mechanism starts with the administration of QDs through adsorption by opsonins, which mark them for phagocytosis. The pathways of QDs after administration are shown in Fig. 20.10.

There are several reports indicating the nonspecific uptake of QDs by RES, liver and spleen, which do not reflect the actual behavior of QDs [127–129]. The first in vivo application of QDs was performed by Akermen et al. in 2002 [130]. They administrated the QDs-peptide conjugates through tail vein in mouse and demonstrated the distribution of the QDs in the lung blood vessels. Based on Ballou et al. findings, it can be said that QDs remain in the organs of RES and can be detectable for at least 4 months from initial administration [127–129]. QDs can last for long periods of time without degradation into toxic elements.

Fisher et al. demonstrated that surface modification of QD is holding key for recognition of the particles. According to their study, albumin coated QDs were removed from the circulation and held in the liver after a tail vein injection, which was much faster than unmodified QDs [128]. From a clinical point of view; it is assumed that the accumulation of the QDs can be counted as negligible, as long as they are within the size of renal excretion. Recent publications about in vivo applications of QDs have based on this assumption. Another important issue for targeted imaging is the half-life of contrast agents in blood stream, which is related to biodistribution of the particles. It was found that the surface properties of QDs bearing a great importance in biodistribution as demonstrated in Ballou et al. study on mice [127].

Gao et al. used an antibody conjugate of QDs encapsulated in PEG to track PSMA [98]. The conjugate bound to human prostate xenografts, selectively. They claimed that the signals from QD conjugate were brighter than the green fluorescent protein (GFP), which was used for a similar manner, previously. Dubertret and co-workers followed a parallel approach of the employment of QDs for imaging [131]. They coated QDs with PEG-phosphatidylethanolamine and PEG-phosphatidylcholine, which are micelle-forming block copolymers [131]. In an another example, a triblock copolymer consist of poly(butyl acrylate), poly(ethyl acrylate) and poly

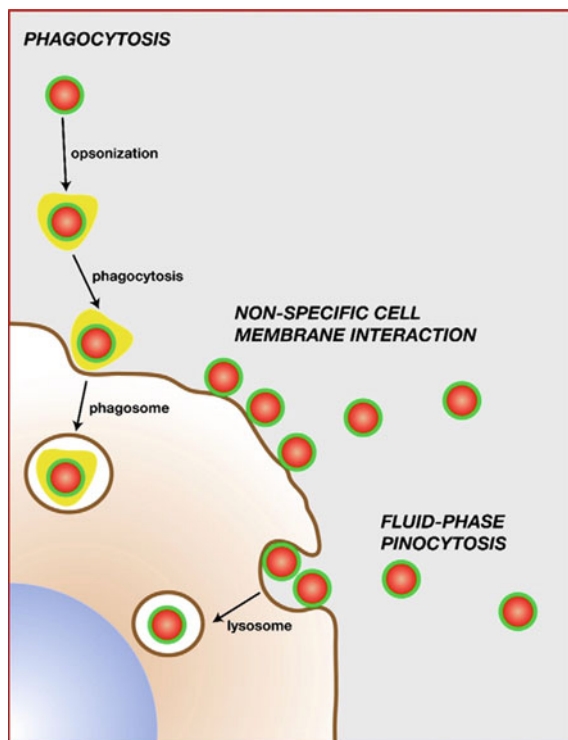
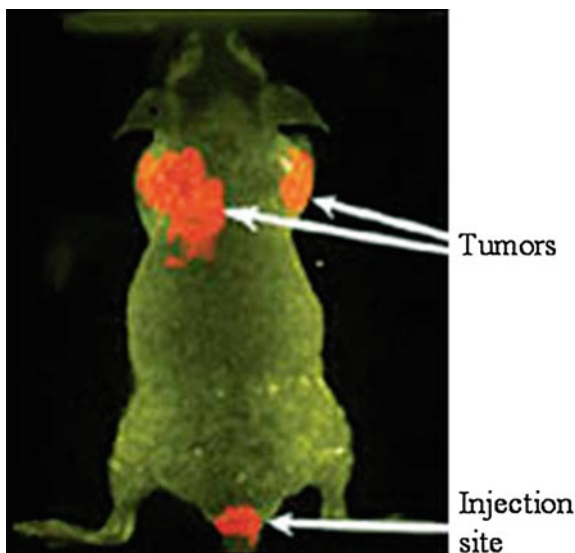


Fig. 20.10 Schematic diagram showing QD interactions with blood immune cells and plasma proteins. The probable modes of interactions including QD opsonization and phagocytosis by leucocytes (e.g., monocytes), nonspecific QD–cell membrane interactions (electrostatic or hydrophobic), and fluid-phase pinocytosis. Reprinted from *Advanced Drug Delivery Reviews*, Vol. 60, Andrew M. Smith, Hongwei Duan, Aaron M. Mohs, Shuming Nie, *Bioconjugated quantum dots for in vivo molecular and cellular imaging*, 1226–1240, Copyright (2008), with permission from Elsevier

(methacrylate) was used to modify the surface of QDs. The modified QDs were conjugated with PSMA antibody and they were able to observe prostate cancer successfully in nude mice [98]. The triblock copolymer acted as reactive surface for conjunction of PSMA antibody. Figure 20.11 shows the localized QD–antibody conjugates on tumor sites of a mouse.

Another developing area of in vivo applications of QDs is imaging of inflammation related to variety of diseases such as diabetes, atherosclerosis, colitis and asthma [132]. Many different cell types and surface proteins take place in inflammation process. A study on investigation of inflammation in retinal endothelium of streptozotocin-treated diabetic rats showed that, different color-coded QDs could be detected and identification of increased retinal vasculature was possible on diabetic animals [133]. The conjugates of various QDs with monoclonal antibodies specific to leukocytes and neutrophils were used for the imaging.

Fig. 20.11 In vivo fluorescence image of tumor bearing mice using QD probes with PEG-PSMA antibody conjugates
Reprinted by permission from Mcmillan Publishers Ltd: [Nature Biotechnology], [98], Copyright (2004)



20.4.5 In Vivo Vascular Imaging

The initial successful in vivo applications of QDs were accomplished on the cardiovascular and lymphatic systems of mammals that it was a non-specific imaging of tumor vasculature. It was shown that the intravenously injected QDs might be employed as marker for fluorescent detection in capillaries of adipose tissue and skin of living mouse [100]. This study triggered the future investigations on non-specific in vivo imaging of tumor vasculatures, lymph nodes and lymphatic drainage by bioconjugated QDs. For instance, PEG-phosphatidylethanolamine modified QDs were used to target and imaging of tumor vasculature related with MCaIV isogenic mouse adenocarcinoma tumor implants in C3H mice by Stroh et al. [134]. In general, the NIR was used to excite the QDs for deep tissue penetration, which cannot be done with organic dyes [135]. The excitation with NIR was used in many studies for imaging including the coronary vasculature of a rat heart and the blood vessels of chicken embryos [136].

The size of the imaging probe is more important in lymphatic system. The first attempt to use QDs for fluorescence imaging of lymph nodes was reported by Frangioni et al. that they used oligomeric phosphine coated CdTe-CdSe QDs on mice and pig [126]. They were able to detect the signals of QDs from 1 cm beneath the skin pre-surgically and even under bloody surgical conditions. As the hydrodynamic volume of probes gets smaller, the penetration of imaging agents into lymphatic system increases. If the QDs has overall hydrodynamic volume around 9 nm, they might able to label five nodes [137]. This technique would help the surgeons to identify lymph nodes by fluorescence and excise nodes draining from metastatic tumors since the lymph nodes are one of the most important prognostic signs of

cancers. Additionally, there are examples of the utilization of the QD-assisted mapping on identification of lymph nodes downstream from the lungs [129, 138], esophagus [139] and from subcutaneous tumors [139]. QDs could be fluorescently observed up to five different nodes by the multiplexing capability [140].

In a recent study, anti-hemagglutinin (HA) conjugated QDs was used to investigate the kinetics of G protein coupled receptor (GPCRs) endosomal trafficking pathways, where QDs were used as fluorescent probes [114]. The latest study is reported by Gao et al. in 2013 about QDs based multicolor multicycle molecular profiling technology to target LaminA, Cox-4, Ki67 and HSP90 in single cells [102]. They used QD-antibody probes by detecting the staining intensity. The latest study on QDs was on modifying QDs with RGD peptides as NIR emitting probes for specific tumor targeting by Wang et al. [141]. The tumor was clearly detected after 30 min. of administration of QDs. A distinct red fluorescence was observed after 2.5 h of injection with a maximum tumor to background ratio [142].

20.4.6 Förster Resonance Energy Transfer (FRET) Based Applications

Förster resonance energy transfer (FRET) theory was first put forward by Theodor Förster in 1948 [143]. FRET is allowed an energy transfer from a donor to an acceptor through a dipol-dipol interaction, which is used to identify a protein-protein interaction, conformational changes in biological molecules and ligand receptor binding in the range of 1–20 nm [144]. In general, fluorophores such as fluorescent dyes, proteins and polymers are used as energy donors in FRET technique. The QDs are unique fluorophores with their broad absorption and narrow emission profiles, high excitation coefficients and brightness and also large surface area allows using as a biosensor through conjugation with proper biomolecules [145]. Also, QDs can be synthesized at different sizes, which can prevent the spectral overlapping between donor and acceptor. All these advantages increase the efficiency of FRET technique and make QDs preferable fluorophores for imaging [103].

Despite the superior optical properties, QDs are not appropriate materials as acceptors depending on the large extinction coefficients and long excited-state lifetimes [146]. In previous studies, it was shown that QDs are strong FRET donors to organic dye acceptors in multiple applications [146, 147]. Algar and Krull showed that, the non-radiative energy transfer between QD donors and organic dyes or fluorescent proteins (FP) are in parallel with the dipole-dipole interaction of Förster theory [66, 146].

The most important point to consider before choosing a QD-dye or QD-FP pair is the spatial proximity between two fluorophores that has a great impact on energy transfer efficiency. The distance between donor and the acceptor called as Förster distance (R_0), which can be determined by the photophysical properties of the donor and the acceptor molecules (J.R. Lakowicz, Principles of fluorescence spectroscopy, 3rd ed. Springer, New York, USA (1999)). The quantum yield of the

QD will be towering in the absence of acceptor, for a FRET pair, where the QD used as a donor. The QDs were found as strong donors when paired with the acceptor molecules with a strong spectral overlap [148]. Another issue while choosing the QDs to employ in FRET is the surface coating of the QDs. A study in 2008 showed that the modifiers on the surface of the QDs affect the energy transfer to FPS. The small ligand coatings, which allow shorter distances between QDs and FPS, result in higher energy transfer in vitro assays [148, 149]. In general conjugates of QDs with biological molecules are employed in FRET for immunoassay studies. For instance, McKinney et al. used the QDs conjugates to monitor the protein interactions in Holiday Junction [150]. DNA undergoes a conformational change in the presence of Mg^{2+} ions during recombination phase [151]. They monitored the change in the emission by the movement of arm due to Mg^{2+} ions where QDs are used as donor and Cy5 used as the acceptor on the two arms of DNA [151]. Another study on enzyme assays pointed out the importance of QDs for FRET experiments [103].

The QDs can be used to develop as biosensors in different applications using FRET technique. *Escherichia coli* maltose-binding protein conjugated QDs were used as a carbohydrate biosensor by Medintz and co-workers [103]. Detection of nucleic acids through QDs FRET approach is also an interested area to develop. A streptavidin conjugated CdSe/ZnS QDs was used as an energy donor for sensitive detection of Cy5 labelled DNA by Zang et al. in 2005. Detection was occurred with hybridization of biotinylated oligonucleotide with Cy5 labelled DNA as seen schematically in Fig. 20.12 [152]. The FRET efficiency found maximum of

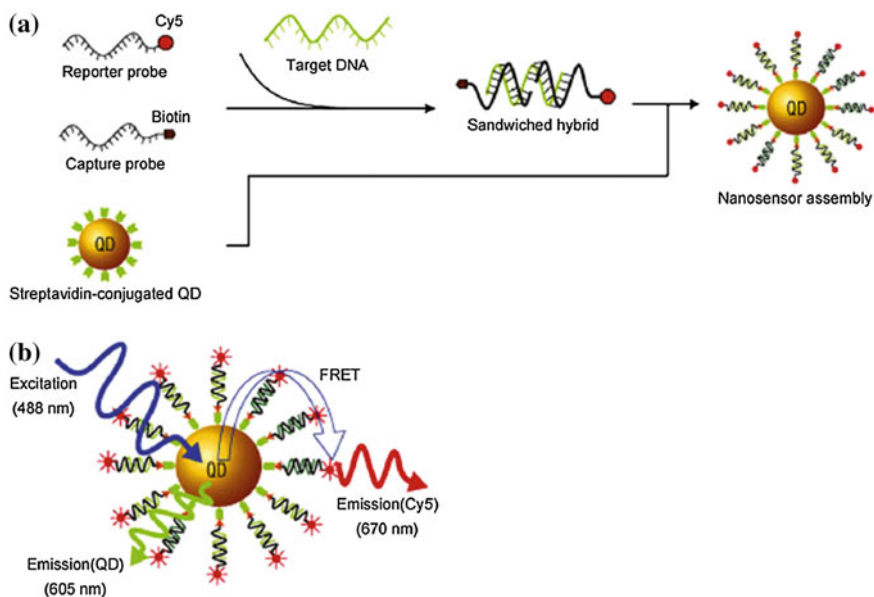


Fig. 20.12 Design of QD based nanosensor (a) and fluorescence emission from Cy5 using QD as an energy donor (b). Reprinted with permits from [6] Copyright (2005) Nature Publishing Group

54-fold greater than when single molecules are bound [153]. The main limitation of using QDs in FRET applications is the increased size of the QDs due to the surface modification process to improve water-solubility. Using thiolated molecules is a good strategy to prevent the increasing of the size of the QDs. The conjugate of QDs with thiolated fluorophore labelled DNA was used as an energy donor with high efficiency [63]. In a study of Zhou and co-workers, tri(ethylene glycol) was used as a linker for coating the QDs surface with DNA, which prevent the non-specific interactions between QDs and DNA [68]. The differentiation of the FRET interactions becomes difficult in case of non-specific bindings [154].

There are various examples of QDs in FRET applications on gene technology. A high number of studies demonstrated that the conjugates of QDs with oligonucleotide sequences bound through surface carboxylic groups may be utilized to bind DNA or mRNA [155]. For instance, Patolsky et al. demonstrated that, QDs could be used to determine the kinetics of telomerisation and DNA replication via FRET [156]. The strength of the quenching is affected by the interparticle distance, which was shown by the gold conjugated DNA with QDs [157]. The percentages of quenching were found as 79 [157] and 85 [158] depending on the short interparticle distances addition to non-radiative interactions. Gueroi et al. were also claimed that the interaction of QD-DNA complexes with clean oligonucleotides, which were not bearing any Au, affect the emission yield, significantly [157].

Besides the DNA technology, QDs has noticeable applications in RNA technology. The QDs are generally employed in the detection of mRNA through in situ hybridization (ISH) and in RNA interference applications with siRNA [103]. The ISH techniques are used in the detection of expression of specific mRNA transcripts in mouse mid brain, where the success of QDs have already proven [159]. The results of QDs are much better than the most sensitive organic fluorophore on labelling of four different mRNA transcripts in neurons. Combination of immunohistochemistry with ISH techniques stand to visualize the localization of growth hormone and prolactin proteins related to their mRNA [103].

The QDs are also utilized in RNA interference applications to determine the extent of gene knockdown in a cell through the difference in brightness, which became an important tool to monitor the gene functions [160].

20.5 Conclusions

In this chapter, some of the recent advancements in the synthesis methods, bio-conjunction strategies and in vivo/in vitro biological applications of QDs are summarized. The synthesis of QDs and understanding their optical properties open a new era in biomedical detection and imaging. The development of new synthesis and modification approaches increased the fluorescence properties of QDs and their compatibility with biological systems.

QDs are already proved itself with their size-tunable and stable photoluminescence, semiconductivity, biocompatibility and availability in multicolor features. The

multiplex detection of biomolecules to multicolor whole animal imaging drew much attention to QDs. They are brighter than the dye molecules and the large surface area provides the opportunity to attach multiple cancer markers, which make them promising materials for multimodal imaging in vivo and in vitro. Surface chemistry, polymer chemistry and bioconjugate chemistry were successfully employed for the modification of QDs up to date. The utilization of QDs to in vivo and in vitro targeting of prostate, breast, pancreatic and liver cancers was demonstrated through QD-antibody, QD-peptide and QD-small molecules bioconjugates. The deep-tissue imaging applications of QDs with NIR are another aspect to be considered as improvement in biomedical field. In vitro/in vivo detection and imaging studies with cancer markers and other reporter molecules are already established in animal models.

The spectral characteristics of QDs provide a unique opportunity to utilize them in FRET applications as donor fluorophores. QDs have gained their place on gene technology, cell tracking, pathogens and toxin detections as FRET donors. The future of the QDs as FRET donors depend on the QD-FRET innovative approaches. It is expected that, the surface modifications and the developments in synthesis will reduce the negative effects of QDs in spite of their complex nature.

The recent advancements on bioconjugated QDs clearly revealed that the QDs are powerful tools on cancer diagnosis and prognosis. However, there is still need for improvement on issues of QDs probes such as specificity of targeting, signaling intensity, multimodality and therapeutic potentials. The next challenge of utilization of QDs into biomedical field is the commercialization of the products. Although there are some commercially available products of QDs for research purposes, they have a long way to go before they can hit the market.

Acknowledgments The authors acknowledge Yeditepe University and TÜBİTAK for their financial support.

References

1. H.M.E. Azzazy, M.M.H. Mansour, S.C. Kazmierczak, From diagnostics to therapy: prospects of quantum dots. *Clin. Biochem.* **40**, 917–927 (2007)
2. P. Juzenas, W. Chen, Y.-P. Sun, MaN Coelho, R. Generalov, N. Generalova, I.L. Christensen, Quantum dots and nanoparticles for photodynamic and radiation therapies of cancer. *Adv. Drug Deliv. Rev.* **60**, 1600–1614 (2008)
3. Q. Huo, A perspective on bioconjugated nanoparticles and quantum dots. *Colloids Surf., B* **59**, 1–10 (2007)
4. T. Jamieson, R. Bakhshi, D. Petrova, R. Pocock, M. Imani, A.M. Seifalian, Biological applications of quantum dots. *Biomaterials* **28**, 4717–4732 (2007)
5. S. Santra, K. Wang, R. Tapeç, W. Tan, Development of novel dye-doped silica nanoparticles for biomarker application. *J. Biomed. Opt.* **6**, 160 (2001)
6. L. Brus, Quantum crystallites and nonlinear optics. *Appl. Phys. Mater. Sci. Process.* **53**, 465–474 (1991)

7. M.G. Bawendi, A.R. Kortan, M.L. Steigerwald, L.E. Brus, X-ray structural characterization of larger cdse semiconductor clusters. *J. Chem. Phys.* **91**, 7282–7290 (1989)
8. A.R. Kortan, R. Hull, R.L. Opila, M.G. Bawendi, M.L. Steigerwald, P.J. Carroll, L.E. Brus, Nucleation and growth of CdSe on ZnS quantum crystallite seeds, and vice versa, in inverse micelle media. *J. Am. Chem. Soc.* **112**, 1327–1332 (1990)
9. C.B. Murray, D.J. Norris, M.G. Bawendi, Synthesis and characterization of nearly monodisperse cde (e = s, se, te) semiconductor nanocrystallites. *J. Am. Chem. Soc.* **115**, 8706–8715 (1993)
10. J.E.B. Katari, V.L. Colvin, A.P. Alivisatos, X-ray photoelectron-spectroscopy of cdse nanocrystals with applications to studies of the nanocrystal surface. *J. Phys. Chem.* **98**, 4109–4117 (1994)
11. X.G. Peng, J. Wickham, A.P. Alivisatos, Kinetics of ii-vi and iii-v colloidal semiconductor nanocrystal growth: “Focusing” of size distributions. *J. Am. Chem. Soc.* **120**, 5343–5344 (1998)
12. T. Sugimoto, Preparation of monodispersed colloidal particles. *Adv. Colloid Interface Sci.* **28**, 65–108 (1987)
13. M.A. Hines, P. Guyot-Sionnest, Synthesis and characterization of strongly luminescing ZnS-capped cdse nanocrystals. *J. Phys. Chem.* **100**, 468–471 (1996)
14. B.O. Dabbousi, J. Rodriguezviejo, F.V. Mikulec, J.R. Heine, H. Mattoussi, R. Ober, K.F. Jensen, M.G. Bawendi, (CdSe)/ZnS core-shell quantum dots: synthesis and characterization of a size series of highly luminescent nanocrystallites. *J. Phys. Chem. B* **101**, 9463–9475 (1997)
15. J. McBride, J. Treadway, L.C. Feldman, S.J. Pennycook, S.J. Rosenthal, Structural basis for near unity quantum yield core/shell nanostructures. *Nano Lett.* **6**, 1496–1501 (2006)
16. L. Liu, Q. Peng, Y. Li, Preparation of cdse quantum dots with full color emission based on a room temperature injection technique. *Inorg. Chem.* **47**, 5022–5028 (2008)
17. X. Chen, J.L. Hutchison, P.J. Dobson, G. Wakefield, Highly luminescent monodisperse cdse nanoparticles synthesized in aqueous solution. *J. Mater. Sci.* **44**, 285–292 (2009)
18. J.T. Siy, E.M. Brauser, M.H. Bartl, Low-temperature synthesis of cdse nanocrystal quantum dots. *Chem. Commun.* **47**, 364–366 (2011)
19. I.U. Arachchige, S.L. Brock, Sol-gel assembly of cdse nanoparticles to form porous aerogel networks. *J. Am. Chem. Soc.* **128**, 7964–7971 (2006)
20. M.R. Gaeeni, M. Tohidian, M. Mapes-Ara, Green synthesis of cdse colloidal nanocrystals with strong green emission by the sol-gel method. *Ind. Eng. Chem. Res.* **53**, 7598–7603 (2014)
21. I.U. Arachchige, S.L. Brock, Sol-gel methods for the assembly of metal chalcogenide quantum dots. *Acc. Chem. Res.* **40**, 801–809 (2007)
22. A.M. Derfus, W.C.W. Chan, S.N. Bhatia, Probing the cytotoxicity of semiconductor quantum dots. *Nano Lett.* **4**, 11–18 (2004)
23. W.C.W. Chan, S.M. Nie, Quantum dot bioconjugates for ultrasensitive nonisotopic detection. *Science* **281**, 2016–2018 (1998)
24. C. Kirchner, T. Liedl, S. Kudara, T. Pellegrino, A.M. Javier, H.E. Gaub, S. Stolzle, N. Fertig, W.J. Parak, Cytotoxicity of colloidal CdSe and CdSe/ZnS nanoparticles. *Nano Lett.* **5**, 331–338 (2005)
25. U. Resch-Genger, M. Grabolle, S. Cavaliere-Jaricot, R. Nitschke, T. Nann, Quantum dots versus organic dyes as fluorescent labels. *Nat. Methods* **5**, 763–775 (2008)
26. W. Liu, H.S. Choi, J.P. Zimmer, E. Tanaka, J.V. Frangioni, M. Bawendi, Compact cysteine-coated cdse(zncds) quantum dots for in vivo applications. *J. Am. Chem. Soc.* **129**, 14530–14543 (2007)
27. A.S. Karakoti, R. Shukla, R. Shanker, S. Singh, Surface functionalization of quantum dots for biological applications. *Adv. Colloid Interface Sci.* **215**, 28–45 (2015)
28. S.F. Wuister, C.D. Donega, A. Meijerink, Influence of thiol capping on the exciton luminescence and decay kinetics of CdTe and CdSe quantum. *J. Phys. Chem. B* **108**, 17393–17397 (2004)

29. R. Pejchal, M.L. Ludwig, Cobalamin-independent methionine synthase (metE): a face-to-face double barrel that evolved by gene duplication. *PLoS Biol.* **3**, 254–265 (2005)
30. Y. Zhang, A. Clapp, Overview of stabilizing ligands for biocompatible quantum dot nanocrystals. *Sensors* **11**, 11036–11055 (2011)
31. M.H. Stewart, K. Susumu, B.C. Mei, I.L. Medintz, J.B. Delehanty, J.B. Blanco-Canosa, P.E. Dawson, H. Mattoussi, Multidentate poly(ethylene glycol) ligands provide colloidal stability to semiconductor and metallic nanocrystals in extreme conditions. *J. Am. Chem. Soc.* **132**, 9804–9813 (2010)
32. A.T. Nagaraja, A. Sooresh, K.E. Meissner, M.J. Mcshane, Processing and characterization of stable, ph-sensitive layer-by-layer modified colloidal quantum dots. *ACS Nano* **7**, 6194–6202 (2013)
33. E. Gravel, C. Tanguy, E. Cassette, T. Pons, F. Knittel, N. Bernards, A. Garofalakis, F. Duconge, B. Dubertret, E. Doris, Compact tridentate ligands for enhanced aqueous stability of quantum dots and in vivo imaging. *Chem. Sci.* **4**, 411–417 (2013)
34. L. Liu, X. Guo, Y. Li, X. Zhong, Bifunctional multidentate ligand modified highly stable water-soluble quantum dots. *Inorg. Chem.* **49**, 3768–3775 (2010)
35. H. Duan, M. Kuang, Y.A. Wang, Quantum dots with multivalent and compact polymer coatings for efficient fluorescence resonance energy transfer and self-assembled biotagging. *Chem. Mater.* **22**, 4372–4378 (2010)
36. G.R. Bardajee, R. Soleyman, Surface passivation of cdse-topo quantum dots by poly(acrylic acid): solvent sensitivity and photo-induced emission in water. *Iran. Polym. J.* **22**, 885–890 (2013)
37. M.A. Correa-Duarte, M. Giersig, L.M. Liz-Marzan, Stabilization of cds semiconductor nanoparticles against photodegradation by a silica coating procedure. *Chem. Phys. Lett.* **286**, 497–501 (1998)
38. M.A. Correa-Duarte, Y. Kobayashi, R.A. Caruso, L.M. Liz-Marzan, Photodegradation of sio2-coated cds nanoparticles within silica gels. *J. Nanosci. Nanotechnol.* **1**, 95–99 (2001)
39. A.G. Kanaras, F.S. Kamounah, K. Schaumburg, C.J. Kiely, M. Brust, Thioalkylated tetraethylene glycol: a new ligand for water soluble monolayer protected gold clusters. *Chem. Commun.* 2294–2295 (2002)
40. W. Stober, A. Fink, E. Bohn, Controlled growth of monodisperse silica spheres in micron size range. *J. Colloid Interface Sci.* **26**, 62 (1968)
41. M. Darbandi, R. Thomann, T. Nann, Single quantum dots in silica spheres by microemulsion synthesis. *Chem. Mater.* **17**, 5720–5725 (2005)
42. R.E. Anderson, W.C.W. Chan, Systematic investigation of preparing biocompatible, single, and small ZnS-capped CdSe quantum dots with amphiphilic polymers. *ACS Nano* **2**, 1341–1352 (2008)
43. R.C. Somers, P.T. Snee, M.G. Bawendi, D.G. Nocera, Energy transfer of CdSe/ZnS nanocrystals encapsulated with rhodamine-dye functionalized poly(acrylic acid). *J. Photochem. Photobiol. Chem.* **248**, 24–29 (2012)
44. C. Luccardini, C. Tribet, F. Vial, V. Marchi-Artzner, M. Dahan, Size, charge, and interactions with giant lipid vesicles of quantum dots coated with an amphiphilic macromolecule. *Langmuir* **22**, 2304–2310 (2006)
45. W. Regel, C. Schneider, Poly(maleic anhydride) - synthesis and proof of structure. *Macromol. Chem. Phys. Makromol. Chem.* **182**, 237–242 (1981)
46. T. Pellegrino, L. Manna, S. Kudera, T. Liedl, D. Koktysh, A.L. Rogach, S. Keller, J. Radler, G. Natile, W.J. Parak, Hydrophobic nanocrystals coated with an amphiphilic polymer shell: a general route to water soluble nanocrystals. *Nano Lett.* **4**, 703–707 (2004)
47. R. Di Corato, A. Quarta, P. Piacenza, A. Ragusa, A. Figuerola, R. Buonsanti, R. Cingolani, L. Manna, T. Pellegrino, Water solubilization of hydrophobic nanocrystals by means of poly (maleic anhydride-alt-1-octadecene). *J. Mater. Chem.* **18**, 1991–1996 (2008)
48. L. Qi, X. Gao, Quantum dot-amphipol nanocomplex for intracellular delivery and real-time imaging of sima. *ACS Nano* **2**, 1403–1410 (2008)

49. N. Tomczak, R. Liu, J.G. Vancso, Polymer-coated quantum dots. *Nanoscale* **5**, 12018–12032 (2013)
50. X. Nie, J. Xu, J. Cui, B. Yang, W. Jiang, Encapsulation of semiconductor quantum dots into the central cores of block copolymer cylindrical and toroidal micelles. *RSC Adv.* **3**, 24625–24633 (2013)
51. Y.-K. Lee, S.M. Hong, J.S. Kim, J.H. Im, H.S. Min, E. Subramanyam, K.M. Huh, Encapsulation of CdSe/ZnS quantum dots in poly(ethylene glycol)-poly(D, L-lactide) micelle for biomedical imaging and detection. *Macromol. Res.* **15**, 330–336 (2007)
52. S. Ghaderi, B. Ramesh, A.M. Seifalian, Fluorescence nanoparticles “quantum dots” as drug delivery system and their toxicity: a review. *J. Drug Target.* **19**, 475–486 (2011)
53. J. Dimitrijevic, L. Krapf, C. Wolter, C. Schmidtke, J.-P. Merkl, T. Jochum, A. Kornowski, A. Schueth, A. Gebert, G. Huettmann, T. Vossmeier, H. Weller, CdSe/cds-quantum rods: fluorescent probes for in vivo two-photon laser scanning microscopy. *Nanoscale* **6**, 10413–10422 (2014)
54. V. Kulvietis, E. Zuraszkas, R. Rotomskis, Distribution of polyethylene glycol coated quantum dots in mice skin. *Exp. Dermatol.* **22**, 157–159 (2013)
55. E.L. Bentzen, I.D. Tomlinson, J. Mason, P. Gresch, M.R. Warnement, D. Wright, E. Sanders-Bush, R. Blakely, S.J. Rosenthal, Surface modification to reduce nonspecific binding of quantum dots in live cell assays. *Bioconjug. Chem.* **16**, 1488–1494 (2005)
56. I. Yildiz, E. Deniz, B. McCaughan, S.F. Cruickshank, J.F. Callan, F.M. Raymo, Hydrophilic CdSe-ZnS core-shell quantum dots with reactive functional groups on their surface. *Langmuir* **26**, 11503–11511 (2010)
57. Z. Jiang, C. Zhao, X. Liu, Synthesis of poly(ethylene glycol)-graft-chitosan and using as ligand for fabrication of water-soluble quantum dots. *Colloids Surf., B* **115**, 260–266 (2014)
58. A. Papagiannaros, T. Levchenko, W. Hartner, D. Mongayt, V. Torchilin, Quantum dots encapsulated in phospholipid micelles for imaging and quantification of tumors in the near-infrared region. *Nanomed. Nanotechnol. Biol. Med.* **5**, 216–224 (2009)
59. V.P. Torchilin, Micellar nanocarriers: pharmaceutical perspectives. *Pharm. Res.* **24**, 1–16 (2007)
60. F. Erogbogbo, K.-T. Yong, R. Hu, W.-C. Law, H. Ding, C.-W. Chang, P.N. Prasad, M.T. Swihart, Biocompatible magnetofluorescent probes: luminescent silicon quantum dots coupled with superparamagnetic iron(III) oxide. *ACS Nano* **4**, 5131–5138 (2010)
61. J. Pathak, K. Rawat, S. Sanwlani, H.B. Bohidar, Interaction of globular plasma proteins with water-soluble cdse quantum dots. *ChemPhysChem* **16**, 1777–1786 (2015)
62. J. Lee, Y. Choi, J. Kim, E. Park, R. Song, Positively charged compact quantum dot-DNA complexes for detection of nucleic acids. *ChemPhysChem* **10**, 806–811 (2009)
63. W.R. Algar, U.J. Krull, Adsorption and hybridization of oligonucleotides on mercaptoacetic acid-capped CdSe/ZnS quantum dots and quantum dot-oligonucleotide conjugates. *Langmuir* **22**, 11346–11352 (2006)
64. P. Liu, N. Na, L. Huang, D. He, C. Huang, J. Ouyang, The application of amine-terminated silicon quantum dots on the imaging of human serum proteins after polyacrylamide gel electrophoresis (page). *Chem. Eur. J.* **18**, 1438–1443 (2012)
65. S.M. Rozenzhak, M.P. Kadakia, T.M. Caserta, T.R. Westbrook, M.O. Stone, R.R. Naik, Cellular internalization and targeting of semiconductor quantum dots. *Chem. Commun.* 2217–2219 (2005)
66. I.L. Medintz, S.A. Trammell, H. Mattoussi, J.M. Mauro, Reversible modulation of quantum dot photoluminescence using a protein-bound photochromic fluorescence resonance energy transfer acceptor. *J. Am. Chem. Soc.* **126**, 30–31 (2004)
67. J.B. Delehanty, I.L. Medintz, T. Pons, F.M. Brunel, P.E. Dawson, H. Mattoussi, Self-assembled quantum dot-peptide bioconjugates for selective intracellular delivery. *Bioconjug. Chem.* **17**, 920–927 (2006)
68. D. Zhou, L. Ying, X. Hong, E.A. Hall, C. Abell, D. Klenerman, A compact functional quantum dot-DNA conjugate: preparation, hybridization, and specific label-free DNA detection. *Langmuir* **24**, 1659–1664 (2008)

69. G.H.T. Au, L. Mejias, V.K. Swami, A.D. Brooks, W.Y. Shih, W.-H. Shih, Quantitative assessment of tn antigen in breast tissue micro-arrays using cdse aqueous quantum dots. *Biomaterials* **35**, 2971–2980 (2014)
70. W. Cai, X. Chen, Preparation of peptide-conjugated quantum dots for tumor vasculature-targeted imaging. *Nat. Protoc.* **3**, 89–96 (2008)
71. A. Samanta, Z. Deng, Y. Liu, Aqueous synthesis of glutathione-capped CdTe/CdS/ZnS and CdTe/CdSe/ZnS core/shell/shell nanocrystal heterostructures. *Langmuir* **28**, 8205–8215 (2012)
72. R. Bilan, F. Fleury, I. Nabiey, A. Sukhanova, Quantum dot surface chemistry and functionalization for cell targeting and imaging. *Bioconjug. Chem.* **26**, 609–624 (2015)
73. A. Bernardin, A. Cazet, L. Guyon, P. Delannoy, F. Vinet, D. Bonnaffe, I. Texier, Copper-free click chemistry for highly luminescent quantum dot conjugates: application to in vivo metabolic imaging. *Bioconjug. Chem.* **21**, 583–588 (2010)
74. W.H. Binder, R. Sachsenhofer, C.J. Straif, R. Zirbs, Surface-modified nanoparticles via thermal and cu(i)-mediated “click” chemistry: generation of luminescent CdSe nanoparticles with polar ligands guiding supramolecular recognition. *J. Mater. Chem.* **17**, 2125–2132 (2007)
75. M.B. Meza, Bead-based HTS applications in drug discovery. *Drug Discov. Today* 38–41 (2000)
76. C. Schmidtke, A.-M. Kreuziger, D. Alpers, A. Jacobsen, Y. Leshch, R. Eggers, H. Kloust, T. Huang, J. Ostermann, T. Schotten, J. Thiem, J. Thimm, H. Weller, Glycoconjugated amphiphilic polymers via click-chemistry for the encapsulation of quantum dots. *Langmuir* **29**, 12593–12600 (2013)
77. M. J. t, Quantum dots finally come of age. *Nat. Biotechnol.* **21**, 32–33 (2003)
78. A.M. Smith, X.H. Gao, S.M. Nie, Quantum dot nanocrystals for in vivo molecular and cellular imaging. *Photochem. Photobiol.* **80**, 377–385 (2004)
79. P. Alivisatos, The use of nanocrystals in biological detection. *Nat. Biotechnol.* **22**, 47–52 (2003)
80. A.P. Alivisatos, W.W. Gu, C. Larabell, Quantum dots as cellular probes, in *Quantum Dots as Cellular Probes* (Annual Reviews, Palo Alto. 2005), pp. 55–76
81. M.P. Bruchez, Turning all the lights on: quantum dots in cellular assays. *Curr. Opin. Chem. Biol.* **9**, 533–537 (2005)
82. A. Smith, H. Duan, A. Mohs, S. Nie, Bioconjugated quantum dots for in vivo molecular and cellular imaging. *Adv. Drug Deliv. Rev.* **60**, 1226–1240 (2008)
83. K.K. Jain, Nanotechnology in clinical laboratory diagnostics. *Clin. Chim. Acta* **358**, 37–54 (2005)
84. X. Michalet, F.F. Pinaud, L.A. Bentolila, J.M. Tsay, S. Doose, J.J. Li, G. Sundaresan, A.M. Wu, S.S. Gambhir, S. Weiss, Quantum dots for live cells, in vivo imaging, and diagnostics. *Science* **307**, 538–544 (2005)
85. A.M. Smith, S. Dave, S.M. Nie, L. True, X.H. Gao, Multicolor quantum dots for molecular diagnostics of cancer. *Expert Rev. Mol. Diagn.* **6**, 231–244 (2006)
86. W.W. Yu, E. Chang, R. Drezek, V.L. Colvin, Water-soluble quantum dots for biomedical applications. *Biochem. Biophys. Res. Commun.* **348**, 781–786 (2006)
87. K.K. Jain, Nanodiagnostics: application of nanotechnology in molecular diagnostics. *Expert Rev. Mol. Diagn.* **3**, 153–161 (2003)
88. M. Bruchez, M. Moronne, P. Gin, S. Weiss, A.P. Alivisatos, Semiconductor nanocrystals as fluorescent biological labels. *Science* **281**, 2013–2016 (1998)
89. S.C. Makrides, C. Gasbarro, J.M. Bello, Bioconjugation of quantum dot luminescent probes for western blot analysis. *Biotechniques* **39**, 501–506 (2005)
90. J.K. Jaiswal, S.M. Simon, Potentials and pitfalls of fluorescent quantum dots for biological imaging. *Trends Cell Biol.* **14**, 497–504 (2004)
91. V. Biju, S. Mundayoor, R.V. Omkumar, A. Anas, M. Ishikawa, Bioconjugated quantum dots for cancer research: present status, prospects and remaining issues. *Biotechnol. Adv.* **28**, 199–213 (2010)

92. N. Kawashima, K. Nakayama, K. Itoh, T. Itoh, M. Ishikawa, V. Biju, Reversible dimerization of egfr revealed by single-molecule fluorescence imaging using quantum dots. *Chem. Eur. J.* **16**, 1186–1192 (2010)
93. H. Tada, H. Higuchi, T.M. Wanatabe, N. Ohuchi, In vivo real-time tracking of single quantum dots conjugated with monoclonal anti-her2 antibody in tumors of mice. *Cancer Res.* **67**, 1138–1144 (2007)
94. K.C. Weng, C.O. Noble, B. Papahadjopoulos-Sternberg, F.F. Chen, D.C. Drummond, D.B. Kirpotin, D. Wang, Y.K. Hom, B. Hann, J.W. Park, Targeted tumor cell internalization and imaging of multifunctional quantum dot-conjugated immunoliposomes in vitro and in vivo. *Nano Lett.* **8**, 2851–2857 (2008)
95. H. Zhang, D. Sachdev, C. Wang, A. Hubel, M. Gaillard-Kelly, D. Yee, Detection and downregulation of type i igf receptor expression by antibody-conjugated quantum dots in breast cancer cells. *Breast Cancer Res. Treat.* **114**, 277–285 (2009)
96. K.-T. Yong, H. Ding, I. Roy, W.-C. Law, E.J. Bergey, A. Maitra, P.N. Prasad, Imaging pancreatic cancer using bioconjugated inq quantum dots. *ACS Nano* **3**, 502–510 (2009)
97. D.S. Lidke, P. Nagy, R. Heintzmann, D.J. Arndt-Jovin, J.N. Post, H.E. Grecco, E.A. Jares-Erijman, T.M. Jovin, Quantum dot ligands provide new insights into erbb/her receptor-mediated signal transduction. *Nat. Biotechnol.* **22**, 198–203 (2004)
98. X.H. Gao, Y.Y. Cui, R.M. Levenson, L.W.K. Chung, S.M. Nie, In vivo cancer targeting and imaging with semiconductor quantum dots. *Nat. Biotechnol.* **22**, 969–976 (2004)
99. M.E. Akerman, W.C.W. Chan, P. Laakkonen, S.N. Bhatia, E. Ruoslahti, Nanocrystal targeting in vivo. *Proc. Nat. Acad. Sci. U.S.A.* **99**, 12617–12621 (2002)
100. D.R. Larson, W.R. Zipfel, R.M. Williams, S.W. Clark, M.P. Bruchez, F.W. Wise, W.W. Webb, Water-soluble quantum dots for multiphoton fluorescence imaging in vivo. *Science* **300**, 1434–1436 (2003)
101. X. Ji, F. Peng, Y. Zhong, Y. Su, Y. He, Fluorescent quantum dots: synthesis, biomedical optical imaging, and biosafety assessment. *Colloids Surf., B* **124**, 132–139 (2014)
102. P. Zrazhevskiy, X.H. Gao, Quantum dot imaging platform for single-cell molecular profiling. *Nat. Commun.* **4**, 12 (2013)
103. I.L. Medintz, A.R. Clapp, H. Mattoussi, E.R. Goldman, B. Fisher, J.M. Mauro, Self-assembled nanoscale biosensors based on quantum dot fret donors. *Nat. Mater.* **2**, 630–638 (2003)
104. L. Igor, H.T.U. Medintz, Ellen R. Goldman, Hedi Mattoussi, Quantum dot bioconjugates for imaging, labelling and sensing. *Nat. Mater.* **4**, 435–446 (2005)
105. I.L. Medintz, J.H. Konnert, A.R. Clapp, I. Stanish, M.E. Twigg, H. Mattoussi, J.M. Mauro, J. R. Deschamps, A fluorescence resonance energy transfer-derived structure of a quantum dot-protein bioconjugate nanoassembly. *Proc. Nat. Acad. Sci. U.S.A.* **101**, 9612–9617 (2004)
106. F. Pinaud, D. King, H.P. Moore, S. Weiss, Bioactivation and cell targeting of semiconductor CdSe/ZnS nanocrystals with phytochelatin-related peptides. *J. Am. Chem. Soc.* **126**, 6115–6123 (2004)
107. J.M. De La Fuente, M. Fandel, C.C. Berry, M. Riehle, L. Cronin, G. Aitchison, A.S.G. Curtis, Quantum dots protected with tiopronin: a new fluorescence system for cell-biology studies. *ChemBiochem* **6**, 989–991 (2005)
108. B.C. Lagerholm, M.M. Wang, L.A. Ernst, D.H. Ly, H.J. Liu, M.P. Bruchez, A.S. Waggoner, Multicolor coding of cells with cationic peptide coated quantum dots. *Nano Lett.* **4**, 2019–2022 (2004)
109. S.K. Chakraborty, JaJ Fitzpatrick, J.A. Phillippi, S. Andreko, A.S. Waggoner, M.P. Bruchez, B. Ballou, Cholera toxin b conjugated quantum dots for live cell labeling. *Nano Lett.* **7**, 2618–2626 (2007)
110. V. Biju, D. Muraleedharan, K. Nakayama, Y. Shinohara, T. Itoh, Y. Baba, M. Ishikawa, Quantum dot-insect neuropeptide conjugates for fluorescence imaging, transfection, and nucleus targeting of living cells. *Langmuir* **23**, 10254–10261 (2007)
111. S.J. Rosenthal, J.C. Chang, O. Kovtun, J.R. McBride, I.D. Tomlinson, Biocompatible quantum dots for biological applications. *Chem. Biol.* **18**, 10–24 (2011)

112. G. Ruan, A. Agrawal, A.I. Marcus, S. Nie, Imaging and tracking of tat peptide-conjugated quantum dots in living cells: new insights into nanoparticle uptake, intracellular transport, and vesicle shedding. *J. Am. Chem. Soc.* **129**, 14759–14766 (2007)
113. B.R. Smith, Z. Cheng, A. De, A.L. Koh, R. Sinclair, S.S. Gambhir, Real-time intravital imaging of rgd-quantum dot binding to luminal endothelium in mouse tumor neovasculature. *Nano Lett.* **8**, 2599–2606 (2008)
114. K.M. Fichter, M. Flajolet, P. Greengard, T.Q. Vu, Kinetics of g-protein-coupled receptor endosomal trafficking pathways revealed by single quantum dots. *Proc. Nat. Acad. Sci. U.S.A.* **107**, 18658–18663 (2010)
115. V. Bagalkot, L. Zhang, E. Levy-Nissenbaum, S. Jon, P.W. Kantoff, R. Langer, O.C. Farokhzad, Quantum dot—aptamer conjugates for synchronous cancer imaging, therapy, and sensing of drug delivery based on bi-fluorescence resonance energy transfer. *Nano Lett.* **7**, 3065–3070 (2007)
116. S.J. Rosenthal, A. Tomlinson, E.M. Adkins, S. Schroeter, S. Adams, L. Swafford, J. McBride, Y.Q. Wang, L.J. Defelice, R.D. Blakely, Targeting cell surface receptors with ligand-conjugated nanocrystals. *J. Am. Chem. Soc.* **124**, 4586–4594 (2002)
117. H.A. Gussin, I.D. Tomlinson, D.M. Little, M.R. Warnement, H. Qian, S.J. Rosenthal, D.R. Pepperberg, Binding of muscimol-conjugated quantum dots to gaba(c) receptors. *J. Am. Chem. Soc.* **128**, 15701–15713 (2006)
118. I.D. Tomlinson, J.N. Mason, R.D. Blakely, S.J. Rosenthal, High affinity inhibitors of the dopamine transporter (dat): novel biotinylated ligands for conjugation to quantum dots. *Bioorg. Med. Chem. Lett.* **16**, 4664–4667 (2006)
119. I.D. Tomlinson, M.R. Warnerment, J.N. Mason, M.J. Vergne, D.M. Hercules, R.D. Blakely, S.J. Rosenthal, Synthesis and characterization of a pegylated derivative of 3-(1,2,3,6-tetrahydro-pyridin-4-yl)-1 h-indole (idt199): a high affinity sert ligand for conjugation to quantum dots. *Bioorg. Med. Chem. Lett.* **17**, 5656–5660 (2007)
120. H.S. Choi, W. Liu, F. Liu, K. Nasr, P. Misra, M.G. Bawendi, J.V. Frangioni, Design considerations for tumour-targeted nanoparticles. *Nat. Nanotechnol.* **5**, 42–47 (2010)
121. S.J. Byrne, B. Le Bon, S.A. Corr, M. Stefanko, C. O’connor, Y.K. Gun’ko, Y.P. Rakovich, J. F. Donegan, Y. Williams, Y. Volkov, P. Evans, Synthesis, characterisation, and biological studies of CdTe quantum dot-naproxen conjugates. *Chemmedchem* **2**, 183–190 (2007)
122. S.J. Clarke, C.A. Hollmann, Z.J. Zhang, D. Suffern, S.E. Bradforth, N.M. Dimitrijevic, W.G. Minarik, J.L. Nadeau, Photophysics of dopamine-modified quantumdots and effects on biological systems. *Nat. Mater.* **5**, 409–417 (2006)
123. C. Tortiglione, A. Quarta, A. Tino, L. Manna, R. Cingolani, T. Pellegrino, Synthesis and biological assay of gsh functionalized fluorescent quantum dots for staining hydra vulgaris. *Bioconjug. Chem.* **18**, 829–835 (2007)
124. R. Weissleder, A clearer vision for in vivo imaging. *Nat. Biotechnol.* **19**, 316–317 (2001)
125. J.V. Frangioni, In vivo near-infrared fluorescence imaging. *Curr. Opin. Chem. Biol.* **7**, 626–634 (2003)
126. S. Kim, Y.T. Lim, E.G. Soltesz, A.M. De Grand, J. Lee, A. Nakayama, J.A. Parker, T. Mihaljevic, R.G. Laurence, D.M. Dor, L.H. Cohn, M.G. Bawendi, J.V. Frangioni, Near-infrared fluorescent type ii quantum dots for sentinel lymph node mapping. *Nat. Biotechnol.* **22**, 93–97 (2004)
127. B. Ballou, B.C. Lagerholm, L.A. Ernst, M.P. Bruchez, A.S. Waggoner, Noninvasive imaging of quantum dots in mice. *Bioconjug. Chem.* **15**, 79–86 (2004)
128. H.C. Fischer, L. Liu, K.S. Pang, W.C.W. Chan, Pharmacokinetics of nanoscale quantum dots: in vivo distribution, sequestration, and clearance in the rat. *Adv. Funct. Mater.* **16**, 1299–1305 (2006)
129. E.G. Soltesz, S. Kim, R.G. Laurence, A.M. Degrand, C.P. Parungo, D.M. Dor, L.H. Cohn, M.G. Bawendi, J.V. Frangioni, T. Mihaljevic, Intraoperative sentinel lymph node mapping of the lung using near-infrared fluorescent quantum dots. *Ann. Thorac. Surg.* **79**, 269–277 (2005)

130. M.E. Akerman, W.C.W. Chan, P. Laakkonen, S.N. Bhatia, E. Ruoslahti, Nanocrystal targeting in vivo. *Proc. Nat. Acad. Sci.* **99**, 12617–12621 (2002)
131. B. Dubertret, P. Skourides, D.J. Norris, V. Noireaux, A.H. Brivanlou, A. Libchaber, In vivo imaging of quantum dots encapsulated in phospholipid micelles. *Science* **298**, 1759–1762 (2002)
132. A. Karwa, E. Papazoglou, K. Pourrezaei, S. Tyagi, S. Murthy, Imaging biomarkers of inflammation in situ with functionalized quantum dots in the dextran sodium sulfate (dss) model of mouse colitis. *Inflamm. Res.* **56**, 502–510 (2007)
133. A. Jayagopal, P.K. Russ, F.R. Haselton, Surface engineering of quantum dots for in vivo vascular imaging. *Bioconjug. Chem.* **18**, 1424–1433 (2007)
134. M. Stroh, J.P. Zimmer, D.G. Duda, T.S. Levchenko, K.S. Cohen, E.B. Brown, D.T. Scadden, V.P. Torchilin, M.G. Bawendi, D. Fukumura, R.K. Jain, Quantum dots spectrally distinguish multiple species within the tumor milieu in vivo. *Nat. Med.* **11**, 678–682 (2005)
135. M. Rubart, Two-photon microscopy of cells and tissue. *Circ. Res.* **95**, 1154–1166 (2004)
136. J.D. Smith, G.W. Fisher, A.S. Waggoner, P.G. Campbell, The use of quantum dots for analysis of chick cam vasculature. *Microvasc. Res.* **73**, 75–83 (2007)
137. X.G. Peng, M.C. Schlamp, A.V. Kadavanich, A.P. Alivisatos, Epitaxial growth of highly luminescent cdse/cds core/shell nanocrystals with photostability and electronic accessibility. *J. Am. Chem. Soc.* **119**, 7019–7029 (1997)
138. C.P. Parungo, Y.L. Colson, S.W. Kim, S. Kim, L.H. Cohn, M.G. Bawendi, J.V. Frangioni, Sentinel lymph node mapping of the pleural space. *Chest* **127**, 1799–1804 (2005)
139. B. Ballou, L.A. Ernst, S. Andreko, T. Harper, JaJ Fitzpatrick, A.S. Waggoner, M.P. Bruchez, Sentinel lymph node imaging using quantum dots in mouse tumor models. *Bioconjug. Chem.* **18**, 389–396 (2007)
140. H. Kobayashi, Y. Hama, Y. Koyama, T. Barrett, C.a S. Regino, Y. Urano, P.L. Choyke, Simultaneous multicolor imaging of five different lymphatic basins using quantum dots. *Nano Lett.* **7**, 1711–1716 (2007)
141. X. Ji, F. Peng, Y. Zhong, Y. Su, Y. He, Fluorescent quantum dots: synthesis, biomedical optical imaging, and biosafety assessment. *Colloids Surf., B* **124**, 132–139 (2014)
142. J. Wang, Y. Lu, F. Peng, Y. Zhong, Y. Zhou, X. Jiang, Y. Su, Y. He, Photostable water-dispersible nir-emitting CdTe/CdS/ZnS core-shell-shell quantum dots for high-resolution tumor targeting. *Biomaterials* **34**, 9509–9518 (2013)
143. T. Forster, Zwischenmolekulare energiewanderung und fluoreszenz. *Annalen Der Physik* **2**, 55–75 (1948)
144. A. Miyawaki, Visualization of the spatial and temporal dynamics of intracellular signaling. *Dev. Cell* **4**, 295–305 (2003)
145. E.A. Jares-Erijman, T.M. Jovin, FRET imaging. *Nat. Biotechnol.* **21**, 1387–1395 (2003)
146. W.R. Algar, U.J. Krull, Quantum dots as donors in fluorescence resonance energy transfer for the bioanalysis of nucleic acids, proteins, and other biological molecules. *Anal. Bioanal. Chem.* **391**, 1609–1618 (2008)
147. N. Mcgrath, M. Barroso, Quantum dots as fluorescence resonance energy transfer donors in cells. *J. Biomed. Opt.* **13**–21 (2008)
148. A.M. Dennis, D.C. Sotito, B.C. Mei, I.L. Medintz, H. Mattoussi, G. Bao, Surface ligand effects on metal-affinity coordination to quantum dots: implications for nanoprobe self-assembly. *Bioconjug. Chem.* **21**, 1160–1170 (2010)
149. A.M. Dennis, G. Bao, Quantum dot-fluorescent protein pairs as novel fluorescence resonance energy transfer probes. *Nano Lett.* **8**, 1439–1445 (2008)
150. S.A. McKinney, A.C. Declais, D.M.J. Lilley, T. Ha, Structural dynamics of individual holliday junctions. *Nat. Struct. Biol.* **10**, 93–97 (2003)
151. S. Hohng, T. Ha, Single-molecule quantum-dot fluorescence resonance energy transfer. *Chemphyschem* **6**, 956–960 (2005)
152. C.Y. Zhang, H.C. Yeh, M.T. Kuroki, T.H. Wang, Single-quantum-dot-based DNA nanosensor. *Nat. Mater.* **4**, 826–831 (2005)

153. A.R. Clapp, I.L. Medintz, J.M. Mauro, B.R. Fisher, M.G. Bawendi, H. Mattoussi, Fluorescence resonance energy transfer between quantum dot donors and dye-labeled protein acceptors. *J. Am. Chem. Soc.* **126**, 301–310 (2004)
154. I.L. Medintz, H. Mattoussi, Quantum dot-based resonance energy transfer and its growing application in biology. *Phys. Chem. Chem. Phys.* **11**, 17–45 (2009)
155. S. Pathak, S.K. Choi, N. Arnheim, M.E. Thompson, Hydroxylated quantum dots as luminescent probes for in situ hybridization. *J. Am. Chem. Soc.* **123**, 4103–4104 (2001)
156. F. Patolsky, R. Gill, Y. Weizmann, T. Mokari, U. Banin, I. Willner, Lighting-up the dynamics of telomerization and DNA replication by CdSe-ZnS quantum dots. *J. Am. Chem. Soc.* **125**, 13918–13919 (2003)
157. Z. Gueroui, A. Libchaber, Single-molecule measurements of gold-quenched quantum dots. *Phys. Rev. Lett.* **93** (2004)
158. L. Dyadyusha, H. Yin, S. Jaiswal, T. Brown, J.J. Baumberg, F.P. Booy, T. Melvin, Quenching of CdSe quantum dot emission, a new approach for biosensing. *Chem. Commun.* 3201–3203 (2005)
159. P.M. Chan, T. Yuen, F. Ruf, J. Gonzalez-Maeso, S.C. Sealton, Method for multiplex cellular detection of mRNAs using quantum dot fluorescent in situ hybridization. *Nucleic Acids Res.* **33** (2005)
160. A.A. Chen, A.M. Derfus, S.R. Khetani, S.N. Bhatia, Quantum dots to monitor RNAi delivery and improve gene silencing. *Nucleic Acids Res.* **33** (2005)

Chapter 21

Bionanotechnology: Lessons from Nature for Better Material Properties

F.N. Kök

Abstract For millions of years, nature has built hierarchically organized intricate systems with interesting material properties that synthetic materials often fail to replicate. With the advances in instrumentation for both characterization and manipulation in nano-scale, it has now become possible to comprehend the molecular mechanisms and structures behind that success and mimic them. Biomimicry should not be understood as a superficial imitation of the biological systems. It should rather be interpreted as the inspiration from the structure-function relationships observed in biological systems to construct new hierarchical structures with improved properties.

21.1 Introduction

For centuries, humankind has been both astonished by and inspired from the nature and tried to understand and mimic the natural systems to establish new materials and machines. Long before the invention of sophisticated instruments to observe and manipulate in micro or nano-scale, bat wings had been an inspiration for Leonardo da Vinci to design a flying machine or imitation of human arm movements had been used to design robotic arms [1]. In today's world, however, the concept of hierarchical design could be understood starting from molecular level and this gives us the opportunity to design smart, highly sophisticated materials starting from molecules or functional modules like lipid membranes, ion channels, motor proteins, etc.

Nature not only gives inspiration for better materials or systems but also for the establishment of less power hungry production technologies with minimum waste generation. The synthesis of sophisticated materials was done under mild reaction

F.N. Kök (✉)

Molecular Biology and Genetics Department, Istanbul Technical University,
34469 Maslak, Istanbul, Turkey
e-mail: kokf@itu.edu.tr

conditions from aqueous precursors as opposed to the conventional technologies which require the usage of organic solvents and extreme conditions (high temperature, pressure, acidity, etc.). This makes the bioinspired technologies more environmental friendly and the materials produced tend to lack the toxic residues; all this makes them intrinsically biocompatible [2]. The applications of these materials are not only limited to medical areas. A diverse industrial field ranging from photovoltaics to catalysis and pollution control could take advantage of these attractive production technologies.

The process of self-assembly, which could be defined as the autonomous organization of components into patterns or structures without human intervention is one of the most important mechanisms behind the biological structures and function [3]. Self-assembly processes depend on non-covalent bonds (e.g. hydrogen bonds, ionic bonds, and hydrophobic interactions) which are individually weak but extremely effective working in harmony for the construction of strong structures. Folded polypeptides, assembly of pore forming peptides, amyloid forming peptides, cell bilayer membrane, prokaryotic S-layers and lipid tubules, viral particles are some of the examples of the diverse processes and structures which are organized by self-assembly.

Self-assembly is also an attractive bottom-up strategy for the fabrication of nanostructures owing to highly ordered nature of the process and considerable degree of flexibility and diversity provided by biological molecules [4, 5]. Understanding these mechanisms helps scientists to design novel systems exhibiting unusual properties, even in nano-scale. These natural processes could be combined by inorganic processes and hybrid materials could also be designed. Supramolecular self-assembly and inorganic sol-gel polymerization, for example, could be combined to form higher self-organized hybrid materials with complex design and intricate properties [6]. A single type of molecule is rarely used in nature and materials are mostly arranged in composite systems. By this mean, a single building block, like collagen, is used as a major component in different tissues that require different properties, by applying different chemistries and assemblies [7]. Implementation of self-organization to obtain self-organized hybrid materials was explained in different studies and reviews [6, 8]. In this chapter, an overview of diverse applications will be given.

21.2 Biomineralization

Biomineralization results in materials with different size and shapes that take part in many interesting activities in living organisms. Although mechanical works like support or chewing and ion storage are the most prominent functions of these biominerals, they could also be used as navigation and optical tools. These minerals may be extremely ordered and are referred to as crystalline minerals (or biogenic crystals), or they may be relatively disordered and are referred to as amorphous minerals [9]. Biomineralization is an organic matrix-mediated process rather than a

simple thermodynamic and kinetic process of crystallization of inorganic mineral ions [10]. These biological determinants, which self-assemble and used as templates to control the nucleation, growth or morphology, have not been completely resolved yet, even in the systems where the molecules which are crucial for the event are known [11].

Mineralized tissue composed of calcium phosphate (e.g. bone and teeth) and calcium carbonate (e.g. seashells) is one of the most common examples for biomineralization. In the form of hydroxyapatite (HA), calcium phosphate is the most common mineral in vertebrate tissues providing mechanical durability to bones and teeth. In addition, being a stable, biocompatible, non-toxic and cheap inorganic material, it is a material of choice for the coating of hard tissue implants as well as carriers or ingredients in cosmetics and pharmaceutical applications [12, 13]. Calcium phosphate mineralization is therefore an important inspiration for scientist to construct highly organized, bioactive materials at near-ambient conditions avoiding high temperatures or organic solvents which are required for conventional production [14]. This process is generally done by bioinspired precipitation of inorganic minerals from aqueous solution using peptides, proteins, protein mimics, synthetic peptides, carbohydrates and DNA [15].

Although they are less abundant than calcium phosphate/carbonate minerals, silica from diatoms [16] and sponges [17] and magnetite from magnetic bacteria [11] are other interesting materials synthesized in nature. Magnetite, for example, is an iron oxide mineral that is found in various organisms ranging from unicellular bacteria to fish and it serves as a geonavigational or mechanical tool. In magnetotactic bacteria, in which its synthesis has been studied in detail, these monodisperse magnetic nanoparticles form chains which termed as magnetosomes [11]. Their potential in medical applications as MRI contrast agents makes in vitro synthesis of these materials attractive. For this, a typical strategy is to first identify the biologically determinant molecules and their characterization. These molecules are generally used for template-based synthesis and this method to syntheses inorganic materials was well-established [11, 15, 18, 19]. Consequently the efforts have been focused on the production—by similar approach—of materials, which are not biomineralized by living organisms (e.g. gold, silver, platinum) [14].

The discovery of new template molecules for both known and new biominerals could be done by two different approaches. In rational design, *in silico* modelling techniques based on existing knowledge was used to design new sequences [20]. The other method, on the other hand, requires no prior knowledge on existing molecules and basically depends on an evolutionary approach involving the selection of peptides from random peptide libraries constructed with surface display methods [11, 21, 22]. By successive amplification-selection cycles, high affinity peptides could be enriched. However, one should keep in mind that strong binding affinity towards the minerals does not always translate into better biomineralization directing ability and peptides with no significant binding affinity towards hydroxyapatite, for example, resulted in faster kinetics in biomineralization [19]. This could be explained by the interaction of the peptide with the soluble precursor ions rather than the mineral surface that the selection was actually done. Apart from

peptides, various polymers were also tried for their potential to assist bioinspired calcium phosphate mineralization from aqueous solution and gave successful results [15, 18].

21.3 Biomimetic Proteins: Receptors, Catalysts, Channels

Although proteins provide a selective recognition for many different molecules, lack of stability and high production cost often limit their usage in high-throughput systems. Strategies of supramolecular chemistry, on the other hand, offer an alternative to design new biomimetic proteins inspired from the biological analogues but providing better stability and applicability in production and operation. Most widely employed strategy involves the usage of template-directed syntheses and molecular imprinting technologies to produce molecularly imprinted polymers (MIPs) [23]. MIPs are generally produced to mimic natural receptors and therefore called as synthetic receptors. They are mostly synthesized by copolymerizing functional monomers in the presence of a particular template molecule (Fig. 21.1) [24]. When the target is removed, cavities are obtained with a high affinity and selectivity for the corresponding template molecule. MIPs that could selectively recognize small molecules like histamine, serotonin and L-nicotine can be produced and used in the biosensor construction [25].

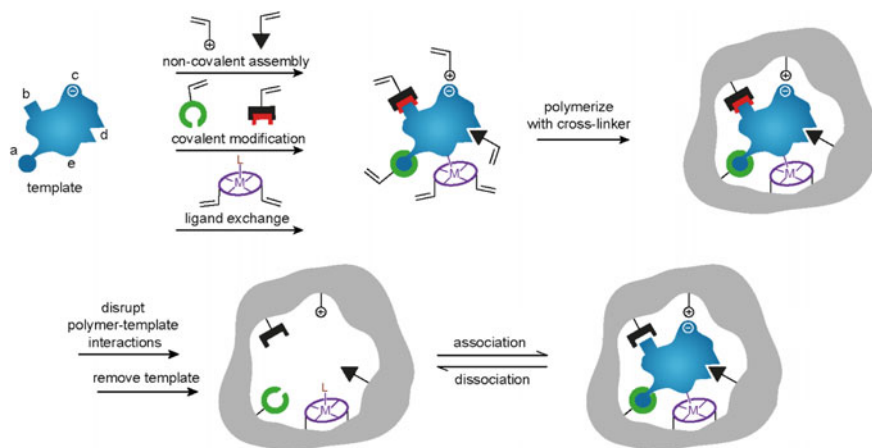


Fig. 21.1 Schematic representation of the molecular imprinting process. Reversible interactions between the template and the polymerizable functional monomer can involve different mechanisms explained in figure. After polymerization, a porous matrix in which the template sites are located is formed. Removal of the template leaves an empty area defined by the shape of the template and the specific interactions that it forms with functional monomers. The target analyte can selectively bind to this imprint. Reprinted with permission from Alexander et al. [24]. Copyright (2006) John Wiley and Sons

MIPs could also be used to produce enzyme mimics with high stability in organic solvents, extreme pH and temperature conditions, which are needed for most of the industrial processes. Although robust enzymes with tailored selectivities can be obtained, the catalytic efficiency of these systems is, in general, still well below that of the natural enzymes [26].

Apart from MIPs, hybrid approaches can also be used, especially in the fabrication of nanoscale pores with high selectivity based on specific biochemical properties [27]. Solid-state fabrication of pores in nanoscale has some challenges such as difficulties in adjustment of the pore size and specificity, non-specific interactions and clogging of the pores. Another alternative is the usage of wild-type and genetically modified biological pores in the construction of selective and sensitive biosensors [28, 29] but instability of proteins and difficulties in their purification may limit their large scale production. These limitations and challenges can be addressed by biomimetic approaches. Olfactory systems of insects, for example, have nanochannels coated with a lipid layer which helps to preconcentrate odorant molecules before transporting. A similar system was adapted to a synthetic pore and the resultant lipid-coated nanopores had better selectivity and caused less non-specific adsorption [30]. In addition, translocation speed of proteins can be tuned by regulating the lipid-bilayer viscosity.

21.4 Optics/Biophotonics

One of the exiting areas of bioinspired materials is the development of components for devices, such as optics, photonics, photovoltaics, circuits and data storage [14]. Human eye uses gradient refractive index lenses with better focusing power than traditional glass ones [31]. Although this type of lenses can be fabricated using various methods [32], their usage as ocular implants is limited due to their poor performance and lack of biocompatibility. Both limitations had been addressed by several different bioinspired approaches. Template-mediated biomineralization using an amphiphilic biopolymer-like surfactant, for example, was shown to be useful to mediate the self-assembly of calcium carbonate from solution to construct uniform hemispherical microlenses [33]. These lenses require no extreme conditions for fabrication and are biocompatible. Insect eyes, on the other hand, contain a microstructured part which significantly reduces the reflectance of the lens surface [34]. Apart from insect eyes, distinctively organized biosilica fibers which form spicules of the glass sponge *Euplectella aspergillum* of deep seas possesses interesting optical properties and could be an inspiration to construct optical materials with better properties [17]. Similar microstructured materials have been shown to be effective not only to produce optical lenses but also anti-reflective coatings [35] and even solar cells [36]. High cost and low efficiency of the existing solar cells are the main challenges in this area and anti-reflection coatings inspired from the corneal nipple arrays found in moths' eyes had been found to significantly increase the conversion efficiency [36]. Another strategy relies on natural photosynthesis

mechanism in which the conversion of sunlight into usable energy is efficiently realized. Artificial photosynthetic systems can be constructed by implementing biological elements like enzyme with nanoparticles to increase the efficiency and robustness of these photocatalytic hybrid systems [37].

The shift from natural pigments to synthetic dyes made the production of more economical textiles with diverse colors but their heavy usage led to environmental problems due to their toxic and even carcinogenic nature. Biomimetic approach may offer a completely different way of producing bright colors, using photonic structures rather than pigments [38–40]. Butterflies and beetles display striking colors originated from a wide range of structural mechanisms such as multilayer films, diffraction gratings and quasi-ordered scattering [41]. Morphotex, for example, is an unstained, structurally colored fiber, which is based on the biomimetic conception for the microscopic structure of Morpho butterfly's wings. By precisely controlling the layer thickness of the structure according to visible wavelength, four types of basic colors such as red, green, blue and violet can be developed [42]. In addition, some beetle species demonstrate reversible color changes, based on the hydration of a porous ordered three-dimensional matrix [43]. These structures could be used in the development of optic displays and biomimetic sensors.

Beetles and butterflies are not the only options for scientists to design color-shifting systems. A phage-bundle nanostructure inspired from wattle structure of turkey, for example, was used to design a tunable, colorimetric sensor for trinitrotoluene (TNT) detection [44]. Wattle is composed of bundles of collagen enriched with lots of blood vessels and in relaxed turkey the bundles are small and transparent, exposing the blood vessels through the skin, creating a red color. When a turkey gets excited, however, the bundles expand, become opaque and the color shifts from red to white, or even blue. Similar structure was mimicked using self-assembled genetically engineered phage displaying TNT-binding peptide motifs. On exposure to various volatile organic chemicals, the structures rapidly swell and undergo distinct color changes.

21.5 Natural Adhesives

Natural adhesives derived from marine organisms, especially marine mussels (*M. edulis*), have been extensively studied for the past three decades and the role of proteins in the success of these adhesives is well-established [45]. These materials makes the adhesion to metal and mineral surfaces possible even in wet saline environments at ambient temperatures, which could not be achieved by man-made ones. Although their potency and non-toxicity is an advantage for both medical and industrial applications, difficulties in extracting and purifying sufficient quantities from the producing organism or low product yield have limited their commercialization. Moreover, recombinant production of mussel proteins often yields products with reduced functionalities compared to native proteins. Biomimetic

approach, on the other hand, could make the design of practical analogs of these adhesive materials possible; indeed improved functionalities can be achieved at lower cost [46].

The knowledge on the mechanism behind this performance was piled up over several decades using various approaches of molecular biology, chemistry and physics fields [47–51]. Mussel adhesive proteins (MAPs), i.e. polyphenolic proteins, collagen, and polyphenol oxidase, are important components of byssus responsible for adhesion. When needed, these proteins were secreted into the seawater where the curing process occurs. Quick diffusion of these proteins to the sea environment was prevented by complex coacervation, a phenomenon occurred in the presence of two oppositely charged polyelectrolyte solutions at isoelectric pH [52]. Tyrosine residues in polyphenolic proteins are converted to 3,4-dihydroxy-phenylalanine (DOPA) by polyphenol oxidase and high DOPA content was associated with good adhesion [45]. Although the exact mechanism was not known, various processes are suggested to work concurrently [45] and complex formation of metal ions, metal oxides and semimetals with DOPA-containing proteins to form crosslinked structures is one of them [51]. Collagen proteins, on the other hand, mainly provides the mechanical features (e.g. elasticity, extensibility, strength) needed for shock absorbing ability and endurance [53].

Although some pieces of the puzzle are still missing, the existing knowledge helped researchers to design biomimetic strategies for the production of non-toxic and more effective adhesives. One way of producing biomimetic adhesives is to synthesize simplified MAP mimics composed of shorter peptides using functional groups such as DOPA and hydroxyproline rather than developing whole protein [54]. The other way is the usage of knowledge on adhesion mechanism to develop completely new molecules. In one of these strategies, L-poly(lactic acid) composites reinforced with vegetable fibers together with dopamine had been shown to exhibit improved mechanical properties compared with conventional ones [55].

21.6 Biointerfaces

As already mentioned, nature organizes functional units by weak interactions and chirality. Various functional units including proteins, DNA double helix, phospholipid membranes, photosystems, and biominerals are self-assembled or directed by these interactions [56]. Another important issue in nature is the interaction of nature's elements with its surrounding environment via dynamic biointerfaces, which is also choreographed by highly selective weak interactions responsive to different external stimuli (light, pH, temperature, adsorption, etc.). With the help of the conformational transitions induced by external stimuli, surface characteristics could be changed which subsequently leading to macroscopic functions. Interaction with weak forces also brings reversibility to the mechanism and various biological events could not take place without reversibility. Stimuli-responsive smart polymers like thermo-responsive poly(N-isopropylacrylamide) (PNIPAAm) can be used for

this purpose but still have limited applicability. For this reason various different strategies from the design of three-component molecular systems including “recognition-mediating-function” (RMF) to mimic complex interactions in vivo [56] to the arrangement of surface topography based on different structural designs [4, 57, 58] were used. In this chapter, two different motivations for interface design, production of self-cleaning surfaces and biocompatible surfaces, are covered.

21.6.1 Self-cleaning Surfaces

One of the well-known biomimetic approaches in surface design was based on the discovery of so called “lotus effect” based on the self-cleaning ability of lotus flower petals [59]. Apart from plants, a wide range of living organisms (e.g. insect wings, water strider legs, shark skin, gecko feet, spider silks, bird feathers) had been investigated for their unique properties resulting different wettability profiles and self-cleaning features and Gecko (*Gekko gecko*) feet is one of the attractive examples since these animals can climb vertical surfaces without the need of cleaning their feet [60]. This behavior was explained by the presence of adhesive foot hairs (or setae), a uniform array on overlapping lamellar pads, which increases the intermolecular interactions with the profile of the substrate and thus enhance the adhesion [60, 61].

Self-cleaning surfaces exhibit low adhesion for a broad range of foulants. This is achieved with the help of a complex interplay between chemistry and surface morphology and can be classified in two classes: hydrophobic surfaces or hydrophilic (underwater oleophobic) surfaces [57]. Superhydrophobic non-wetting capability is a fundamental property of typical self-cleaning biological surfaces and very different surface structures can lead to similar outcome. The underwater oleophobic self-cleaning surfaces, on the other hand, are inspired from the hydrated skin of marine organisms and the cooperation of physical heterogeneity and high hydration energy moieties work synergistically to prevent oil-fouling [57, 62].

Studies were then focused on the design of artificial self-cleaning adhesive surfaces using different materials and production methods [58, 61, 63] and biomimetic surfaces with similar wetting behavior of that of the natural ones had been obtained (Fig. 21.2) [64]. Wu et al. [64] used an improved laser interference lithography to produce a surface topography resembling to an array of a string of pearls. By tailoring their height along x and y directions, and designing their shape and size, a controlled anisotropic biosurfaces were prepared. Prepared surface showed a similar wetting behavior of that of rice leaf which was known as one of the most typical examples for anisotropic wetting caused by those directional microstructures (Fig. 21.2). These petal inspired surfaces could also be produced by flexible replication methods like microinjection compression molding needed for the mass-production of different hierarchical topographies (e.g. biomimics of rice or lotus leaf) using designed templates and possess good self-cleaning performance in submicron contaminants and good mechanical robustness [65]. Production of

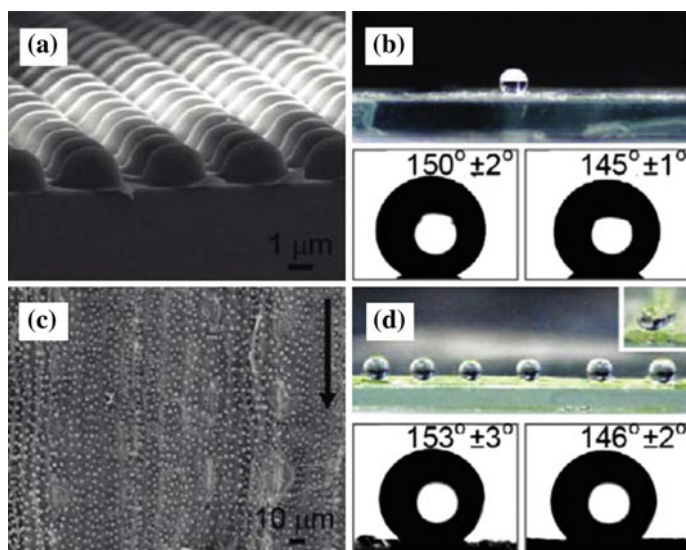


Fig. 21.2 Comparison of **a** biomimetic anisotropic surfaces and **c** natural rice leaf. Water droplets and contact angles associated with **b** biomimetic and **d** natural surface. Reprinted with permission from Wu et al. [64]. Copyright (2010) American Chemical Society

micropatterned vertically aligned carbon nanotubes on a silicon substrate is another approach to mimic setae in gecko feet to obtain sticky tapes with higher shear resistance than the natural gecko feet without compromising its remarkable self-cleaning abilities [63].

21.6.2 Bioinspired Interfaces for Better Biocompatibility

One of the most crucial properties of biomaterials that we use to restore or replace damaged or lost tissue functions is biocompatibility. The term “biocompatibility” has evolved over years and it is now considered as a characteristic of a material-biological host system and not a property of a material [66]. For the success of a biomaterial, relationships between materials and host tissue should be understood so that materials could be designed to elicit the most appropriate response. Designing the host-material interface has therefore a great importance on the integration of the biomaterial to the body by triggering an appropriate tissue response and directing cells to attach, proliferate or even differentiate to desired lineages. Organs and tissues have a definite micro-architecture and use specific recognition motifs so when these elements are implemented in the design, they could increase the success of implants in the body.

Designing implants in nano/micro level is therefore utmost importance to achieve a long-term success and various biomimetic approaches have been used to enhance the performance of implants and tissue engineering scaffolds [67]. Modification of implant surfaces by self-organized peptides to produce a bio-inspired interface is one of the approaches to address this problem [68, 69]. In one of these studies, a chimeric peptide composed of a peptide specifically binding to titanium and another one inducing a regulatory pathway activating osteogenic differentiation of mesenchymal stem cells for bone formation was used [69]. Attachment of peptide by molecular recognition and self-assembly have been found to prevent the problems arising from chemical coupling and provided a mild and effective way to stimulate osteogenesis in mesenchymal stem cells in vitro. Self-assembled double layer formed by collagen I and IV have been proposed as another approach to provide microenvironment mimicking the structure and organization of the native tissue [70]. The results showed that self-assembled collagen structures supported the cell maturation and functionality compared to dip-coated collagen surface. This also shows that not only the molecule itself but how it is organized on the surface affects the success of the system and biomimetic design strategies lead to more superior structures.

21.7 Biomimetic Membranes

Cell membrane has a complex structure composed of various elements and is therefore hard to mimic. Artificial bilayer membranes can be designed but incorporation of membrane proteins into these structures is still an important challenge due to the problems associated with their production, purification and instability in aqueous solutions. Membrane biomimics are generally used in two main areas, as described in the following sections.

21.7.1 Cell Membrane Mimics

Cell membrane mimics constituted in vitro are interesting and valuable systems in basic membrane research. Cell membranes mimics are generally composed of phospholipids that are already found in cell membrane. They could be formed either in freely floating form like liposomes and nanodiscs [71–73] or on a solid surface as solid-supported bilayer lipid membranes (sBLMs), tethered bilayer membranes (tBLMs) [74–77], etc. The properties of these systems can be tailored both in terms of phospholipid type and composition and membrane protein repertoire. This gives the opportunity to study membrane protein–ligand and membrane protein–membrane interactions in vitro conditions which cannot be done easily in vivo conditions due to many external and internal factors involved in cellular organization. These cell membrane mimics with or without proteins are not only promising for basic research

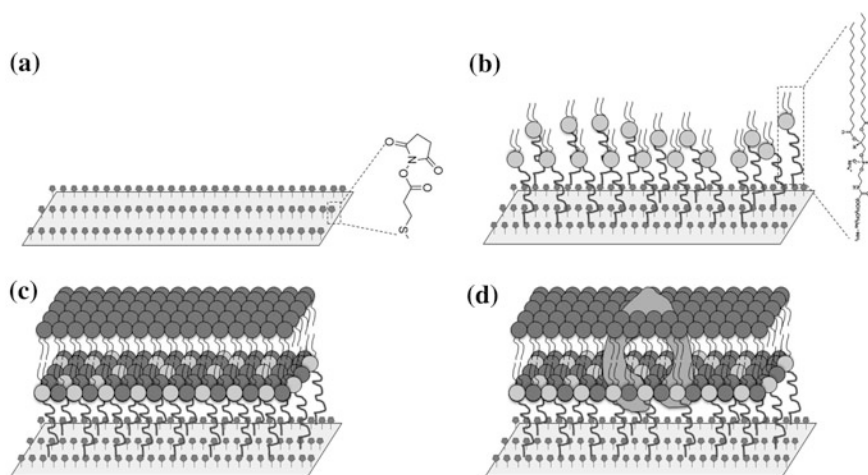


Fig. 21.3 Construction of tBLM on gold-coated surface. **a** Activation of the surface. **b** Attachment of tethering molecule (PEG2000 modified lipid). **c** Construction of protein-free tBLM. **d** Construction of protein-incorporated tBLM (molecules were not presented in their actual sizes). Reprinted from Inci et al. [77]. Copyright (2015), with permission from Elsevier

but also for different applications especially in medical fields to construct sensors and drug screening systems.

Tethered bilayer membranes are especially attractive due to their ability to incorporate membrane proteins of different sizes by preserving their biological activities. Different molecules (e.g. peptides, polymers, etc.) and approaches can be used to tether the lipid bilayer to the solid surface [78]. For a membrane protein with small extra-membrane part, a short tethering molecule is enough for successful incorporation [76]. When, however, a membrane protein with large extra-membrane domain has to be incorporated into tBLM, more space is needed in between the solid surface and the bilayer. This restriction can be overcome by the usage of long tethering molecules. To construct a tBLM consisting p-glycoprotein, for example, a modified polymer spacer (polyethylene glycol-2000-modified lipid molecule) can provide this necessary space in order not to disrupt the active conformation of the protein [77]. Construction of this system needs extensive optimizations, and generally done in several steps such as: (i) activation of the surface (Fig. 21.3a), (ii) attachment of tethering layer (Fig. 21.3b), and (iii) construction of whole bilayer without (Fig. 21.3c) or with (Fig. 21.3d) membrane protein via liposome spreading. This protein incorporated tBLMs can be used to investigate drug (e.g. statin-based cholesterol lowering drug)-membrane protein (p-glycoprotein) interactions in an isolated environment [77] and this work can also be extended to the study of binding kinetics of drugs to different membrane related components.

21.7.2 Membranes for Water Treatment

Increasing amount of wastewater and decreasing access to clean water supplies makes water purification and wastewater treatment inevitable to obtain usable water from new sources. Membrane technology for the production of high quality water has therefore gained importance in recent years and non-traditional water sources originated from agricultural/industrial wastewater, seawater, etc. have been studied for this purpose. Due to the diverse nature of these sources, different strategies have to be applied to improve the separation performance and biomimetic approaches have provided new insights to the challenges associated with that.

One of the important problems in the conventional oil removing membranes is the easy fouling by oils. Self-cleaning ability of various living organisms discussed in Sect. 21.6.1 could be adopted to design membranes with low oil fouling property [57]. Hierarchical structures of fish scale had inspired researchers to design polyacrylamide hydrogel-coated mesh membranes that have been successfully used in gravity-driven oil-water separation process [62].

In many water treatment applications, not only large contaminants like bacteria but also small ones like metal ions are needed to be removed and this is mainly done by sieving mechanism based on the membrane pore size. To obtain nanoporous membranes with good resolution and high throughput, a well-known natural counterpart, cell membrane and components associated with that for selective transport, can be used to develop new strategies. One of the most widely studied examples is the usage of aquaporins (AQPs) in desalination process [79, 80]. Cell membrane separates the intracellular functions from the effect of surrounding environment and only selectively allows the passage of different molecules when needed. Even water could not pass from this membrane and need specialized membrane proteins, i.e. aquaporins, to enter. AQPs let only water pass, rejecting all other solutes [81], and therefore they have been intensively studied as candidates to construct desalination membranes [80]. When the protein-to-polymer ratio was optimized, the permeability values 3000 times higher than that of the pure polymer and AQP based composite membranes with competitive water permeability and ion rejection capacity compared with the existing systems could be obtained [82–84].

21.8 Hints from Nature for Endurance

Hostile and changing environmental conditions induce severe changes in the properties and durability of the exposed materials. To deal with various stresses, living organisms synthesized a wide range of composite materials with desired properties [85]. Protein fiber-reinforced composites constituting byssus of mussels

were an inspiration for the production of synthetic materials with gradual mechanical properties which can withstand various environmental conditions [86]. Another example is the unusual mechanical tensile strength of egg capsules synthesized by marine snails from the family *Melongenidae* [87]. This behavior was found to be originated from conformational changes of the protein backbone under mechanical stress. Similar studies revealed the involvement of protein fibrils to obtain hybrid materials with better mechanical properties and this led to design of high-performance materials with locally defined properties (like mechanics, porosity). This type of gradual composites has gained importance in different fields like automotive and aircraft technology as well as the biomedical field [86]. Materials are generally limited by their resistance to fracture rather than strength and therefore most critical structures, such as bridges, ships, nuclear pressure vessels, are manufactured from materials that are comparatively low in strength but high in toughness [88]. Hierarchical structural architectures found in nature could be an answer to produce advanced hybrid materials with better strength and toughness properties. Although it sounds very promising, biomimetic approach has yielded few technological advances, simply because of the lack of processing techniques in practical dimensions needed for construction industry [88]. In general, direct mimicry of the biological systems would not be enough to find engineering solutions to improve mechanical properties. Living organisms use the materials that are found in their environment and except from the extremophiles, they survive in environments that have restricted mechanical loads, fairly narrow temperature regimes, and so on, so their structures are not necessarily resistant to every condition they may face during an industrial application [89]. It is therefore important to be “inspired” from the nature to build new structure and if needed using different materials but not directly try to mimic it.

From nature’s various strategies to construct armored protection systems, scaled skins, in particular, are attractive due to their remarkable mechanical properties [90]. This ultrathin structure is not only compliant, resistant to penetration and lightweight but also its discrete, spatially overlapped scales attached to a flexible dermis layer allow flexibility needed for the animals’ motion [91]. For example, the arrangement of the scales of teleost fish is such that the soft dermis is covered by at least three layers of scales [92]. Inspired from this elegant multilayer structure, a synthetic fish skin material was designed by Funk et al. [90] (Fig. 21.4) to replicate the structural, mechanical, and functional aspects of natural teleost fish skin. Key characteristics of this complex structure was defined as low-modulus elastic mesh or “dermis” layer that holds relatively rigid plastic scales and constructed system was tested for its mechanical properties. This design resulted to an attractive low-weight, low-profile protective coating for soft materials.

Another strategy that nature uses to increase the competitiveness of living systems in changing environmental conditions is the use of protective covers such as prokaryotic S-layers. These S-layer proteins can self-assemble in solution and at

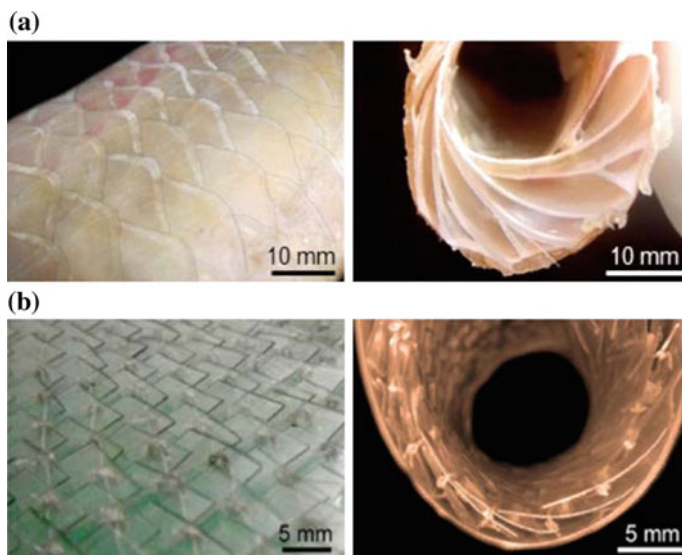


Fig. 21.4 Photographs of the skin of **a** the striped red mullet (*Mullus surmuletus*) and **b** the synthetic material when undeformed (*left*) and under a bending deformation (*right*). Adapted with permission from Funk et al. [90]. Copyright (2015) American Chemical Society

interfaces and can be used as protective coats for mechanically unstable structures like liposomes and planar lipid membranes [93]. Instability of these artificial lipid systems often hinders their commercial usage and this approach might help to improve their longevity and, in particular, offer a solution to the problems emerging upon their drying [94].

21.9 Conclusions

Nature provides an immense source of materials and systems, and mimicking the composition, structure and formation principles of these systems help researchers to design materials with improved properties. A tremendous amount of knowledge has been accumulated for the last couple of decades about the molecular mechanisms and structures behind natural materials and systems. Understanding the mechanisms behind the success of the living systems and exploit them for practical purposes requires interdisciplinary research with the cooperative effort of biologists, physicists, chemists, material scientists and engineers. Biomimicry does not mean to produce the exact copy of the biological systems by bare imitation. Once the molecular mechanisms and structure-function relationships have been understood, they could be adopted to construct completely new hierarchical structures with properties tailored for specific applications through an interdisciplinary effort.

References

1. M. Billard, J. Matari, Learning human arm movements by imitation: evaluation of a biologically inspired connectionist architecture. *Robot Auton. Syst.* **941**, 1–16 (2001)
2. X. Li, H. Xu, Z.S. Chen, G. Chen, Biosynthesis of nanoparticles by microorganisms and their applications. *J. Nanomater.* **2011**, 270974 (16 pp) (2011)
3. G.M. Whitesides, B. Grzybowski, Self-assembly at all scales. *Science* **295**, 2418–2421 (2002)
4. J.H. Zhang, Y.R. Zhao, S.Y. Han, C.X. Chen, H. Xu, Self-assembly of surfactant-like peptides and their applications. *Sci. China Chem.* **57**, 1634–1645 (2014)
5. A. Rawat, R. Nagaraj, Covalently attached fatty acyl chains alter the aggregation behavior of an amyloidogenic peptide derived from human $\beta(2)$ -microglobulin. *J. Pept. Sci.* **12**, 770–783 (2013)
6. M. Barboiu, Constitutional hybrid materials—Toward selection of functions. *Eur. J. Inorg. Chem.* **2015**, 1112–1125 (2015)
7. S. Lapidot, S. Meirovitch, S. Sharon, A. Heyman, D.L. Kaplan, O. Shoseyov, Clues for biomimetics from natural composite materials. *Nanomedicine* **7**, 1409–1423 (2012)
8. M. Hnilova, B. Taktak Karaca, J. Park, C. Jia, B.R. Wilson, M. Sarikaya, C. Tamerler, Fabrication of hierarchical hybrid structures using bio-enabled layer-by-layer self-assembly. *Biotechnol. Bioeng.* **109**, 1120–1130 (2012)
9. L. Addadi, S. Weiner, Biomineralization: mineral formation by organisms. *Phys. Scr.* **89**, 098003 (13 pp) (2014)
10. T.-Y. Ning, X.-H. Xu, L.-F. Zhu, X.-P. Zhu, C.H. Chu, L.-K. Liu, Q.-L. Li, Biomimetic mineralization of dentin induced by agarose gel loaded with calcium phosphate. *J. Biomed. Mater. Res., Part B* **100B**, 138–144 (2012)
11. J. Baumgartner, M.A. Carillo, K.M. Eckes, P. Werner, D. Faivre, Biomimetic magnetite formation: from biocombinatorial approaches to mineralization effects. *Langmuir* **30**, 2129–2136 (2014)
12. A. Sinha, T. Mishra, N. Ravishankar, Polymer assisted hydroxyapatite microspheres suitable for biomedical application. *J. Mater. Sci. Mater. Med.* **19**, 2009–2013 (2008)
13. S. Onder, A.C. Calikoglu-Koyuncu, K. Kazmanli, M. Urgen, G.T. Kose, F.N. Kok, Behavior of mammalian cells on magnesium substituted bare and hydroxyapatite deposited (Ti, Mg)N coatings. *New Biotechnol.* (2015). doi:[10.1016/j.nbt.2014.11.006](https://doi.org/10.1016/j.nbt.2014.11.006)
14. J.M. Galloway, J.P. Bramble, S.S. Staniland, Biomimetic synthesis of materials for technology. *Chem. Eur. J.* **19**, 8710–8725 (2013)
15. S. Schweizer, A. Taubert, Polymer-controlled, bio-inspired calcium phosphate mineralization from aqueous solution. *Macromol. Biosci.* **7**, 1085–1099 (2007)
16. M. Hildebrand, Diatoms, biomineralization processes, and genomics. *Chem. Rev.* **108**, 4855–4874 (2008)
17. J. Aizenberg, V.C. Sundar, A.D. Yablon, J.C. Weaver, G. Chen, Biological glass fibers: correlation between optical and structural properties. *Proc. Natl. Acad. Sci. USA* **101**, 3358–3363 (2004)
18. K. Bleek, A. Taubert, New developments in polymer-controlled, bioinspired calcium phosphate mineralization from aqueous solution. *Acta Biomater.* **9**, 6283–6321 (2013)
19. M. Gungormus, E.E. Oren, J.A. Horst, H. Fong, M. Hnilova, M.J. Somerman, M.L. Snead, R. Samudrala, C. Tamerler, M. Sarikaya, Cementomimetics—constructing a cementum-like biomineralized microlayer via amelogenin-derived peptides. *Int. J. Oral Sci.* **4**, 69–77 (2012)
20. E.E. Oren, C. Tamerler, D. Sahin, M. Hnilova, U.O.S. Seker, M. Sarikaya, R. Samudrala, A novel knowledge-based approach to design inorganic-binding peptides. *Bioinformatics* **23**, 2816–2822 (2007)
21. J. Kim, Y. Rheem, B. Yoo, Y. Chong, K.N. Bozhilov, D. Kim, M.J. Sadowsky, H.-G. Hur, N. V. Myung, Peptide-mediated shape- and size-tunable synthesis of gold nanostructures. *Acta Biomater.* **6**, 2681–2689 (2010)

22. M. Sarikaya, C. Tamerler, A.K.Y. Jen, K. Schulten, F. Baneyx, Molecular biomimetics: nanotechnology through biology. *Nat. Mater.* **2**, 577–585 (2003)
23. F.L. Dickert, Biomimetic receptors and sensors. *Sensors* **14**, 22525–22531 (2014)
24. C. Alexander, H.S. Andersson, L.I. Andersson, R.J. Ansell, N. Kirsch, I.A. Nicholls, J. O'Mahony, M.J. Whitcombe, Molecular imprinting science and technology: a survey of the literature for the years up to and including 2003. *J. Mol. Recog.* **19**, 106–180 (2006)
25. G. Wackers, T. Vandenryt, P. Cornelis, E. Kellens, R. Thoelen, W. de Ceuninck, P. Losada-Pérez, B. van Grinsven, M. Peeters, P. Wagner, Array formatting of the heat-transfer method (HTM) for the detection of small organic molecules by molecularly imprinted polymers. *Sensors* **14**, 11016–11030 (2014)
26. M. Resmini, Molecularly imprinted polymers as biomimetic catalysts. *Anal. Bioanal. Chem.* **402**, 3021–3026 (2012)
27. S.W. Kowalczyk, T.R. Blosser, C. Dekker, Biomimetic nanopores: learning from and about nature. *Trends Biotechnol.* **29**, 607–614 (2011)
28. S. Howorka, Z. Siwy, Nanopore analytics: sensing of single molecules. *Chem. Soc. Rev.* **38**, 2360–2384 (2009)
29. T. Nowotny, M. de Bruyne, A.Z. Berna, C.G. Warr, S.C. Trowell, Drosophila olfactory receptors as classifiers for volatiles from disparate real world applications. *Bioinspir. Biomim.* **9**, 046007 (13 pp) (2014)
30. E.C. Yusko, J.M. Johnson, S. Majd, P. Prangkio, R.C. Rollings, J. Li, J. Yang, M. Mayer, Controlling protein translocation through nanopores with bio-inspired fluid walls. *Nat. Nanotechnol.* **6**, 253–260 (2011)
31. Y. Jin, H. Tai, A. Hiltner, E. Baer, J.S. Shirk, New class of bioinspired lenses with a gradient refractive index. *J. Appl. Polym. Sci.* **103**, 1834–1841 (2007)
32. H. Ren, S.-T. Wu, Adaptive liquid crystal lens with large focal length tenability. *Opt. Express* **14**, 11292–11298 (2006)
33. K. Lee, W. Wagermaier, A. Masic, K.P. Kommareddy, M. Bennet, I. Manjubala, S.-W. Lee, S. B. Park, H. Colfen, P. Fratzl, Self-assembly of amorphous calcium carbonate microlens arrays. *Nat. Commun.* **3**, 725 (2012)
34. D.G. Stavenga, S. Foletti, G. Palasantzas, K. Arikawa, Light on the moth-eye corneal nipple array of butterflies. *Proc. R. Soc. B* **273**, 661–667 (2006)
35. C. Morhard, C. Pacholski, D. Lehr, R. Brunner, M. Helgert, M. Sundermann, J.P. Spatz, Tailored antireflective biomimetic nanostructures for UV applications. *Nanotechnology* **21**, 425301 (2010)
36. R. Dewan, S. Fischer, W.B. Meyer-Rochow, Y. Ozdemir, S. Hamraz, D. Knipp, Studying nanostructured nipple arrays of moth eye facets helps to design better thin film solar cells. *Bioinspiration Biomimetics* **7**, 016003 (2012)
37. P.W. King, Designing interfaces of hydrogenase–nanomaterial hybrids for efficient solar conversion. *Bioch. Biophys. Acta* **1827**, 949–957 (2013)
38. K. Watanabe, T. Hoshino, K. Kanda, Y. Haruyama, S. Matsui, Brilliant blue observation from a morpho-butterfly-scale quasi-structure. *Jap. J. Appl. Phys.* **44**, L48–L50 (2005)
39. J. Xu, Z. Guo, Biomimetic photonic materials with tunable structural colors. *J. Colloid Interf. Sci.* **406**, 1–17 (2013)
40. K. Yu, T. Fan, S. Lou, D. Zhang, Biomimetic optical materials: Integration of nature's design for manipulation of light. *Prog. Mater. Sci.* **58**, 825–873 (2013)
41. A.E. Seago, P. Brady, J.-P. Vigneron, T.D. Schultz, Gold bugs and beyond: a review of iridescence and structural colour mechanisms in beetles (Coleoptera). *J. R. Soc. Interface* **6**, S165–S184 (2009)
42. <http://www.nanotechproject.org/cpi/products/morphotex-r-fiber/>
43. M. Rassart, J.-F. Colomer, T. Tabarrant, J.P. Vigneron, Diffractive hygrochromic effect in the cuticle of the hercules beetle *Dynastes hercules*. *New J. Phys.* **10**, 033014 (14 pp) (2008)
44. J.-W. Oh, W.-J. Chung, K. Heo, H.-E. Jin, B.Y. Lee, E. Wang, C. Zueger, W. Wong, J. Meyer, C. Kim, S.-Y. Lee, W.-G. Kim, M. Zemla, M. Auer, A. Hexemer, S.-W. Lee, Biomimetic virus-based colourimetric sensors. *Nat. Commun.* **5**, 3043 (2014)

45. N. Bandara, H. Zeng, J. Wu, Marine mussel adhesion: biochemistry, mechanisms, and biomimetics. *J. Adhesion Sci. Technol.* **27**, 2139–2162 (2013)
46. H.J. Cha, D.S. Hwang, S. Lim, Development of bioadhesives from marine mussels. *Biotech. J.* **3**, 631–638 (2008)
47. J.H. Waite, Adhesion a la moule. *Integr. Comp. Biol.* **42**, 1172–1180 (2002)
48. Q. Lin, D. Gourdon, C. Sun, N. Holten-Andersen, T.H. Anderson, J.H. Waite, J.N. Israelachvili, Adhesion mechanisms of the mussel foot proteins mfp-1 and mfp-3. *PNAS USA* **104**, 3782–3786 (2007)
49. S. Haemers, G.J.M. Koper, G. Frens, Effect of oxidation rate on cross-linking of mussel adhesive proteins. *Biomacromolecules* **4**, 632–640 (2003)
50. D.S. Hwang, Y. Gim, H.J. Yoo, H.J. Cha, Practical recombinant hybrid mussel bioadhesive fp-151. *Biomaterials* **28**, 3560–3568 (2007)
51. D.S. Hwang, H. Zeng, A. Masic, M.J. Harrington, J.N. Israelachvili, J.H. Waite, Protein and metal dependent interactions of a prominent protein in mussel adhesive plaques. *J. Biol. Chem.* **285**, 25850–25858 (2010)
52. D. Hwang, H. Zeng, A. Srivastava, D.V. Krogstad, M. Tirrell, J.N. Israelachvili, J.H. Waite, Viscosity and interfacial properties in a mussel-inspired adhesive coacervate. *Soft Matter* **6**, 3232–3236 (2010)
53. H.G. Silverman, F.F. Roberto, Understanding marine mussel adhesion. *Marine Biotechnol.* **9**, 661–681 (2007)
54. M. Yu, T. DeSimone, T.J. Deming, Synthetic analogs of marine mussel cement protein. *Polym. Prepar.* **38**, 101–102 (1997)
55. A. Bourmaud, J. Riviere, D. Le Antoine, G. Raj, C. Baley, Investigations of the use of a musselinspired compatibilizer to improve the matrix-fiber adhesion of a biocomposites. *Polym. Testing* **28**, 668–672 (2009)
56. B. Chang, M. Zhang, G. Qing, T. Sun, Dynamic biointerfaces: from recognition to function. *Small* **11**, 1097–1112 (2015)
57. J. Zhao, X. Zhao, Z. Jianga, Z. Li, X. Fana, J. Zhu, H. Wu, Y. Su, D. Yang, F. Pan, J. Shi, Biomimetic and bioinspired membranes: preparation and application. *Prog. Polym. Sci.* **39**, 1668–1720 (2014)
58. B. Bhushan, R.A. Sayer, Surface characterization and friction of a bio-inspired reversible adhesive tape. *Microsyst. Technol.* **13**, 71–78 (2007)
59. C. Neinhuis, W. Barthlott, Characterization and distribution of water-repellent, self-cleaning plant surfaces. *Ann. Bot.* **79**, 667–677 (1997)
60. W.R. Hansen, K. Autumn, Evidence for self-cleaning in gecko setae. *PNAS* **102**, 385–389 (2005)
61. T.S. Kustandi, V.D. Samper, W.S. Ng, A.S. Chong, H. Ga, Fabrication of a gecko-like hierarchical fibril array using a bonded porous alumina template. *J. Micromech. Microeng.* **17**, 75–81 (2007)
62. Z. Xue, S. Wang, L. Lin, L. Chen, M. Liu, L. Feng, L. Jiang, A novel superhydrophilic and underwater superoleophobic hydrogel-coated mesh for oil/water separation. *Adv. Mater.* **23**, 4270–4273 (2011)
63. S. Sethi, L. Ge, L. Ci, P.M. Ajayan, A. Dhinojwala, Gecko-inspired carbon nanotube-based self-cleaning adhesives. *Nano Lett.* **8**, 822–825 (2008)
64. S.-Z. Wu, D. Wu, J. Yao, Q.-D. Chen, J.-N. Wang, L.-G. Niu, H.-H. Fang, H.-B. Sun, One-step preparation of regular micropearl arrays for two-direction controllable anisotropic wetting. *Langmuir* **26**, 12012–12016 (2010)
65. W.-S. Guan, H.-X. Huang, A.-F. Chen, Tuning 3D topography on biomimetic surface for efficient self-cleaning and microfluidic manipulation. *J. Micromech. Microeng.* **25**, 035001 (9 pp) (2015)
66. D.F. Williams, There is no such thing as a biocompatible material. *Biomaterials* **35**, 10009–10014 (2014)
67. D. Singh, D. Singh, S. Zo, S.S. Han, Nano-biomimetics for nano/micro tissue regeneration. *J. Biomed. Nanotechnol.* **10**, 3141–3161 (2014)

68. H. Yazici, H. Fong, B. Wilson, E.E. Oren, F.A. Amos, H. Zhang, J.S. Evans, M.L. Snead, M. Sarikaya, C. Tamerler, Biological response on a titanium implant-grade surface functionalized with modular peptides. *Acta Biomater.* **9**, 5341–5352 (2013)
69. Y. Zhou, M.L. Snead, C. Tamerler, Bio-inspired hard-to-soft interface for implant integration to bone. *Nanomed. Nanotech. Bio Med.* **11**, 431–434 (2015)
70. A.E. Sorkio, E.P. Vuorimaa-Laukkanen, H.M. Hakola, H. Liang, T.A. Ujula, J. J. Valle-Delgado, M. Osterberg, M.L. Yliperttula, H. Skottman, Biomimetic collagen I and IV double layer Langmuir-Schaefer films as microenvironment for human pluripotent stem cell derived retinal pigment epithelial cells. *Biomaterials* **51**, 257–269 (2015)
71. Y. Dayani, N. Malmstadt, Liposomes with double-stranded DNA anchoring the bilayer to a hydrogel Core. *Biomacromolecules* **14**, 3380–3385 (2013)
72. I.G. Denisov, Y.V. Grinkova, A.A. Lazarides, S.G. Sligar, Directed self-assembly of monodisperse phospholipid bilayer nanodiscs with controlled size. *J. Am. Chem. Soc.* **126**, 3477–3487 (2004)
73. F. Hagn, G. Wagner, Structure refinement and membrane positioning of selectively labeled OmpX in phospholipid nanodiscs. *J. Biomol. NMR* **61**, 249–260 (2015)
74. R.P. Richter, J.L.K. Him, A. Brisson, Supported lipid membranes. *Mater Today* **6**, 32–37 (2003)
75. J.A. Jackman, W. Knoll, N.-J. Cho, Biotechnology applications of tethered lipid bilayer membranes. *Materials* **5**, 2637–2657 (2012)
76. E.-K. Sinner, U. Reuning, F.N. Kok, B. Sacc  , L. Moroder, W. Knoll, D. Oesterheld, Incorporation of integrins into artificial planar lipid membranes: characterization by plasmon-enhanced fluorescence spectroscopy. *Anal. Biochem.* **333**, 216–224 (2004)
77. F. Inci, U. Celik, B. Turken, H.O.   zer, F.N. Kok, Construction of p-glycoprotein incorporated tethered lipid bilayer membranes. *Biochem. Biophys. Rep.* (2015). doi:[10.1016/j.bbrep.2015.05.012](https://doi.org/10.1016/j.bbrep.2015.05.012)
78. S. Rebaud, O. Maniti, A.P. Girard-Egrot, Tethered bilayer lipid membranes (tBLMs): interest and applications for biological membrane investigations. *Biochimie* **107**, 135–142 (2014)
79. M. Kumar, M. Grzelakowski, J. Zilles, M. Clark, W. Meier, Highly permeable polymeric membranes based on the incorporation of the functional water channel protein. Aquaporin Z. *Proc. Natl. Acad. Sci. USA* **104**, 20719–20724 (2007)
80. M. Grzelakowski, M.F. Cherenet, Y.-X. Shen, M. Kumar, A framework for accurate evaluation of the promise of aquaporin based biomimetic membranes. *J. Membr. Sci.* **479**, 223–231 (2015)
81. P. Agre, Aquaporin water channels (Nobel lecture). *Angew. Chem. Int. Ed.* **43**, 4278–4290 (2004)
82. P.H.H. Duong, T.S. Chung, K. Jeyaseelan, A. Armugam, Z. Chen, Hong M. YangJ, Planar biomimetic aquaporin-incorporated triblock copolymer membranes on porous alumina supports for nanofiltration. *J. Membr. Sci.* **409–410**, 34–43 (2012)
83. P.S. Zhong, T.S. Chung, K. Jeyaseelan, A. Armugam, Aquaporin-embedded biomimetic membranes for nanofiltration. *J. Membr. Sci.* **407–408**, 27–33 (2012)
84. X. Li, R. Wang, F. Wicaksana, C. Tang, J. Torres, A.G. Fane, Preparation of high performance nanofiltration (NF) membranes incorporated with aquaporin Z. *J. Membr. Sci.* **450**, 181–188 (2014)
85. E. Munch, M.E. Launey, D.H. Alsem, E. Saiz, A.P. Tomsia, R.O. Ritchie, Tough, bio-inspired hybrid materials. *Science* **322**, 1516–1520 (2008)
86. A. Hagenau, M.H. Suhrea, T.R. Scheibel, Nature as a blueprint for polymer material concepts: Protein fiber-reinforced composites as holdfasts of mussels. *Prog. Polym. Sci.* **39**, 1564–1583 (2014)
87. A. Miserez, J.C. Weaver, O. Chaudhurid, Biological materials and molecular biomimetics – filling up the empty soft materials space for tissue engineering applications. *J. Mater. Chem. B* **3**, 13–24 (2015)
88. M.E. Launey, R.O. Ritchie, On the fracture toughness of advanced materials. *Adv. Mater.* **21**, 2103–2110 (2009)

89. G. Mayer, Rigid biological systems as models for synthetic composites. *Science* **310**, 1144–1147 (2005)
90. N. Funk, M. Vera, L.J. Szewciw, F. Barthelat, M.P. Stoykovich, F.J. Vernerey, Bioinspired fabrication and characterization of a synthetic fish skin for the protection of soft materials. *ACS Appl. Mater. Interfaces* **7**, 5972–5983 (2015)
91. D.G. Elliott, In the laboratory fish. In: G.K. Ostrander (ed.) Academic Press, San Diego, Chap. 5, p. 99 (2000)
92. D. Zhu, L. Szewciw, F. Vernerey, F. Barthelat, Puncture resistance of the scaled skin from striped bass: collective mechanisms and inspiration for new flexible armor designs. *J. Mech. Behav. Biomed. Mater.* **24**, 30–40 (2013)
93. U.B. Sleytr, B. Schuster, E.-M. Egelseer, D. Pum, S-layers: principles and applications. *FEMS Microbiol. Rev.* **38**, 823–864 (2014)
94. B. Schuster, U.B. Sleytr, Biomimetic interfaces based on S-layer proteins, lipid membranes and functional biomolecules. *J. R. Soc. Interface* **11**, 20140232 (2014)

Chapter 22

Quantum Dots in Bionanotechnology and Medical Sciences: Power of the Small

İ. Ergal and A.T. Akarsubasi

Abstract Among the other bionanotechnology and nanomedicine tools Quantum Dots (QDs) have pioneered numerous advances in both fundamental and applied sciences; due to their unique optical, electronic and physical properties namely photostability, single source excitation, narrow emission, multiplexing capabilities and high quantum yield as compared to traditional fluorophore materials. New generation QDs with better biocompatibility allows studying intracellular processes in human beings at the single molecule level, to use in vitro imaging for bionanotechnological purposes, long term in vivo observation of cellular events, tumor targeting and even in diagnostics. However, further developments are necessary for robust and reproducible methods to conjugate QDs with many different biomolecules under stringent biological requirements to improve its commercial availability and further use in clinical diagnostics.

22.1 Introduction

Quantum Dots (QDs) have distinct advantages over traditional fluorescent organic dyes and green fluorescent proteins on account of their longer photobleaching capacity, spectral overlapping and low signal intensity [1–3].

Principally for early diagnostics, the requirement for highly sensitive and specific biological information has pushed the need for such nano crystal structures like QDs to be used in clinical applications. Because of their unique properties, QDs

İ. Ergal · A.T. Akarsubasi (✉)
Molecular Biology and Genetics Department,
İstanbul Technical University, Maslak, İstanbul, Turkey
e-mail: akarsubasi@itu.edu.tr

İ. Ergal
e-mail: ipeker@itu.edu.tr

have become one of the most investigated subject in biology and medicine. QDs have already been used in many areas of biotechnology particularly, *in vitro* and *in vivo* imaging of sentinel lymph nodes [4, 5], tumor-specific receptors [6, 7], malignant tumor detectors [8], tumor immune responses [9].

22.2 Types and Characteristics of QDs

QDs are semi conductor nanocrystals which shows particular optical properties. Unique capabilities of QDs are mainly coming from five distinct properties. First, nano-scale size of these crystals ranging from 1 to 20 nm make them efficient probes used in biological and medical applications. Second, their size-tunable fluorescence defined as ‘quantum confinement effect gives them a wide range of absorption and emission wavelengths from visible to the near infrared (NIR). Thirdly, combination of tunable-size for variety of fluorescence, with large Stokes shifts and narrow emission spectrum make simultaneous excitation of multiple colors possible. Fourth, QDs have high signal intensity due to huge absorption extinction coefficients and high quantum yield. Last but not least, since they are inorganic, they exhibit more resistance for photobleaching called long-term photostability [10–13]. Discovery of QDs begins with first observation of them in glass crystals and pioneered by Efros and Ekimov in early 1980s [14].

The typical QDs consist of a semiconductor core composed of II–IV group of elements where CdX is the most investigated QDs due to their excellent optical and electrochemical properties.

Advancement of the science and technology has led to many research projects on QDs synthesis and applications which was driven after mid-eighties. However, it took nearly two decades of promotion in QD research until the successful synthesis of colloidal CdX (X = S, Se, Te) QDs [15].

One of the major drawbacks which restricts the application of QDs for *in vivo* bionanotechnological research is the toxicity associated with heavy metals; particularly the cadmium, lead, or arsenic containing QDs.

If the surfaces (shells) are not properly covered or protected by ligands, the toxic heavy metal ions can easily be permeated into biological systems.

Evolving alternative designs of fluorescent nanocrystals with high biocompatibility has become crucial because of the limitation of heavy metal containing QDs. Basically, a few methods are commonly used, one of them is to generate highly fluorescent nanoparticles with nontoxic elements or to explore luminescent conjugated polymer dots.

Recently, scientists also focused on fluorescent polymer encapsulation and functionalization to yield conjugated polymer dots (CP dots) for bioimaging [16, 17]. CP dots which have appropriate absorption range (350–550 nm) for fluorescence microscopy and laser excitation, can exhibit higher brightness than any other nanoparticles under the same conditions [18]. Another type of QDs is the silicon nanoparticles (<10 nm) which has unique properties such as resistance to

photobleaching and wide emission range with relatively high quantum yields. Silicon based QDs (Si QDs) have advantages over traditional fluorescent organic dyes when they are used as fluorescent probes for bioimaging due to their nontoxic structure [19]. It has been reported that Si QDs are 10 times safer than Cd-containing QDs [20].

22.3 Advantages and Disadvantages of QDs

Quantum dots are nanoscale light-emitting crystals that are emerging as a new class of biological probes (specifically in the imaging field). They could replace organic dyes and fluorescent proteins. The fundamental advantages of using quantum dots are their high quantum yield and strong emission intensity, in addition to the broad emission spectrum as a function of tunable particle size [21].

The QDs can be excited by UV light and emit light, changing from visible to infrared wavelengths. Their broad excitation spectrum and narrow emission spectrum with almost no spectral overlap makes them attractive for high resolution imaging of multiple species at the same time without the need for complex optic detection systems.

QDs have higher signal-to-noise ratios compared to traditional fluorochromes. The high sensitivity of the QDs allows precise detection even in the presence of strong autofluorescent signals especially during *in vivo* imaging [22, 23]. During the live cell imaging long term monitoring of the biological sample has been enabled by QDs because of their resistance to photobleaching. Additionally, QDs can be used in multicolor optical coding for high throughput screening and also used in microarrays, bioimaging, FACS (fluorescence-activated cell sorting) analysis, and monitoring pharmacokinetics as contrast agents [24].

One of the major limitations of quantum dots can emerge from the stability of the core shell structure. Most commonly used materials in commercially available QDs are a core of CdSe and a shell of ZnS. In order to provide hydrophilic character for better biocompatibility for this inorganic structure, amphiphilic polymers are used to cap the shell layer and provide reactive sites for further linking to proteins. Therefore the stability of the shell layer provides better control on solubility and possible release of core materials (Cd ions) to the surrounding which may result in toxicity [25]. That is why QDs have limited immediate clinical use, but rather has focused applications on animal testing and *in vitro* assay developments. However, before we move forward toward widespread use of quantum dots in biological systems, further research is needed [22], since its toxicity to humans is still being debated.

Despite, size dependent tuning methodology mentioned as an advantage in terms of covering multi-color in the system; due to the production strategy, different hydrodynamic sizes makes such multiplexing procedures difficult in biological systems.

Another disadvantage is the blinking property of the QDs. They tend to blink at the single dot level that exhibit some limitation in absolute fluorescence

quantification. However, when used for imaging of biomarkers blinking doesn't affect much as there are very many of quantum dots in a sample to allow proper averaging.

22.4 Synthesis of QDs

Over three decades ago, the first QD synthesis has been reported. Since then variety of methods have been developed for synthesis of QDs. QDs can be manufactured with different sizes starting from few nanometers up to few micrometers and size distribution can be controlled using specific growth techniques under high temperatures. QDs can be designed in a way that, product can be in bare core; core/shell or core/shell/shell architecture. One of the main parameters over here besides the size distribution is the choice of shell and coatings. This has particular importance since the shell stabilizes the nanocrystal core and gives specific photophysical properties.

These methods can be classified broadly into two types, based on the solvent used in production processes: the organometallic method [26–29] and the aqueous method [30–32].

In organometallic synthesis, high boiling point and high coordinating capability media such as trioctylphosphine oxide (TOPO) is used with both metal and chalcogen elements (Oxygen family from group 6 in the periodic table). QDs synthesized with this methodology have high fluorescence quantum yields due to perfect crystal structures. Another advantage of this method is to provide narrow size distribution. Yet, in typical organometallic method, QDs are synthesized with the hydrophobic surface which is the major drawback for biological applications. Changing the surface properties of QDs with coating with a water-soluble shell and ligand-exchange is resulted with significant loss in fluorescence signal.

One of the most important disadvantages of the organometallic synthesis is the cost and troublesome process; however, still it is considered as one of the main synthesis methodology due to high quantum yield and better means of controlled size distribution.

On the other hand, the aqueous synthesis has some advantages over organometallic synthesis such as being environment-friendly and has relatively low cost production procedure resulting from improved simplicity, reproducibility and less toxicity. Therefore the aqueous synthesis has gained better reputation as preferred method for synthesis of QDs, in spite of the lower quantum yields and broader size distribution [33, 34].

Especially, with the properties such as higher stability and better compatibility in biological environment, hydrophilic thiol-capped QDs are become much popular compared to those prepared in organic solvent. Various methods have been reported to synthesize the thiol-capped QDs, including hydrothermal synthesis and microwave-assisted irradiation [35–38].

In the hydrothermal synthesis method, an airtight container which contains all reaction reagents is heated to supercritical temperature under the high pressure. Thus, the reaction time and surface defects of QDs can be efficiently reduced in these conditions [39]. Since, the reaction temperature is high; the processes of nucleus formation and crystal growth can be separated. As a result QDs with narrow size distribution produced and the existing problem in the traditional aqueous synthesis can be overcome.

Another attractive method for QDs synthesis is the microwave-assisted irradiation. Almost a decade ago, CdTe and CdSe-CdS have been simply synthesized by QDs by microwave-assisted irradiation method [40]. Microwave irradiation method is based on a heating source which is used to optimize synthesis conditions. Water is the most suitable solvent in this system. The quantum yields of QDs are able to reach as high as 17 % with the quick heating the whole solution (above 100 °C). Production of QDs with aqueous methods enables the use of QDs in biological researches without additional procedures such as ligand exchange.

22.5 Toxicity

For the application of QDs in biological studies, safety, and biocompatibility issues must be addressed. Potential adverse human health effects have been suggested depending on toxicological studies, but the fundamental cause-effect relationships are still not clear [41].

QDs are toxic mainly coming from in vitro studies where cadmium-containing particles killed cells in culture. Thus, many researchers are believed that QDs will never be used for diagnostics and treating patients.

It has been assumed that since QDs are toxic to cells, they must also be toxic to humans despite there is not sufficient scientific evidence. Actually, cell-based toxicity studies are rarely transferable to more complex biological systems to be able to conclude with such extrapolation. Besides, yet they have not been tested in animal models at physiologically appropriate doses.

Therefore, it is crucial to detect their potential toxicity in cells and on organisms and to define the interaction of QDs with biological systems.

Several studies have shown that the toxicity of QDs depends on many factors which can be categorized as inherent physicochemical properties and environmental conditions [41–43]. QDs should be characterized individually by their types since QDs are heterogeneous. They vary in their core and shell composition, sizes, charges, concentrations, in their surface chemistry and functional groups, oxidation, and mechanical stability and all these variations have been involved as contributing factors to its toxicity [42]. QDs toxicity primarily depends on heavy metal based toxicity and also induce cytotoxicity generated by reactive oxygen species (ROS). Also, the quantum size effect which resulted from nanoparticle toxicity is their size dependent intracellular routing where the nanoscale particles are able to reach organelles that are inaccessible to metal ions.

Very few studies mentioned, *in vivo* QDs toxicity validation tests. The one which has been applied on rodents via bioimaging studies, helped to identify body areas at risk for toxicity. The organs of the reticuloendothelial system (RES) took up and retained the majority of given QDs [44, 45]. RES is the part of body's defense mechanism which consist of phagocytic cells located in the liver, spleen and lymph nodes. Due to such data, it is proposed that QDs mainly degraded in these organs.

Despite accumulation of QDs in the RES organs has been reported; little focus has been placed on the QDs clearance from the body. There are two excretion pathways available for metal containing particles: following renal clearance (through kidneys to urine) or through the liver's biliary system into feces. These results lead to a question: Why are QDs apparently toxic *in vitro* but seems safer *in vivo*? Actually rather than mentioning about toxicity of the QDs; the question should be "What is the lethal dose?" Recent, studies have demonstrated that QD physicochemical properties affects the cell interaction and the potential for toxicity.

Despite all the evolution of QDs over the past decade, we are still far from carrying QD-based technology from lab scale to a patient's bedside. Moreover, new strategies to engineer safer QDs, combined with advanced methodology are important steps forward in closing the gap between *in vitro* and *in vivo* testing and, at the end, try to find out the answer to QD toxicity.

22.6 Surface Modification and Functionalization

In order to apply QDs to biological studies; biocompatibility and toxicity are the main issues that must be solved. Very hydrophobic nanocrystals are generated as a result of high quantum yield QDs and these nanocrystals are only soluble in non-polar solvents such as toluene or chloroform.

Because of the need for the aqueous solutions and buffers in biological applications, their water solubility must be increased while optical properties conserved. This purpose can be achieved by modification of the surface nature while keeping the core intact. Usually, carboxylate, amino or hydroxyl like functional groups has been used to modify surfaces. Hydrophilic ligand exchange or encapsulation using hydrophobic interactions are two main approaches for the functionalization of the QD surfaces (Fig. 22.1).

Surface modification of QDs synthesized via the organometallic route can improve the aqueous solubility for further biological applications while protect them from degradation and photobleaching. Polyethyleneglycol (PEG) can provide functional groups for further bioconjugation that also helps to reduce non-specific uptake in the non-targeted organs. Many literature reports have shown that PEG coating can change the power of biodistribution of QDs in animals [46, 47].

Combination of PEG and other polymers, poly(lactic-co-glycolic acid) (PLGA) and liposomes has also been investigated for surface modification of QDs [26, 27].

Encapsulating QDs with highly biocompatible silica can render a hydrophilic surface and also facilitate the incorporation of various functional groups such as

Nonspecific binding reduces the signal-to-noise ratio and limits detection sensitivity and immunostaining specificity. Nonspecific binding can also lead to false-positive staining for biomarkers in fluids, cells, and tissues. In particular, QDs with highly negative or positive surface charges, such as surface coatings containing carboxylic acids or amines, can exhibit strong nonspecific binding to cells and tissues [30, 53, 54] as well as to proteins in serum and blood. Since most biomolecules are charged or have charged domains [55], QDs could interact electrostatically with many soluble proteins in solution or with biomolecules on the cell surface and in the cytoplasm, resulting in the nonspecific binding.

To reduce nonspecific binding, PEGs are often attached to the organic coating layer of QDs [30, 53–56]. PEGylated QDs have a nearly neutral surface charge and can maintain colloidal stability through steric repulsion between the PEG chains.

22.7 Biocompatibility in QDs (Bioconjugation)

Preparation of the carefully selected functional surface coatings is one of the most important stages to be able to use the advantage of the unique properties of QDs in biological applications. Depending on the application environment, and desired application mechanisms such as *in vivo*/*in vitro* imaging, diagnostics, biosensors etc. Numerous QD bioconjugation strategies exist (Fig. 22.2) such as covalent linkages, polyhistidine-metal-affinity coordination, electrostatic adhesion, and chemoselective ligation [57].

The most important aspects that should be taken into consideration for bioconjugation are: (i) overall dimensions of bioconjugate (a polymer coated QD with a long linker attached to a large protein might be spatially disadvantageous for energy transfer applications or intracellular measurements); (ii) functionality (conserving photophysical properties and binding properties of the biomolecule); (iii) reproducibility; and (iv) multiplexing capacity (number of biomolecules per QD or vice versa).

A pathway is provided by the proper application of these varying conjugation schemes for controlling critical criteria of the biomolecules attachment including valency, affinity, biomolecular orientation, QD-biomolecular spacing, and orientation of the target binding spot.

Nanoscale sizes of QDs allow attachment of multiple biomolecules to one QD, as well as the attachment of multiple QDs in larger biomolecules such as viruses. The first step for biocompatibility of QDs as mentioned above functionalization of the surface for water-solubility. Transferring the inorganic QDs into water-soluble and biocompatible fluorescent nanoparticles offers variety of possibilities. Once the functionalized QDs are stable in aqueous solution, many different classical as well as novel bioconjugation methods are available, ranging from random conjugation (mostly targeting primary amines or thiols with NHS-ester or maleimide-functionalized QDs) to controlled bio-orthogonal conjugation. Another technique for attaching proteins, peptides, DNA and similar biomolecules to QDs. It is also a relatively easy and stable

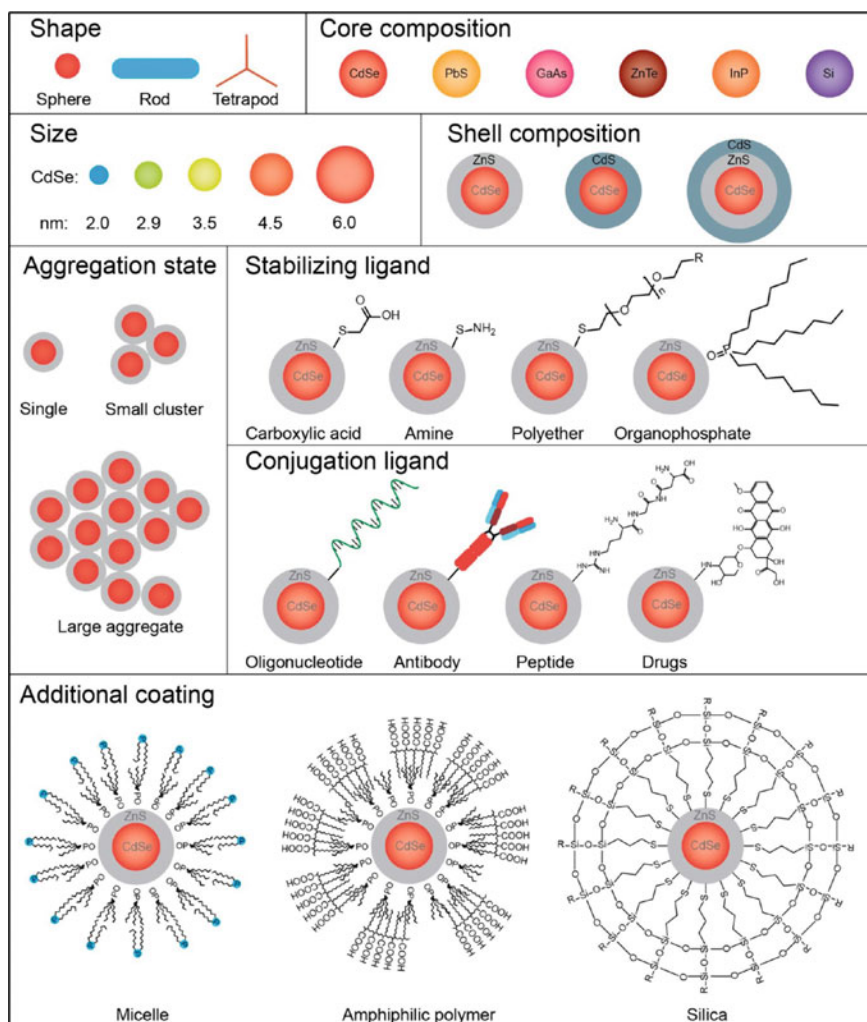


Fig. 22.2 QDs toxicity depend on physicochemical properties. Shape, core composition, size, and shell composition can be manipulated during QD synthesis. Also surface ligands can be added to increase solubilization. An additional coating can further protect the QD core from oxidation. Copyright (2012) American Chemical Society

(k_d 10^{-9} M) conjugation method. As a result of the convenient commercial availability of readily prepared QD bioconjugates and their stable binding (k_d 10^{-14} M), probably the most often applied conjugation is based on biotin–streptavidin recognition. Many other technologies such as HaloTag (Promega), SNAP-tag and copper-free click chemistry can also be used for QD-biolabeling.

22.8 Next Generation QDs (Silica, Carbon, Metal Nanocluster)

New challenge in the QD technology is improving the optimized brightness, minimized hydrodynamic sizes, resistance to nonspecific interactions, and site-specific ligand conjugation. In the case of hydrodynamic size of QD, QD-tagged proteins will be used substituted for GFP-tagged proteins inside living cells [58–60] when it is reduced to level of green fluorescent proteins (GFPs). Because of the inefficacy of nanoscale dots for optical absorbance and coating with a thin polymer layer; the challenging task is improving some protein-sized dots. New opportunities in next generation QDs for molecular medicine and bio-diagnostic applications have being performed with novel studies and related approaches.

22.8.1 *Metal Nanoclusters*

New class of fluorescent labels for biological applications which are called fluorescent metal nanoclusters, known with their ultra-small sizes, good biocompatibility, and excellent photostability, have become a good alternative to QDs. Especially, Au and Ag nanoclusters are favorable among them. Due to the extremely low quantum yield (QY), photoluminescence from the noble metals, started with little attention and much more interests have been attracted by researchers until the much enhanced QY reached [61]. Till today, a lot of Au and Ag NCs Metal nanocrystals with different scaffolds (protein, peptide, and oligonucleotide) have been developed and applied for the detection of thiol compounds [62], metal ions [56, 63], protein [64, 65], DNA [66], RNA [67] and in vivo bioimaging.

DNA stabilized silver nanoclusters have advantages in design and applications, due to easy assembly of DNA sequences; such as, Werner and colleagues designed a nanocluster beacon to detect the human Braf oncogene DNA sequence. Thus, newly developed fluorescent metal nanoclusters leads to promising opportunities for in vivo bioimaging [68].

22.8.2 *Carbon Dots (C-Dots/GQDs)*

C-dots which were first obtained in 2004 during purification of single-walled carbon nanotubes through preparative electrophoresis have excellent optical properties [69]. Thus, C-dots displayed great potentials in biological applications. Because of the high optical property of C-dots; high performance complex catalyst systems with C-dots are constructed, in order to provide the usage of full spectrum of sunlight effectively [70–73]. Due to low autofluorescence and high tissue

transparency in the near infrared (NIR) region, C-dots emission in the NIR under the excitation, is particularly significant for in vivo bionanotechnology [74, 75]. As a type of C-dots, recently the graphene QDs (GQDs) have also gained a lot of interest from researchers because of their excellent optical and electronic properties [76, 77]. Despite belonging to same family, there are some discrepancies between GQDs and C-dots [78]. The C-dots are either amorphous or crystalline, while GQDs are graphene lattices inside the dots [79]. GQDs have been employed in the production of numerous sensors and bioimaging [78–82]. Because of the high biocompatibility and optical properties, C-dots and GQDs have enhanced potentials in imaging applications of nanobiotechnology such as InP [83], InP/ZnS [84], CuInS₂/ZnS [85, 86], Ag₂Se [87], Ag₂S [88].

22.8.3 *Silicon Dots (Si QDs)*

Silicon has been known with poor optical properties for a long time. There are three distinct photoluminescence bands for Si QDs, sequentially in the infrared, in the red, and one in the blue light range. The most important advantage of Si QDs is their good biocompatibility which were claimed to be much safer than Cd-based QDs under UV irradiation (at least 10 times) [25], and even nanoscale Si has been proposed as a food additive [89]. Recently, several methods have been reported to produce optically stable, hydrophilic Si QDs [48]. However, one of the main disadvantages for their applications in bioimaging results from their oxidative degradation in vivo. Thus, surface modifications such as PEGylated micelle encapsulation and bioconjugation are necessary [22]. These Si QDs could be used in multiple cancer-related in vivo applications (sentinel lymph node mapping, tumor vasculature targeting and multicolor NIR imaging in live mice) and also showed great potentials as biocompatible fluorescent probes [49].

22.9 Applications of QDs in Bionanotechnology and Nanomedicine

22.9.1 *Cell Labeling*

One of the most common uses of QDs in bionanotechnology is labeling cells and cellular structures as fluorescent probes. Thus, earliest usage of quantum dots as more stable, brighter and multiplexing nanostructures is to label many of the intra and inter-cellular structures such as cytoskeletal proteins, organelles and elucidation of intracellular processes instead of fluorescent dyes. Organic dyes have been used for cell labelling for decades; however, multicolor labelling has been a bottleneck for such procedures until the potential use of Quantum dots. Because of their broad

absorption and narrow excitation spectra combined with retarded photobleaching makes those molecules appropriate for quantitative long term cellular imaging which reduces signal crosstalk and increases the number of labels in a single system. Long term photostability gives inarguably a significant advantage to QDs over fluorescent dyes during time scale procedures. One of the most important drawbacks in this application is QDs toxicity which has been overcome by new generation of biocompatible water soluble QDs. There are mainly two types of imaging: *in vitro* and *in vivo*.

22.9.2 *In Vitro* Imaging

QDs are used to create specific signals for biosensing depending on their unique electrochemical properties. QDs are robust fluorescent nano-sized structures that are compatible with conventional biosensing techniques used in the research and diagnostics. Using different resonance energy transfer (RET) processes means a nonradiative energy transferred without overlapping as in the excitation and emission of fluorophores. Electron exchange with overlapping orbitals depends on distance that usually exponentially decreases. Fluorescence RET (FRET), chemiluminescent RET (CRET) and bioluminescence RET (BRET) efficiencies show an inversely related to sixth power of radius distance dependence, nanometal RET (NSET) an inversely related to fourth power of radius distance dependence, and dipole to metal particle energy transfer (DMPET) mainly follows the FRET mechanism for small distances. They are therefore limited to shorter distances. Following methods demonstrate the versatility and functionality of these nanoprobes used as biosensors in nanomedicine and research. There are five main transfer mechanisms applied with QDs (Fig. 22.3).

22.9.2.1 FRET, QDs Used as Energy Donors

In other words an organic dye fluorophore can be excited efficiently at an absorption minimum of the acceptor then transferred to QDs which minimizes direct acceptor excitation. The QD is quenched in that terms, The FRET efficiency increases with the help of multiplexing QDs.

22.9.2.2 FRET QDs Used as Acceptors

The broad absorption spectrum of the QD becomes an advantage since QD can be excited at almost any wavelength; however this may lead to inefficient FRET due to small ratio of excited molecules. The problem has been solved using long excited-state lifetimes such as up to several milliseconds and time interval excitation.

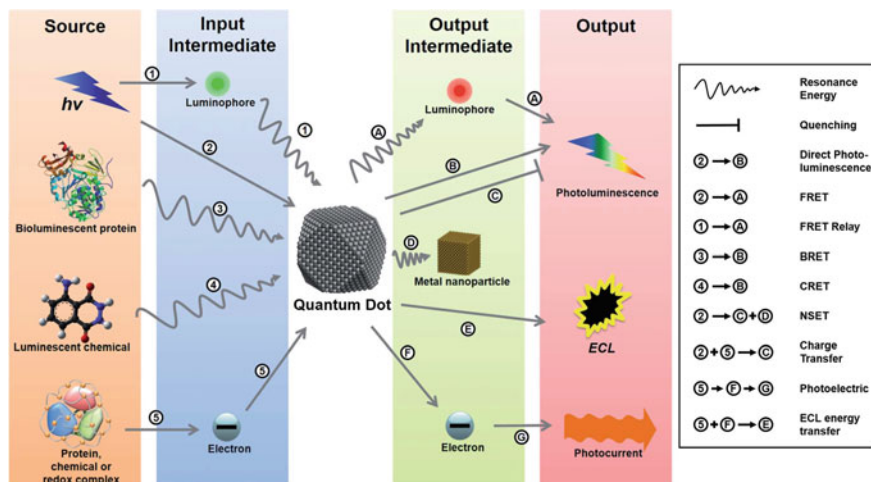


Fig. 22.3 Quantum dot assays: Illustration of the energy transfer schemes utilized in various types of QD assay paradigms. Each assay requires an energy input (electromagnetic, chemical, biochemical or electrical) that is transmitted to the QDs either directly or through an input intermediate. This energy modifies the electron behaviour of the QDs, which then release the energy, either directly or through an output intermediate, in the form of emitted light or charge transfer. Copyright (2014) Royal Society of Chemistry

22.9.2.3 CRET or BRET Using QDs as Acceptors

It is not possible to create luminescence in deep tissues with external light excitation. However, a chemical reaction like the oxidation of luminol in CRET or luciferin in BRET can be triggered at the desired position using bioluminescent donors. Thus, the excitation of the chemiluminescent or bioluminescent donor is quenched and the QD is stimulated for FRET effect that creates a smart probe.

22.9.2.4 NSET Using QDs as Donors

The advantage of energy transfer between QDs and Au-NPs mechanisms where the QD donor is quenched and the Au-NP acts as dark quencher is mainly allowing energy transfer over larger distances compared to FRET (nearly 20 nm).

22.9.2.5 CT Using QDs as Donors or Acceptors

In Charge transfer (CT), different QDs can be quenched by a single CT quencher, which gives an advantage for color multiplexing. It's mechanism basically, transferring electrons from or to proximal molecules with matching oxidation levels to or from mainly the QD surface states which leads to QD photoluminescence quenching.

For the improvement of bionanotechnological applications, smart probes development offers a promising future for QD-based energy and CT systems regarding bionanotechnology applications, especially for in vitro diagnostics (IVD) and cellular imaging [90–92].

22.9.3 *In Vivo Imaging*

There are two types of approaches here as passive targeting which depends on permeability and accumulation on the target cells or tissue. On the other hand active targeting has been applied to recognition of specific cellular components like organelles or surface proteins through ligands targetted with QDs. There are several characteristics which makes a fluorescence probe ideal for in vivo targeting such as; photostability, high biocompatibility, low toxicity and minimal rethiculoendothelial (RES) uptake. There are many reports in the literature mentioning about importance of surface coating and particle sizes effect in vivo behavior of QDs.

22.9.3.1 Peptide-Conjugated QDs

Peptide-quantum dot nanoconjugate for labeling are preferred targeting ligands because of their small sizes. Also many peptides can be simultaneously attached to a single QDs to enhance its specificity. Furthermore, ease of synthesis and low immunogenicity are some of the other advantages of peptides. The strength of peptide-target interaction is highly variable as in antibody recognition. Moreover, covalent conjugation of multiple peptides to one quantum dot which is achieved through water soluble crosslinking reagents (such as EDAC, SIA and Traut's reagent) may considerably increase the binding capacity of the resulting complexes to the target via multivalent interactions. Streptavidin-biotin noncovalent self-assembly is another route for the utilization of peptide targeted QDs. After the first in vivo peptideconjugated QDs targeted imaging was reported in 2006, the arginine-glycine-aspartate (RGD) peptide has also been widely used in nanomedicine [53]. Since Cd-free RGD peptides have better biocompatibility [83, 93, 94], they may find broader bionanotechnological future applications.

Peptide conjugated quantum dots have also been successfully used for targeting of cellular proteins such as growth factor receptors, G protein-coupled receptors, integrins, and ion channels [35, 37, 95]. In particular, quantum dots targeted to protein transduction domains namely HIV TAT, Pep-1, poly-arginine, and SV40 T antigen have been used in the research of intracellular delivery [35, 36, 96]. As a result, peptide conjugated quantum dots offer exclusive advantages over antibody-mediated targeting, and their potential as bionanotechnological probes is being actively explored.

22.9.3.2 Antibody Conjugated QDs

Antibodies are widely used as a class of diverse and specific ligands for targeted imaging and therapy. One of the simplest and quickest labeling route is to use an antibody-quantum dot conjugate, such as using antibodies for F and G proteins to detect the presence and follow the progression of respiratory syncytial viral infection [97]. Another example is using antibody-conjugated quantum dots for tumors in mice to target a prostate-specific membrane antigen *in vivo* [31]. In particular, anti-HER2 quantum dot conjugates has been used imaging breast cancer cells both *in vivo* [33] and *in vitro* [32]. Biotinylated antibodies can also be used with streptavidin-coated quantum dots or they can be directly conjugated to the quantum dot. Even a simple commercial kit is available for linking antibodies to AMP quantum dots requiring no prior quantum dot experience (Invitrogen). Besides, a wide range of antibody-quantum dot conjugates is also commercially available.

Despite many monoclonal antibodies have been used for tumor targeting and imaging, disadvantages to this approach include the relatively large size of antibodies which limits the number of ligands attached to the surface of each QD. This size affects the penetration into solid tumors. Alternatively, antibody heavy- and light- chain variable domains connected by a flexible peptide linker, is a much smaller targeting ligand than intact antibodies (25 vs. 150 kDa), which can maintain high binding affinity and specificity to the antigen. Nonetheless, antibody-quantum dot conjugates are one of the most common methodology of the choice for the most of the quantum dots in bionanotechnology literature.

22.9.3.3 Ligand Conjugated QDs

Targeting the QDs directly to the binding site of the protein of interest via ligands (also called as small molecules) can initiate a dynamic process and enable highly selective faster monitoring [98]. Using ligand-conjugated QD helps to study factors regulating protein expression and cell surface mobility, besides allosteric modulation of the binding site which allows a multiplexing assay for drug discovery. Variety of ligand-conjugated to a different size (color) QDs can be used in such assays. Also, displacement assays are possible in which the binding site targeted test compound will displace the probe, enabling a fluorescence-based assay for drug discovery.

Major disadvantage of the methods with ligand-conjugated QD is the synthesis of an optimized ligand that requires fairly sophisticated organic chemistry to design (Fig. 22.4).

22.9.3.4 High-Affinity Fusion Tag Targeting Approaches

Recently new approaches based on high-affinity fusion tags labelled QDs for cellular targets have been reported. A genetically engineered high-affinity fusion tag

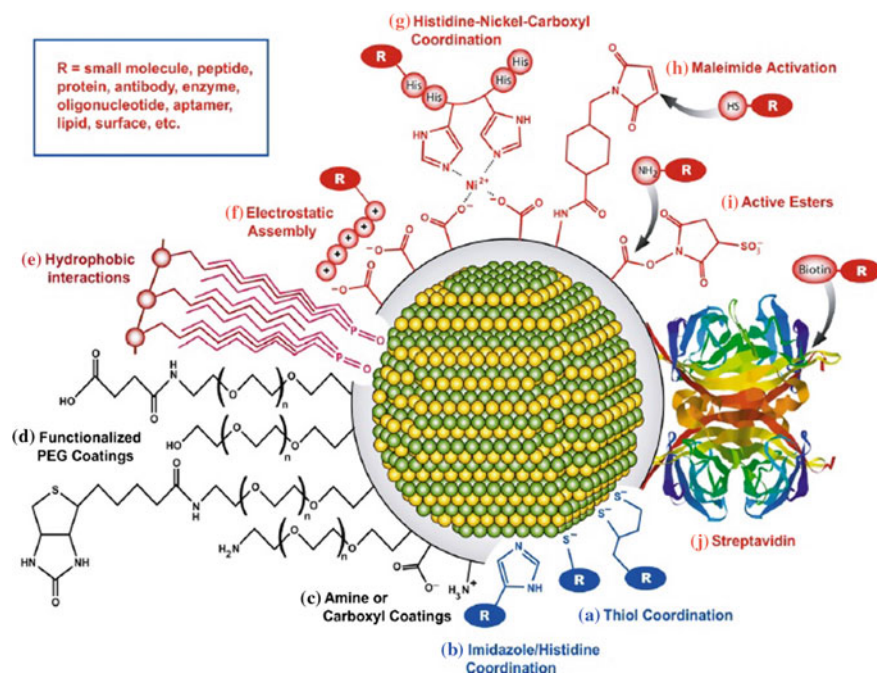


Fig. 22.4 Bioconjugation schemes for QD functionalization including (a) thiol coordination; (b) polyhistidine groups; (c) amine/carboxyl moieties; (d) PEG; (e) hydrophobic interactions; (f) electrostatic adhesion; (g) nickel mediated polyhistidine; (h) maleimide covalent linkage; (i) active esters; and (j) biotin-streptavidin complexes. Figure used with permissions of publisher, Copyright (2014) Sensors

that is usually a short peptide sequence with a recognition site, is used to express a cellular target for further labelling with QDs [99, 100]. For instance, Dahan and colleagues used trisnitriloacetic acid quantum dot conjugate which has a sub-nanomolar affinity for linear polyhistidine motifs to target a decahistidine tag fused to the N-terminus of type-1 interferon receptor subunit [101]. There are other studies that have used fusion tags in conjunction with quantum dots include CrAsH [102] and Halotag [103].

22.9.3.5 Nonspecific Binding

One of the main drawback of quantum dots is nonspecific cellular binding in a range of in vivo nanobiotechnology applications, such as live cell imaging and tissue staining.

Quantum dots may show nonspecific cellular binding that has been observed in several systems and primarily attributed to electrostatic interactions in between

probes surface (e.g., carboxyl) and target cell surface [30, 104]. Therefore, researchers must rigorously examine selected QDs for nonspecific interactions with the surface functional groups and should choose a modification approach that results in minimal nonspecific binding. Moreover, the issue of nonspecific binding should not be taken into consideration in the presentation of research results.

22.10 Dual-Modality Imaging with QDs

Imaging techniques and particularly preclinical imaging systems has been improved massively over the last decade. However, yet no single imaging technique is perfect to obtain all the necessary information for a given study. For instance, one of the main challenges faced in fluorescence imaging is the quantification and deep tissue penetration. Magnetic resonance imaging (MRI) has outstanding soft tissue contrast and good resolution, but rather poor sensitivity. On the other hand; Positron emission tomography (PET) is superior in sensitivity, quantitation, and tissue penetration, nonetheless has relatively low resolution. Thus, in many clinical applications to overcome these disadvantages single probe using multiple imaging techniques has been applied to provide more detailed information.

QD-based nanoprobe will be better tools for these applications rather than traditional fluorophores.

22.10.1 Fluorescence/MRI

Using the RGD peptides as ligands for tumor targeting with Paramagnetic gadolinium coated QDs and PEGylated lipids to develop a dual-modality probe for MRI has been reported by Mulder et al. in 2006 [105]. The high relativity, and the specificity of this nano-probe makes it an excellent contrast agent for nanobiomedical imaging purposes Although it has been confirmed by several studies that the safety and tumor targeting capability of these nanocarriers, more studies required to fully realize its potential for dual-modality imaging. Recently, a core/shell nanoprobe was investigated for dual-modality imaging of breastcancer, which was composed of an IO NP core [106] used with MRI to get detailed anatomical structure of the tumor.

22.10.2 Fluorescence/CT

One of the most reliable and fast diagnostic tool is X-ray computed tomography (CT) in many of the clinical applications which also has high spatial resolution. Iodinated molecules and compounds used as contrast agents for CT; however, these

contrast agents have short lifetime in the blood or lymphatic vessels, which is a major disadvantage. Recently, QDs that have been combined with iodinated molecules can potentially overcome such disadvantages and leads to better future promising results [107]. For instance, diagnostic tools in cardiovascular diseases, as one of the leading cause of death worldwide, is very important specifically in clinical settings for improving patient's onset [108, 109]. QDs embedded inside hydrophobic iodinated oil core was shown to be a good alternative for targeting macrophages in atherosclerotic plaques both in vitro and in vivo.

22.10.3 Fluorescence/PET

PET is widely used imaging technique in cancer diagnosis, evaluation and during the treatment follow up; due to its high sensitivity, good quantitation capability, and superb tissue penetration. The first targeted dualmodality fluorescence/PET probe based on QDs was reported in 2007 that was used RGD peptide as the targeting ligand for the DOTA (1,4,7,10-tetraazacyclododecane-N,N',N'',N'''-tetraacetic acid) chelator (DOTA-QD-RGD). The results between the ones measured by in vivo PET imaging and those measured by ex vivo NIRF imaging or tissue homogenate fluorescence were in good correlation. Thereafter, pathology examination has shown that DOTA-QD-RGD targets primarily the tumor vasculature with little extravasation similar as the previous report using RGD conjugated QDs without ^{64}Cu -labeling [95].

22.11 Conclusions

In this review, selected properties, applications and performance of QDs have been discussed specifically in the context of bionanotechnology and nanomedicine. Among the other bionanotechnological tools these nanocrystal structures have led the way for numerous advances in both fundamental and applied sciences.

Main reason for these advancements is that; QDs exhibit unique optical, electronic and physical properties as compared to traditional fluorophore materials namely photostability, single source excitation, narrow emission, multiplexing capabilities and high quantum yield. Recently, significant progress has been made in the synthesis of QDs and their biocompatibilities in terms of both theoretical models and experimental results.

Quantum dots, specifically those which are called new generation, offer more future promising application potential than both traditional fluorescent probes and also CdSe QDs to study intracellular processes in human beings at the single molecule level. Since they have better biocompatibility, these QDs can be used for in vitro imaging in bionanotechnological applications, long term in vivo

observation of cellular events, tumor targeting and even in diagnostics. Yet, in the areas like labeling of proteins, tagging nucleic acids and in DNA moments an extensive research is going on.

Biomedical applications of QDs can serve in an extensive area of research and applications, such as microscopy, multiplexed histology, flow cytometry, drug delivery, in vivo long term whole organism and clinical imaging, photodynamic therapy, tissue mapping and demarcation, real time detection of intra and inter-cellular events, signaling biosensing, environment and biodefense etc.

What can we expect for the future clinical applications of quantum dots?

The requirement for collecting more biological information from extremely low concentrations of samples from any organisms (human being till microbiological specimens), principally for early diagnostics or during the treatment processes, has pushed the need for such highly sensitive and capability of multiplexing molecules like quantum dots.

However, the lack of robust and reproducible methods to conjugate many different biomolecules such as antibodies, protein markers, DNA and RNA to QDs in a systematic manner with control over their ratio, orientation, and its noncovalent interactions continues to retard their further use in clinical diagnostics. It may seem like QDs are slowly improving, however, those unresolved problems leads to extensive new research areas in science. As a result of such efforts, we may expect more commercial products incorporating QDs for clinical diagnostic, and research purposes, such as particles that can be used for dual-modality imaging.

We believe that the future of quantum dots in life sciences is as bright as the quantum dots themselves.

Acknowledgments I would like to give my sincere gratitudes to Prof. Dr. Hilmi Ünlü for his critical reviews. I would like to express my sincere thanks to all my colleagues especially, Mss. Busra Ahata (BSc), Mr. Mert Kumru (MSc) and Mr. Halil Kurt (PhD) for their valuable contributions throughout the preparation of this chapter.

References

1. W.C.W. Chan, S.M. Nie, Quantum dot bioconjugates for ultrasensitive nonisotopic detection. *Science* **281**(5385), 2016–2018 (1998)
2. B. Dubertret, P. Skourides, D.J. Norris, V. Noireaux, A.H. Brivanlou, A. Libchaber, In vivo imaging of quantum dots encapsulated in phospholipid micelles. *Science* **298**(5599), 1759–1762 (2002)
3. A. Konkar, S.Y. Lu, A. Madhukar, S.M. Hughes, A.P. Alivisatos, Semiconductor nanocrystal quantum dots on single crystal semiconductor substrates: high resolution transmission electron microscopy. *Nano Lett.* **5**(5), 969–973 (2005)
4. A.M. Smith, S. Dave, S.M. Nie, L. True, X. Gao, Multicolor quantum dots for molecular diagnostics of cancer. *Expert Rev. Mol. Diagn.* **6**(2), 231–244 (2006)
5. B. Ballou, L.A. Ernst, S. Andreko, Sentinel lymph node imaging using quantum dots in mouse tumor models. *Bioconjug. Chem.* **18**(2), 389–396 (2007)
6. P. Diagaradjane, J.M. Orenstein-Cardona, N.E. Colon-Casasnovas, Imaging epidermal growth factor receptor expression in vivo: pharmacokinetic and biodistribution

- characterization of a bioconjugated quantum dot nanoprobe. *Clin. Cancer Res.* **14**(3), 731–774 (2008)
7. D.S. Lidke, P. Nagy, R. Heintzmann, D.J. Arndt-Jovin, J.N. Post, H.E. Grecco, E.A. Jares-Erijman, T.M. Jovin, Quantum dot ligands provide new insights into erbB/HER receptor-mediated signal transduction. *Nat. Biotech.* **22**, 198–203 (2004)
 8. L.F. Qi, X.H. Gao, Quantum dot—amphipol nanocomplex for intracellular delivery and real-time imaging of siRNA. *ACS Nano* **2**(7), 1403–1410 (2008)
 9. D. Sen, T.J. Deerinck, M.H. Ellisman, I. Parker, M.D. Cahalan, Quantum dots for tracking dendritic cells and priming an immune response in vitro and in vivo. *PLoS One* **3**(9) (2008)
 10. P. Alivisatos, The use of nanocrystals in biological detection. *Nat. Biotechnol.* **22**(1), 47–52 (2004)
 11. W.R. Algar, A.J. Tavares, U.J. Krull, Beyond labels: a review of the application of quantum dots as integrated components of assays, bioprobes, and biosensors utilizing optical transduction. *Anal. Chim. Acta* **673**(1), 1–25 (2010)
 12. S.J. Rosenthal, J.C. Chang, O. Kovtun, J.R. McBride, I.D. Tomlinson, Biocompatible quantum dots for biological applications. *Chem. Biol.* **18**(1), 10–24 (2011)
 13. A.M. Smith, H.W. Duan, A.M. Mohs, S.M. Nie, Bioconjugated quantum dots for in vivo molecular and cellular imaging. *Adv. Drug Deliv. Rev.* **60**(11), 1226–1240 (2008)
 14. A.I. Ekimov, A.A. Onushchenko, Quantum size effect in three-dimensional microscopic semiconductor crystals. *JETP Lett.* **34**(6), 345–349 (1981)
 15. C.B. Murray, D.J. Norris, M.G. Bawendi, Synthesis and characterization of nearly monodisperse CdE (E = sulfur, selenium, tellurium) semiconductor nanocrystallites. *J. Am. Chem. Soc.* **115**(19), 8706–8715 (1993)
 16. C.F. Wu, C. Szymanski, J. McNeill, Preparation and encapsulation of highly fluorescent conjugated polymer nanoparticles. *Langmuir* **22**(7), 2956–2960 (2006)
 17. C.F. Wu, C. Szymanski, Z. Cain, J. McNeill, Conjugated polymer dots for multiphoton fluorescence imaging. *J. Am. Chem. Soc.* **129**(43), 12904–12905 (2007)
 18. C.F. Wu, B. Bull, C. Szymanski, K. Christensen, J. McNeill, Multicolor conjugated polymer dots for biological fluorescence imaging. *ACS Nano* **2**(11), 2415–2423 (2008)
 19. S.T. Selvan, Silica-coated quantum dots and magnetic nanoparticles for bioimaging applications. *Biointerphases* **5**(3), FA110–FA115 (2010)
 20. X.H. Gao, S.M. Nie, Doping mesoporous materials with multicolor quantum dots. *J. Phys. Chem. B* **107**(42), 11575–11578 (2003)
 21. I.L. Medintz, A.R. Clapp, H. Mattoussi, E.R. Goldman, B. Fisher, J.M. Mauro, *Nat. Mater.* **2**, 630–638 (2003)
 22. F. Erogbogbo, K.T. Yong, I. Roy, In vivo targeted cancer imaging, sentinel lymph node mapping and multichannel imaging with biocompatible silicon nanocrystals. *Nano Lett.* **5**(1), 413–423 (2011)
 23. B. Bruhn, Fabrication and characterization of single luminescing quantum dots from 1D silicon nanostructures. Doctoral Thesis (2012)
 24. J.H. Park, L. Gu, G. VonMaltzahn, E. Ruoslahti, S.N. Bhatia, M.J. Sailor, Biodegradable luminescent porous silicon nanoparticles for in vivo applications. *Nat. Mater.* **8**(4), 331–336 (2009)
 25. J.H. Shen, Y.H. Zhu, X.L. Yang, C.Z. Li, Graphene quantum dots: emergent nanolights for bioimaging, sensors, catalysis and photovoltaic devices. *Chem. Commun.* **48**, 3686–3699 (2012)
 26. P. Pierobon, S. Achouri, S. Courty, A.R. Dunn, J.A. Spudich, M. Dahan, G. Cappello, Velocity, processively, and individual steps of single myosin V molecules in live cells. *Biophys. J.* **96**, 4268–4275 (2009)
 27. Y. Ishihama, T. Funatsu, Single molecule tracking of quantum dot labeled mRNAs in a cell nucleus. *Biochem. Biophys. Res. Commun.* **381**, 33–38 (2009)
 28. K.M. Tsoi, Q. Dai, B.A. Alman, W.C. Chan, Are quantum dots toxic? Exploring the discrepancy between cell culture and animal studies. *Acc. Chem. Res.* **46**(3), 662–671 (2012)

29. G. Ruan, A. Agrawal, A.I. Marcus, S. Nie, Imaging and tracking of Tat peptide-conjugated quantum dots in living cells: new insights into nanoparticle uptake, intracellular transport, and vesicle shedding. *J. Am. Chem. Soc.* **129**, 14759–14766 (2007)
30. E.L. Bentzen, I.D. Tomlinson, J. Mason, P. Gresch, M.R. Warnement, D. Wright, E. Sanders-Bush, R. Blakely, S.J. Rosenthal, Surface modification to reduce nonspecific binding of quantum dots in live cell assays. *Bioconjug. Chem.* **16**, 1488–1494 (2005)
31. X. Gao, Y. Cui, R.M. Levenson, L.W.K. Chung, S. Nie, In vivo cancer targeting and imaging with semiconductor quantum dots. *Nat. Biotech.* **22**, 969–976 (2004)
32. X. Wu, H. Liu, J. Liu, K.N. Haley, J.A. Treadway, J.P. Larson, N. Ge, F. Peale, M. P. Bruchez, Immunofluorescent labeling of cancer marker Her2 and other cellular targets with semiconductor quantum dots. *Nat. Biotech.* **21**, 41–46 (2002)
33. H. Tada, H. Higuchi, T.M. Wanatabe, N. Ohuchi, In vivo real-time tracking of single quantum dots conjugated with monoclonal anti-HER2 antibody in tumors of mice. *Cancer Res.* **67**, 1138–1144 (2007)
34. S. Pathak, M.C. Davidson, G.A. Silva, Characterization of the functional binding properties of antibody conjugated quantum dots. *Nano Lett.* **7**, 1839–1845 (2007)
35. R.L. Orndorff, S.J. Rosenthal, Neurotoxin quantum dot conjugates detect endogenous targets expressed in live cancer cells. *Nano Lett.* **9**, 2589–2599 (2009)
36. R.L. Orndorff, M.R. Warnement, J.N. Mason, R.D. Blakely, S.J. Rosenthal, Quantum dot ex vivo labeling of neuromuscular synapses. *Nano Lett.* **8**, 780–785 (2008)
37. T.Q. Vu, R. Maddipati, T.A. Blute, B.J. Nehilla, L. Nusblat, T.A. Desai, Peptide-conjugated quantum dots activate neuronal receptors and initiate downstream signaling of neurite growth. *Nano Lett.* **5**, 603–607 (2005)
38. I.D. Tomlinson, J.N. Mason, R.D. Blakely, S.J. Rosenthal, Peptide-conjugated quantum dots. *Methods Mol. Biol.* **303**, 51–60 (2005)
39. D. Sparks, T. Hubbard, Micromachined needles and lancets with design adjustable bevel angles. *J. Micromech. Microeng.* **14**(8), 1230 (2004)
40. M.J. Cloninger, Biological applications of dendrimers. *Curr. Opin. Chem. Biol.* **6**(6), 742–748 (2002)
41. A.M. Derfus, W.C.W. Chan, S.N. Bhatia, Probing the cytotoxicity of semiconductor quantum dots. *Nano Lett.* **4**(1), 11–18 (2004)
42. R. Hardman, Toxicological review of quantum dots: toxicity depends on physicochemical and environmental factors. *Environ. Health Perspect.* **114**(2), 165–172 (2006)
43. Y. Pan, S. Neuss, A. Leifert, Size-dependent cytotoxicity of gold nanoparticles. *Small* **3**(11), 941–1949 (2007)
44. X. Michalet, F.F. Pinaud, L.A. Bentolila, Quantum dots for live cells, in vivo imaging, and diagnostics. *Science* **307**(5709), 538–544 (2005)
45. P.N. Prasad, *Biophotonics* (Wiley-Interscience, Hoboken, NJ, USA, 2003)
46. S.M. Smith, R. Renden, H. von Gersdorff, Synaptic vesicle endocytosis: fast and slow modes of membrane retrieval. *Trends Neurosci.* **31**, 559–568 (2008)
47. S. An, D. Zenisek, Regulation of exocytosis in neurons and neuroendocrine cells. *Curr. Opin. Neurobiol.* **14**, 522–530 (2004)
48. H.T. Li, Z.H. Kang, Y. Liu, S.T. Lee, Carbon nanodots: synthesis, properties and applications. *J. Mater. Chem.* **22**, 24230–24253 (2012)
49. H. Ming, Z. Ma, Y. Liu, K.M. Pan, H. Yu, F. Wang, Z.H. Kang, Large scale electrochemical synthesis of high quality carbon nanodots and their photocatalytic property. *Dalton Trans.* **41**, 9526–9531 (2012)
50. A.M. Smith, S. Nie, Next-generation quantum dots. *Nat. Biotechnol.* **27**, 732–733 (2009)
51. A.M. Smith, S. Nie, Minimizing the hydrodynamic size of quantum dots with multifunctional multidentate polymer ligands. *J. Am. Chem. Soc.* **130**, 11278–11279 (2008)
52. H.S. Choi, W. Liu, P. Misra, E. Tanaka, J.P. Zimmer, Renal clearance of quantum dots. *Nat. Biotechnol.* **25**, 1165–1170 (2007)

53. B.A. Kairdolf, M.C. Mancini, A.M. Smith, S. Nie, Minimizing nonspecific cellular binding of quantum dots with hydroxyl-derivatized surface coatings. *Anal. Chem.* **80**, 3029–3034 (2008)
54. S. Pathak, S.K. Choi, N. Arnheim, M.E. Thompson, Hydroxylated quantum dots as luminescent probes for in situ hybridization. *J. Am. Chem. Soc.* **123**, 4103–4104 (2001)
55. F.B. Sheinerman, R. Norel, B. Honig, Electrostatic aspects of protein–protein interactions. *Curr. Opin. Struct. Biol.* **10**, 153–159 (2000)
56. X. Wu, H. Liu, J. Liu, K.N. Haley, J.A. Treadway, Immunofluorescent labeling of cancer marker Her2 and other cellular targets with semiconductor quantum dots. *Nat. Biotechnol.* **21**, 41–46 (2003)
57. S.F. Lim, R. Riehn, W.S. Ryu, N. Khanarian, C.K. Tung, D. Tank, R.H. Austin, In vivo and scanning electron microscopy imaging of upconverting nanophosphors in *caenorhabditis elegans*. *Nano Lett.* **6**(2), 169–174 (2006)
58. L. Li, T.J. Daou, I. Texier, Highly luminescent CuInS₂/ZnS core/shell nanocrystals: cadmium-free quantum dots for in vivo imaging. *Chem. Mater.* **21**, 2422–2429 (2009)
59. J. Park, J. Nam, N. Won, Compact and stable quantum dots with positive, negative, or zwitterionic surface: specific cell interactions and non-specific adsorptions by the surface charges. *Adv. Funct. Mater.* **21**, 1558–1566 (2011)
60. E. Muro, A. Fragola, T. Pons, Comparing intracellular stability and targeting of sulfobetaine quantum dots with other surface chemistries in live cells. *Small* **8**, 1029–1037 (2012)
61. M. Howarth, W.H. Liu, S. Puthenveetil, Monovalent, reduced-size quantum dots for imaging receptors on living cells. *Nat. Methods* **5**, 397–399 (2008)
62. X.F. Hu, T.C. Liu, Y.C. Cao, Characterization of the coupling of quantum dots and immunoglobulin antibodies. *Anal. Bioanal. Chem.* **386**, 1665–1671 (2006)
63. T. Maldiney, M.U. Kaikkonen, J. Seguin, In vitro targeting of avidin-expressing glioma cells with biotinylated persistent luminescence nanoparticles. *Bioconjug. Chem.* **23**, 472–478 (2012)
64. L. Chen, X.W. Zhang, G.H. Zhou, Simultaneous determination of human enterovirus 71 and coxsackievirus B3 by dual-color quantum dots and homogeneous immunoassay. *Anal. Chem.* **84**, 3200–3207 (2012)
65. P.M. Allen, W.H. Liu, V.P. Chauhan, InAs(ZnCdS) quantum dots optimized for biological imaging in the near-infrared. *J. Am. Chem. Soc.* **132**, 470–471 (2010)
66. P.K. Bae, K.N. Kim, S.J. Lee, The modification of quantum dot probes used for the targeted imaging of his-tagged fusion proteins. *Biomaterials* **30**, 836–842 (2009)
67. H.Y. Park, K. Kim, S. Hong, Compact and versatile nickel-nitrilotriacetate-modified quantum dots for protein imaging and forster resonance energy transfer based assay. *Langmuir* **26**, 7327–7333 (2010)
68. B.C. Lagerholm, M. Wang, L.A. Ernst, D.H. Ly, H. Liu, M.P. Bruchez, A.S. Waggoner, Multicolor coding of cells with cationic peptide coated quantum dots. *Nano Lett.* **4**(10), 2019–2022 (2004)
69. K. Susumu, I.L. Medintz, J.B. Delehanty, K. Boeneman, H. Mattoussi, Modification of poly (ethylene glycol)-capped quantum dots with nickel nitrilotriacetic acid and self-assembly with histidine-tagged proteins. *J. Phys. Chem. C* **114**, 13526–13531 (2010)
70. S.H. Bhang, N. Won, T.J. Lee, Hyaluronic acid-quantum dot conjugates for in vivo lymphatic vessel imaging. *ACS Nano* **3**, 1389–1398 (2009)
71. H.W. Yu, I.S. Kim, R. Niessner, D. Knopp, Multiplex competitive microbead-based flow cytometric immunoassay using quantum dot fluorescent labels. *Anal. Chim. Acta* **750**, 191–198 (2012)
72. J.S. Liou, B.R. Liu, A.L. Martin, Protein transduction in human cells is enhanced by cell-penetrating peptides fused with an endosomolytic HA2 sequence. *Peptides* **37**, 273–284 (2012)
73. Y. Zhang, M.K. So, A.M. Loening, HaloTag protein-mediated site-specific conjugation of bioluminescent proteins to quantum dots. *Angew. Chem. Int. Ed. Engl.* **45**, 4936–4940 (2006)

74. A. Charalambous, I. Antoniadis, N. Christodoulou, P.A. Skourides, Split-inteins for simultaneous, site-specific conjugation of quantum dots to multiple protein targets in vivo. *J. Nanobiotechnol.* **9**, 37 (2011)
75. A.M. Smith, H. Duan, A.M. Mohs, S. Nie, Bioconjugated quantum dots for in vivo molecular and cellular imaging. *Adv. Drug Deliv. Rev.* **60**, 1226–1240 (2008)
76. M. Baker, Nanotechnology imaging probes: smaller and more stable. *Nat. Methods* **7**, 957–962 (2010)
77. A.M. Smith, M.M. Wen, S. Nie, Imaging dynamic cellular events with quantum dots—the bright future. *Biochemist* **32**, 12–17 (2010)
78. A. Mooradian, Photoluminescence of metals. *Phys. Rev. Lett.* **22**(5), 185–187 (1969)
79. Z.Z. Huang, F. Pu, Y.H. Lin, J.S. Ren, X.G. Qu, Modulating DNA-templated silver nanoclusters for fluorescence turn-on detection of thiol compounds. *Chem. Commun.* **47**, 3487–3489 (2011)
80. Y.T. Su, G.Y. Lan, W.Y. Chen, H.T. Chang, Detection of copper ions through recovery of the fluorescence of DNA templated copper/silver nanoclusters in the presence of mercaptopropionic acid. *Anal. Chem.* **82**, 8566–8572 (2010)
81. G.Y. Lan, C.C. Huang, H.T. Chang, Silver nanoclusters as fluorescent probes for selective and sensitive detection of copper ions. *Chem. Commun.* **46**, 1257–1259 (2010)
82. J. Sharma, H.C. Yeh, H. Yoo, J.H. Werner, J.S. Martinez, Silver nanocluster aptamers: in situ generation of intrinsically fluorescent recognition ligands for protein detection. *Chem. Commun.* **47**, 2294–2296 (2011)
83. J.J. Li, X.Q. Zhong, H.Q. Zhang, X.C. Le, J.J. Zhu, Binding-induced fluorescence turn on assay using aptamer-functionalized silver nanocluster DNA probes. *Anal. Chem.* **84**, 5170–5174 (2012)
84. H.C. Yeh, J. Sharma, I.M. Shih, D.M. Vu, J.S. Martinez, J.H. Werner, A fluorescence light up Ag nanocluster probe that discriminates single-nucleotide variants by emission color. *J. Am. Chem. Soc.* **134**, 11550–11558 (2012)
85. H.F. Dong, S. Jin, H.X. Ju, K.H. Hao, L.P. Xu, H.T. Lu, X.J. Zhang, Trace and label-free microRNA detection using oligonucleotide encapsulated silver nanoclusters as probes. *Anal. Chem.* **84**, 8670–8674 (2012)
86. J.L. Wang, G. Zhang, Q.W. Li, H. Jiang, C.Y. Liu, C. Amatore, X.M. Wang, In vivo self bio-imaging of tumors through in situ biosynthesized fluorescent gold nanoclusters. *Sci Rep* **3**, 1157–1162 (2013)
87. X.Y. Xu, R. Ray, Y.L. Gu, H.J. Ploehn, L. Gearheart, K. Raker, W.A. Scrivens, Electrophoretic analysis and purification of fluorescent single-walled carbon nanotube fragments. *J. Am. Chem. Soc.* **126**(40), 12736–12737 (2004)
88. L. Cao, X. Wang, M.J. Meziani, F. Lu, H. Wang, P.G. Luo, Y. Lin, B.A. Harruff, L.M. Veca, D. Murray, S.Y. Xie, Y.P. Sun, Carbon dots for multiphoton bioimaging. *J. Am. Chem. Soc.* **129**(37), 11318–11319 (2007)
89. S.Y. Ding, M. Jones, M.P. Tucker, J.M. Nedeljkovic, J. Wall, M.N. Simon, G. Rumbles, M. E. Himmel, *Nano Lett.* **3**, 1581 (2003)
90. B. Hötzer et al., Fluorescence in nanobiotechnology: sophisticated fluorophores for novel applications. *SMALL* (2012)
91. K.E. Sapsford et al., Materials for fluorescence resonance energy transfer analysis: beyond traditional donor–acceptor combinations. *Angew. Chem. Int. Ed.* **45**, 4562–4588 (2006)
92. I.L. Medintz et al., Interactions between redox complexes and semiconductor quantum dots coupled via a peptide bridge. *J. Am. Chem. Soc.* **130**, 16745–16756 (2008)
93. T. Pons, H. Mattoussi, Investigating biological processes at the single molecule level using luminescent quantum dots. *Ann. Biomed. Eng.* **37**, 1934–1959 (2009)
94. J. Schlessinger, Cell signaling by receptor tyrosine kinases. *Cell* **103**, 211–225 (2000)
95. W. Cai, D.-W. Shin, K. Chen, O. Gheysens, Q. Cao, S.X. Wang, S.S. Gambhir, X. Chen, Peptide-labeled near-infrared quantum dots for imaging tumor vasculature in living subjects. *Nano Lett.* **6**, 669–676 (2006)

96. B.R. Smith, Z. Cheng, A. De, A.L. Koh, R. Sinclair, S.S. Gambhir, Real-time intravital imaging of RGD: quantum dot binding to luminal endothelium in mouse tumor neovasculature. *Nano Lett.* **8**, 2599–2606 (2008)
97. E.L. Bentzen, F. House, T.J. Utey, J.E. Crowe, D.W. Wright, Progression of respiratory syncytial virus infection monitored by fluorescent quantum dot probes. *Nano Lett.* **5**, 591–595 (2005)
98. S.J. Rosenthal, I. Tomlinson, E.M. Adkins, S. Schroeter, S. Adams, L. Swafford, J. McBride, Y. Wang, L.J. DeFelice, R.D. Blakely, Targeting cell surface receptors with ligand-conjugated nanocrystals. *J. Am. Chem. Soc.* **124**, 4586–4594 (2002)
99. M. Howarth, K. Takao, Y. Hayashi, A.Y. Ting, Targeting quantum dots to surface proteins in living cells with biotin ligase. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7583–7588 (2005)
100. M. Howarth, A.Y. Ting, Imaging proteins in live mammalian cells with biotin ligase and monovalent streptavidin. *Nat. Protoc.* **3**, 534–545 (2008)
101. V. Roullier, S. Clarke, C. You, F. Pinaud, G.R. Gouzer, D. Schaible, V.R. Marchi-Artzner, J. Piehler, M. Dahan, High-affinity labeling and tracking of individual histidine-tagged proteins in live cells using Ni²⁺ + Tris-nitrilotriacetic acid quantum dot conjugates. *Nano Lett.* **9**, 1228–1234 (2009)
102. E. Genin, O. Carion, B. Mahler, B. Dubertret, N. Arhel, P. Charneau, E. Doris, C. Mioskowski, CrAsH₃ quantum dot nanohybrids for smart targeting of proteins. *J. Am. Chem. Soc.* **130**, 8596–8597 (2008)
103. M.-K. So, H. Yao, J. Rao, HaloTag protein-mediated specific labeling of living cells with quantum dots. *Biochem. Biophys. Res. Comm.* **374**, 419–423 (2008)
104. D. Gerion, W.J. Parak, S.C. Williams, D. Zanchet, C.M. Micheel, A.P. Alivisatos, Sorting fluorescent nanocrystals with DNA. *J. Am. Chem. Soc.* **124**, 7070–7074 (2002)
105. W.J.M. Mulder, R. Koole, R.J. Brandwijk et al., Quantum dots with a paramagnetic coating as a bimodal molecular imaging probe. *Nano Lett.* **6**, 1–6 (2006)
106. Q. Ma, Y. Nakane, Y. Mori et al., Multilayered, core/shell nanoprobe based on magnetic ferric oxide particles and quantum dots for multimodality imaging of breast cancer tumors. *Biomaterials* **33**:8486–8494 (2012)
107. J. Ding, Y. Wang, M. Ma et al., CT/fluorescence dual-modal nanoemulsion platform for investigating atherosclerotic plaques. *Biomaterials* **34**, 209–216 (2013)
108. J.H. Rudd, F. Hyafil, Z.A. Fayad, Inflammation imaging in atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **29**, 1009–1016 (2009)
109. T. Temma, H. Saji, Radiolabelled probes for imaging of atherosclerotic plaques. *Am. J. Nucl. Med. Mol. Imaging* **2**, 432–447 (2012)

Chapter 23

Nanomedicine

Eda Tahir Turanlı and Elif Everest

Abstract Nanomedicine is application of nanotechnology to diagnosis, treatment and prevention of diseases, in which multidisciplinary approaches are used combining chemistry, physics, biology, genetics, and medicine. Currently, there are a number of approved nanomedicines some of which have been developed to treat diseases using two main different approaches; passive or active targeting. Others include in vitro and in vivo diagnostic nanoparticles that have been shown to have many advantages compared to the traditional methods. Although there are some safety issues regarding development and use of nanoparticles in medicine and controversies surrounding the need for nanomedicine-specific regulations; many nanomedicines have been approved to date, and further improvements are being pursued with promising success. In this chapter, we summarise the use, advantages, current and future statuses of pharmaceutical and diagnostic nanomedicines, including ethical issues and regulations.

23.1 Introduction

Nanomedicine is a multidisciplinary area, which combines chemistry, physics, biology, genetics, medicine and develops nano-sized tools to diagnose, treat, and prevent diseases (Fig. 23.1). There are several definitions of nanomedicine that slightly differ from each other. However according to the European Science Foundation's (ESF) Forward Look Nanomedicine [1], the definition is as follows:

E.T. Turanlı (✉)

Molecular Biology and Genetics Department, Istanbul Technical University, Maslak, 34469 Istanbul, Turkey
e-mail: turanlie@itu.edu.tr

E.T. Turanlı · E. Everest

Graduate School of Science Engineering and Technology, Molecular Biology-Genetics and Biotechnology Programme, Dr. Orhan Öcalgiray Molecular Biology and Genetics Research Centre, Istanbul Technical University, Maslak, 34469 Istanbul, Turkey

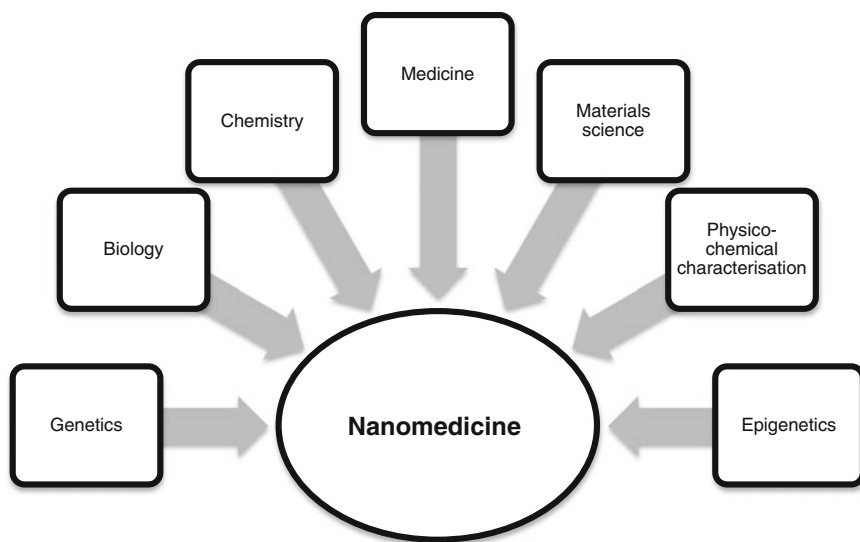


Fig. 23.1 Interdisciplinary nature of nanomedicine

Nanomedicine uses nano-sized tools for the diagnosis, prevention and treatment of disease and to gain increased understanding of the complex underlying patho-physiology of disease. The ultimate goal is improved quality-of-life.

The definition encompasses diagnostic nanoparticles such as in vitro biosensors and surgical tools, and in vivo imaging agents and monitoring technologies, as well as therapeutic biomaterials that are used for drug delivery. According to the European Commission's (EC) Joint Research Centre report [2]; nanomedicines are defined as particles with 1–1000 nm size, and biochips with nanoscale components and polymer therapeutics are also classified as nanomedicine. The significant development of nanomedicine has been achieved within the last couple of decades, though the first results were identified in the 1970s [3]. The early pioneer studies have shown that liposomal formulations of existing drugs have the potential of increasing therapeutic efficiency and reducing toxicity in non-clinical models, though these effects could not be well-demonstrated in patients [4]. Therefore, along with the difficulties in the commercial-scale production of nanoparticles, only limited investment and research have been made in nanoparticles. In 2003, the ESF conducted a study involving more than 100 experts which examined the status of nanotechnology in medical applications, and published a policy summarising their recommendations for further improvement of the field in 2005 [5]. In addition; other reports [6–8] and establishments including European Technology Platform Nanomedicine, European Foundation for Clinical Nanomedicine, Nanomedicine Development Centres of National Institutes of Health (NIH) were developed for further improvement of the field.

All noncommunicable diseases (NCD), also known as chronic diseases, such as cancer, diabetes, and cardiovascular diseases cause damage of cells, tissues, organs, and ultimately related systems in the body. NCDs represent 63 % of annual deaths [9], constituting a large proportion of death causes in the world. Collins, director of NIH, states that among approximately 4,000 diseases whose biological mechanisms have been elucidated, only 250 of them have available treatments currently [10]. Unfortunately, most of the available medical approaches cannot be classified as “satisfactory treatment”, enabling only symptomatic recovery of the patients, rather than being fully curative [11]. In addition to the evident need for new effective treatment strategies, early diagnosis that can increase survival rate is yet another important issue for patients who have diseases like cancer, which needs to be further improved. Similarly, in many neurodegenerative diseases such as multiple sclerosis (MS), patients can manifest clinical symptoms after the initiation of biological disease activity [12]. Therefore, it is very important to identify individuals at risk for these diseases, in which treatment at earlier stages usually defines milder disease course. Nanomedicine is a promising avenue for at least partly filling these huge gaps in diagnostic and therapeutic approaches in the field of medicine. To date, many successful nanomedicines have been developed using multidisciplinary approaches, and further improvements are being pursued. In this chapter, we review the current status of pharmaceutical and diagnostic nanomedicines, as well as the ethical issues and regulations regarding the field of nanomedicine.

23.2 Nanopharmaceuticals

Application of nanopharmaceuticals aims to specifically direct diseased tissues or cells, thereby preventing side effects caused by damages to healthy tissue in conventional treatments, and improving the therapy efficacy. These nanoparticles used as drug carriers have many advantages, especially over the standard cancer therapies. Conventional chemotherapies cause death of healthy cells along with cancer cells, leading to a load of serious adverse side effects. By delivering nanoparticles directly to the tumour tissue, prevention of the drugs interacting with healthy tissue before they reach the target tissue are achieved. Moreover, efficacy of the treatments is enhanced, since drug absorption proportion solely into cancer cells is increased.

There have been studies on three generation vectorisation strategies for nanomedicines, which are nanospheres and nanocapsules, hydrophilic polymer like polyethylene glycol (PEG)-coated nanoparticles, and vectors constituting of biodegradable cores and polymer envelopes with membrane recognition ligands. Using the vectorisation systems, many nanopharmaceutical products have already been marketed (Table 23.1). Nanoparticulate drug delivery has been achieved by a number of different approaches, of which passive targeting relies on “enhanced permeability and retention (EPR)” principle [13]. In EPR effect, particles smaller than 100–200 nm are taken up to tumour cells more easily due to their leaky blood vessels with greater porosity compared to the normal vasculature.

Table 23.1 Examples of nanomedicines

Product name	Indication	Status	Nanotechnology	Company
Doxil	Cancer (various)	Approved	PEGylated liposomal doxorubicin	Johnson and Johnson/Schering-Plough
Myocet	Breast cancer	Approved	Liposomal doxorubicin	Cephalon
Pegasys	Hepatitis C	Approved	PEGylated interferon alpha 2a	Roche
Abraxane	Breast cancer	Approved	Albumin-paclitaxel nanoparticles	Abraxis (Celgene)
Depodur	Pain relief	Approved	Liposomal morphine	EKR Therapeutics Inc
Venofer	Anemia	Approved	Iron oxide, sucrose	Fresenius
Rapamune	Immunosuppressive	Approved	NanoCrystal Elan	Pfizer/Wyeth
ThermoDox	Hepatocellular carcinoma	Phase 3 trial	Temperature sensitive liposomal doxorubicine+RFA	Celsion
CPX-351	Acute myeloid leukemia	Phase 3 trial	Liposomal cytarabine and daunorubicine combination	Celator Pharmaceuticals
BIND-014	Prostate cancer, lung cancer	Phase 2 trial	Targeted nanoparticle	BIND Biosciences
MM-398	Pancreatic cancer	Phase 3 trial	Liposome	Merrimack Pharmaceuticals
NKTR-105	Cancer (various)	Phase 1 trial	PEGylated docetaxel	Nektar

When nanoparticles are taken up to the tumour vessels, they have a long retention time due to poor lymphatic drainage in tumours. To increase the EPR effect, lipid composition of nanoparticles have been modified, allowing prolonged circulation in blood. Doxil (doxorubicin HCl liposome injection) for instance, developed to be used in cancer treatment, was designed based on this principle [14]. Both in xenograft models, and in the first clinical trial, enhanced accumulation of Doxil in tumour tissues was shown compared to the non-liposomal parent, free doxorubicine [15, 16].

In addition to the passive targeting, other promising approaches have been investigated mostly focusing on active targeting, in which nanoparticles with molecules that bind receptors specifically expressed on diseased cells are generated, allowing more accurate targeting. In this context, N-(2-Hydroxypropyl) methacrylamide (HPMA) copolymer-doxorubicine conjugates including galactosamine for targeting asialoglycoprotein receptor on hepatocytes and hepatoma was generated, revealing that after 24 h of dose application, 15–20 % liver targeting is achieved, although the higher level of accumulation was observed in normal liver (16.9 %) than in hepatoma (3.2 %) [17]. Still, the accumulation level in hepatoma was 12–50 times increased, compared to administration of free doxorubicine. Moreover, nanoshells that absorb light at different wavelengths generating heat

have been designed for an application of light with specific wavelength inducing hyperthermia in cancer cells once they are delivered the nanoshells by active targeting. An example for this system is ThermoDox, which is a radio frequency ablation (RFA)-sensitive liposomal formulation of doxorubicin [18]. When induced by RFA, the active drug is locally released into tumour cells, which is suggested to improve the outcome of hepatocellular carcinoma patients who receive RFA.

Currently, there are 93 clinical trials with the “nano” prefix on NIH (as of 16.06.2015), which passed the pre-clinical stages [19]. However, it is known that only about 5 % of drugs that enter clinical trials for all types of cancer therapies are later approved to be marketed [20]. This emphasises the importance of biomarkers for predicting parameters of developing nanomedicines, since the results of potential nanomedicines in vitro culture/in vivo animal model studies and human subjects can greatly vary. Potential nanomedicine biomarkers include;

- Markers showing healthy tissue toxicity
- Evidence of diseased tissue targeting
- Presence of receptors for active targeting
- Presence of activating enzymes and conditions
- Absence of resistance markers
- Early evidence of response

These potential nanomedicine-specific biomarkers mostly predicting targeting and toxicity should be identified and tested with respect to each product for enhanced efficacy and safety, which will improve patient selection for the therapies. However, studies on nanomedicine biomarkers are at early stages and required to be carefully developed with respect to each condition.

23.3 Diagnostic and Theranostic Nanomedicine

Use of nanoparticles in medical diagnosis have many advantages compared to the traditional methods. Such new generation in vitro and in vivo diagnostic devices are both smaller, faster, more specific and sensitive. In terms of in vitro diagnosis, these devices also require smaller samples from the individuals. An in vitro nano-sized diagnostic device, a biosensor, is composed of a biological material that can recognise the presence, activity, or level of a biological molecule, which is then converted to a quantifiable signal by a transducer. Biosensors can be used both in diagnosis of a disease and monitoring a therapy progress. In vivo imaging diagnostic devices comprise identifying tracers or contrast agents that mark the related sites in the body. Commonly used agents include fluorescent dyes, radionuclides, superparamagnetic iron oxides, and heavy elements. Targeted molecular imaging can provide detection of localisation of a tumour, as well as disease states, sites of inflammation, visualisation of vascular structures etc., allowing detection of diseases as early as possible. In vivo diagnostics also include implantable devices such as swallowable imaging pills, that enable monitoring of circulating molecules.

Although nano-harvesting of biomarkers have the potential to significantly improve early disease detection, there is a limitation that albumin and immunoglobulin constitute about 90 % of all molecules in blood, thereby preventing high signals from the trace substances. However, developing polymer-based nanoparticles seem to overcome this problem.

Nano-biopsy, in which a small amount of cells and molecules are obtained from the target area, is specifically advantageous in the detection of brain tumours, since its diagnosis is generally the most difficult one to make among all cancer types, and removing brain tissue by conventional biopsy is more dangerous due to the delicate nature of the brain. In a study of Stott et al. [21] for instance, a microchip named HB-Chip was developed to capture tumour cells from circulating blood (CTCs), which may later be used in the assessment of therapy effectiveness. The system is based on antibody-coated walls of the chip that specifically bind to cancer cells. The unique property of the chip is that it does not just detect cancer cells, but it also traps them, allowing further analyses on these cells, which may eliminate the need for conventional biopsy in certain cancer types in the future.

Combination of the diagnostic capability of a device with therapeutic property is referred by the term ‘nanotheranostics’. In the study of Muthu et al. [22] for example, MCF-7 breast cancer cell line was targeted by D-alpha tocopheryl polyethylene glycol 1000 succinate (TPGS)-coated liposomes with folic acid specific to folate receptor on the cells. The nanoparticle included docetaxel for therapy and quantum dots for imaging, leading to enhanced cellular uptake and toxicity for cancer cells [22]. Examples for nano-imaging products are shown in Table 23.2.

23.4 Ethics and Regulation

Nanotechnology in medical applications raises issues such as safety, sensitivity of genetic information and health care resources, as well as the interface between human and ‘machine’ like implants and chips. Application of nanomedicines to patients may

Table 23.2 Examples of nano-imaging agents

Product name	Indication	Status	Nanotechnology	Company
Endorem	Liver imaging	Approved	SPION dextran coating	AMAG Pharmaceuticals/Guerbet SA
Gastromark	Gastrointestinal imaging	Approved	SPION silicone coating	AMAG Pharmaceuticals
Lumirem	Gastrointestinal imaging	Approved	SPION silicone coating	AMAG Pharmaceuticals
Resovist	Small liver lesion imaging	Approved	SPION	Bayer Healthcare Pharmaceuticals
KER-0001	Angiogenesis imaging	Phase 1 trial	Targeted nanoparticle/fluorine-19	Kereos

cause serious adverse effects by rapid inflammatory-like hypersensitivity reactions or slower specific responses against the nanoparticles. Indeed, possible toxic effects of specific nanomedicines have been implicated in a number of studies [23, 24]. Also, long-term consequences of nanomedicines with no observed toxic effects in pre-clinical and clinical studies have not been well-known yet. Germline therapy with nanomedicines or possible effects of any medication with nanoparticles on germline cells raise yet another ethical issue, since they can cause serious damages to be later passed down to next generations. Regarding the safety issues, people who work in this field also stand at risk due to the possibility of being exposed to the nanoparticles by inhalation, swallowing, or interaction with skin [25, 26]. In addition, the clinical trial literature of first generation nanomedicines is rarely cited in new studies, even though there is a huge clinical data from these first generation products. Thus, conducting new clinical trials without considering earlier experiences is another ethical issue in clinical research of nanomedicines. Therefore, there should be a comprehensive ethical analyses regarding nanomedicines based on bioethic principles, in order to eliminate their possible adverse impacts on public health and balance the improvement of the nanomedicine field and restrictions for sustaining ethical, social, and legal boundaries.

Regulatory framework for current nanomedicines are based on benefit/risk assessments, even though there are no specific procedures. The first generation nanomedicines have been progressed from laboratory to clinical use by considering existing regulatory procedures. However, some criticisms concerning clinical safety of these products have emerged, suggesting that toxicological evaluation of the products should be done on a case by case basis. According to the US Food and Drug Administration (FDA) though, there is no need for a regulation specific to nanomedicines, since they have no significant differences from any other new technology. Still, there is an ongoing debate regarding nanomedicine-specific regulations, which will shape the procedures to be followed later on [27–29].

23.5 Conclusion

Considering there are many approved nanomedicines, and many others under clinical trials, a significant number of therapeutic and diagnostic nanoparticles will be further developed. Although different carrier designs and targeting strategies for nanomedicines with promising success have been developed, toxicity issues have been mostly disregarded. In case of overcoming the safety, as well as other ethical issues, the field of nanomedicine seems to make a significant contribution to healthcare in the future.

References

1. European Science Foundation's Forward Look Nanomedicine: An EMRC Consensus Opinion (2005). <http://www.esf.org>
2. V. Wagner, B. Hüsing, S. Gaisser, Nanomedicine: Drivers for development and possible impacts. European Commission's (EC) Joint Research Centre (JRC) Report No. 46744, 1–116 (2008)
3. G. Gregoriadis, B.E. Ryman, Liposomes as carriers of enzymes or drugs: a new approach to the treatment of storage diseases. *Biochem. J.* **124**, 58P (1971)
4. P.D. Brown, P.R. Patel, Nanomedicine: a pharma perspective. *Wiley Interdiscip Rev Nanomed Nanobiotechnol* **7**, 125–130 (2015)
5. The European Science Foundation, Nanomedicine: An ESF—European Medical Research Councils (EMRC) Forward Look report (2005). <http://www.esf.org>
6. Commission of the European Communities Communication, Towards a European Strategy for Nanotechnology. EU, DG Research, Brussels, (2004). www.cordis.lu/nanotechnology
7. National Institutes of Health, NIH roadmap: nanomedicine. NIH, USA, (2004). <http://nihroadmap.nih.gov>
8. National Institutes of Health—National Cancer Institute, Cancer nanotechnology plan: a strategic initiative to transform clinical oncology and basic research through the directed application of nanotechnology, NCI, NIH, USA, (2004). http://nano.cancer.gov/alliance_cancer_nanotechnology_plan.pdf
9. World Health Organization, 10 facts on noncommunicable diseases, (2013). <http://www.who.int>
10. F.S. Collins, We need better drugs—now. TEDMed Conference (Long Beach, CA, USA, 2012)
11. T. Bartfai, G.V. Lees, *The Future of Drug Discovery: Who Decides which Diseases to Treat?* (Academic Press, 2013)
12. D.T. Okuda, A. Siva, O. Kantarci, M. Inglese, I. Katz, M. Tutuncu, B.M. Keegan, S. Donlon, H. Hua le, A. Vidal-Jordana, X. Montalban, A. Rovira, M. Tintoré, M.P. Amato, B. Brochet, J. de Seze, D. Brassat, P. Vermersch, N. De Stefano, M.P. Sormani, D. Pelletier, C. Lebrun, Radiologically Isolated Syndrome Consortium (RISC), Club Francophone de la Sclérose en Plaques (CFSEP) Radiologically isolated syndrome: 5-year risk for an initial clinical event. *PLoS One* **9**, e90509 (2014)
13. Y. Matsumura, H. Maeda, A new concept for macromolecular therapies in cancer chemotherapy: mechanism of tumouritropic accumulation of proteins and the antitumour agent smancs. *Cancer Res.* **6**, 6387–6392 (1986)
14. Y. Barenholz, Doxil®—the first FDA-approved nano-drug: lessons learned. *J. Control Release.* **160**, 117–134 (2012)
15. J. Vaage, E. Barberá-Guillem, R. Abra, A. Huang, P. Working, Tissue distribution and therapeutic effect of intravenous free or encapsulated liposomal doxorubicin on human prostate carcinoma xenografts. *Cancer* **73**, 1478–1484 (1994)
16. A. Gabizon, R. Catane, B. Uziely, B. Kaufman, T. Safra, R. Cohen, F. Martin, A. Huang, Y. Barenholz, Prolonged circulation time and enhanced accumulation in malignant exudates of doxorubicin encapsulated in polyethylene-glycol coated liposomes. *Cancer Res.* **54**, 987–992 (1994)
17. L.W. Seymour, D.R. Ferry, D. Anderson, S. Hesselwood, P.J. Julyan, R. Poyner, J. Doran, A. M. Young, S. Burtles, D.J. Kerr, Cancer Research Campaign Phase I/II Clinical Trials committee, Hepatic drug targeting: phase I evaluation of polymer bound doxorubicin. *J. Clin. Oncol.* **20**, 1668–1676 (2002)
18. B. Wood, R.T.P. Poon, Z. Neeman, M. Eugeni, J. Locklin, S. Dromi, S. Kachala, R. Probhakar, W. Hahne, S.K. Libutti, Phase I study of thermally sensitive liposomes containing doxorubicin (ThermoDox [TD]) given during radiofrequency ablation (RFA) in patients with unresectable hepatic malignancies. *ASCO Gastrointestinal Cancers Symposium* (2007). abstr. 188
19. ClinicalTrials.gov, a service of the U.S National Institutes of Health. <http://clinicaltrials.gov>

20. D. Sarker, P. Workman (2007) Pharmacodynamic biomarkers for molecular cancer therapeutics. *Adv. Cancer Res.* **96**, 213–268
21. S.L. Stott, C.H. Hsu, D.I. Tsukrov, M. Yu, D.T. Miyamoto, B.A. Waltman, S.M. Rothenberg, A.M. Shah, M.E. Smas, G.K. Korir, F.P. Jr Floyd, A.J. Gilman, J.B. Lord, D. Winokur, S. Springer, D. Irimia, S. Nagrath, L.V. Sequist, R.J. Lee, K.J. Isselbacher, S. Maheswaran, D.A. Haber, M. Toner, Isolation of circulating tumor cells using a microvortex-generating herringbone-chip. *Proc. Natl. Acad. Sci.* **107**, 18392–18397 (2010)
22. M.S. Muthu, S.A. Kulkarni, A. Raju, S.S. Feng, Theranostic liposomes of TPGS coating for targeted co-delivery of docetaxel and quantum dots. *Biomaterials* **33**, 3494–3501 (2012)
23. G. Oberdörster, Pulmonary effects of inhaled ultrafine particles. *Int. Arch. Occup. Environ. Health* **74**, 1–8 (2001)
24. P. Jani, G.W. Halbert, J. Langridge, A.T. Florence, Nanoparticle uptake by the rat gastrointestinal mucosa: quantitation and particle size dependency. *J. Pharm. Pharmacol.* **42**, 821–826 (1990)
25. J. Lademann, H. Weigmann, C. Rickmeyer, H. Barthelmes, H. Schaefer, G. Mueller, W. Sterry, Penetration of titanium dioxide microparticles in a sunscreen formulation into the horny layer and the follicular orifice. *Skin Pharmacol. Appl. Skin Physiol.* **12**, 247–256 (1999)
26. G. Oberdörster, Z. Sharp, V. Atudorei, A. Elder, R. Gelein, A. Lunts, W. Kreyling, C. Cox, Extrapulmonary translocation of ultrafine carbon particles following whole-body inhalation exposure of rats. *J. Toxicol. Environ. Health A* **65**, 1531–1543 (2002)
27. R. Gaspar, R. Duncan, Polymeric carriers: preclinical safety and the regulatory implications for design and development of polymer therapeutics. *Adv. Drug Deliv. Rev.* **61**, 1220–1231 (2009)
28. M. Eaton, Nanomedicine: industry-wise research. *Nat. Mater.* **6**, 251–253 (2007)
29. R. Nijhara, K. Balakrishnan, Bringing nanomedicines to market: regulatory challenges, opportunities, and uncertainties. *Nanomed* **2**, 127–136 (2006)

Chapter 24

Microfluidics and Its Applications in Bionanotechnology

Z.P. Çakar and B. Sönmez

Abstract Several advantages of microfluidic systems such as rapid control, reduced size, low costs, large-scale integration and in vivo simulation of cellular microenvironments make them suitable for diverse bionanotechnological applications. This chapter reviews major current and potential microfluidic applications in bionanotechnology, including miniaturized devices related to common molecular biological techniques such as polymerase chain reaction (PCR), DNA microarray and electrophoresis; microfluidic bioreactors; and monitoring microbial behaviour by microfluidics. Additionally, present applications and future potential of microfluidics in microbial strain development and single cell analysis/characterization are discussed.

24.1 Introduction to Bionanotechnology and Microfluidics

Biological systems consist of complex molecular machineries, as shown by modern molecular biology research. Individual molecules are the functional constituents of these complex machineries. There are two major reasons that make living systems crucial for nanoscience and nanotechnology. One reason is the strong evidence provided by biological systems for the functionality of nanostructured systems. This

Z.P. Çakar (✉) · B. Sönmez

Department of Molecular Biology & Genetics, Faculty of Science and Letters,
Istanbul Technical University, 34469 Maslak, Istanbul, Turkey
e-mail: cakarp@itu.edu.tr

B. Sönmez

e-mail: sonmezbur@itu.edu.tr

Z.P. Çakar · B. Sönmez

Dr. Orhan Öcalgiray Molecular Biology, Biotechnology and Genetics Research Center
(ITU-MOBGAM), Istanbul Technical University, Maslak, Istanbul 34469, Turkey

is also known as biomimetics, a technology development approach inspired by biology. The second reason is that understanding life at nanoscale has a wide range of important applications in medicine, agriculture, engineering and sustainability of life. Using biomimetic approaches, self-assembly of biological molecules in vitro, namely outside a living organism, has been used to reconstitute biological function, which is the basis of bionanotechnology [1].

The invention of Polymerase Chain Reaction (PCR) method is a breakthrough in modern molecular biology, and also has significant impact on bionanotechnology. This method enables in vitro DNA replication. In this method, there is no need for the whole molecular machinery to denature DNA and to polymerize it, but finding a thermostable DNA polymerase, such as the one isolated from the thermophilic bacterium *Thermus aquaticus*, is necessary. At the RNA level, DNA microarrays constitute a powerful method by reading the RNA present in a biological sample and determining which genes are expressed at a particular time or under particular conditions. Thus, they provide valuable information about the state of a biological system [1].

The microelectronics industry initiated the progress in miniaturization, where fabrication sizes became comparable to those of biological molecules. Additionally, progress in chemistry allows the synthesis of an increasing number of constituents with a precision at atomic level. Thus, nanostructured materials can be fabricated and used for a wide range of nanotechnology applications, including diagnostics and detection [1].

Microfluidics can be defined as manipulation and precise control of fluids in channels that have cross-sectional dimensions in 10–100 μm range [2]. This central technology in many miniaturized systems has found a wide range of chemical, biological and medical applications. Early microfluidics systems were fabricated using microelectronics-derived technologies, such as photolithography and etching in silicon and glass. However, silicon is an expensive material and is not environmentally friendly. Additionally, it is an opaque material in the visible/UV range of spectrum which prevents its use for optical detection. Glass, on the other hand, is transparent, but it is amorphous, making vertical side walls difficult to etch. Most importantly, sealing of silicon and glass required clean room facilities, and high voltages or temperatures [2]. These difficulties were overcome by the use of new materials such as inexpensive polymers. In the late 90s, soft-lithography using polymer molding was introduced to fabricate cheap microfluidic devices. Particularly, soft-lithography of poly-di-methyl-siloxane (PDMS) became the basis of microfluidic systems for cell biology applications [3].

There are many advantages of microfluidics systems for biological applications: these systems are fast-responding microsystems, where changes in media, environmental changes and temperature control are fast. As the reduced size also reduces the effects of inertial forces compared to frictional ones, laminar flow is observed as the flow regime in microfluidic channels. Laminar flow allows static and dynamic gradient formation at subcellular resolution. Other practical

advantages of microfluidic systems are low reagent consumption, simultaneous and independent manipulation of large number of cells, automatic generation of various unique conditions, and large-scale integration. Additionally, the versatility of microfluidic devices allows simulation of *in vivo* cellular microenvironment, including vascularization and nutrient stress conditions [3].

The major present challenge of microfluidic systems is the relatively slow diffusive mixing resulting from the laminar flow regime, which limits applications that require a rapidly homogenizing flow. To overcome this limitation, the use of integrated mixers has been suggested [3, 4]. Other challenges of microfluidic systems include difficulties in adapting biological protocols to fit experiments in microsystems, such as media and cell concentrations; the bias resulting from biomolecule absorption or adsorption from the medium due to the affinity of PDMS for small hydrophobic molecules; and osmolarity changes resulting from media drying because of the permeability of PDMS to water vapor [3].

Technological advances in bionanotechnology include a variety of microfluidic applications, also for systems biomedicine [5]. As systems biology approach views and models a biological system as a whole, it also requires large amounts of data and information for analysis. Microfluidic system examples that help fulfill this requirement are an integrated microfluidic device for high-throughput measurement of molecular affinities between DNA and transcription factors to build inter-relation networks [6]; and integrated use of digital PCR with microfluidic manipulation and microarrays to obtain data from a single bacterium [7]. The latter is crucial for single cell analysis of heterogeneous microbial populations. Technological advances in bionanotechnology that involve microfluidics systems include nano- and microfluidic systems for handling fluids in high-throughput biological experiments with low sample volume [8]; real time quantitative reverse transcription polymerase chain reaction (qRT-PCR) applications with integrated microfluidic chips for single cell quantification of gene expression levels [9]; microfluidic devices with extracellular matrix-mimicking hydrogel applications for *in vitro* high-throughput drug screening, such as organ-on-a-chip platforms for systematic investigations of pharmacokinetics [10]; high-throughput detection of target molecules by combining conventional label-free detection methods with micro- or nanofluidic platforms [11]; and next generation DNA sequencing which is a breakthrough in genome and transcriptome analysis [5].

The present chapter reviews a wide range of current and potential microfluidics applications in bionanotechnology (Fig. 24.1). Firstly, it focuses on microfluidics applications in commonly used biological techniques such as PCR, DNA microarray systems, and electrophoresis. It then discusses the development of microfluidic bioreactors and their applications. Application of microfluidics technologies for monitoring microbial behavior and for microbial strain development is also reviewed, with an emphasis on evolutionary engineering approach to select for improved microbial strains. Finally, microfluidics applications in single cell studies are discussed.

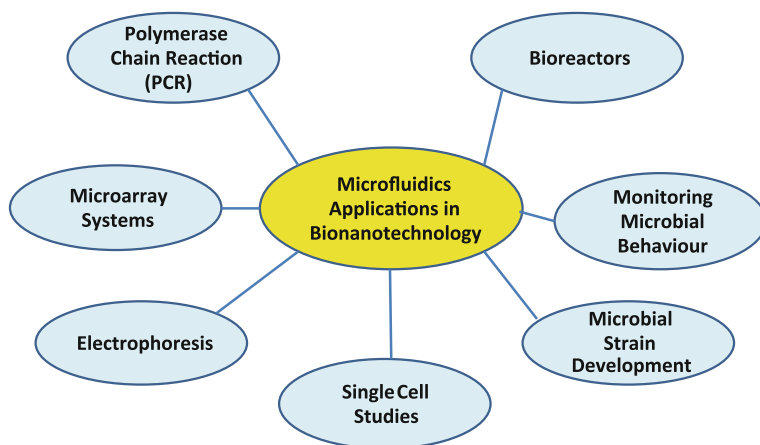


Fig. 24.1 Bionanotechnological microfluidics application areas reviewed in the present chapter

24.2 Microfluidic PCR Applications

Micro-electro-mechanical-systems (MEMS) technology enabled miniaturization of analytical devices used in biology and chemistry which had a strong influence on a variety of fields such as medical diagnostics, microbial detection, etc. Among miniaturized analytical devices, those related to PCR are studied extensively [12]. A review article published in 2005 summarizes properties of an automated chip-based PCR device for simple and rapid amplification of nucleic acids. The rapid temperature ramping ability and microliter-range sample volumes are the advantages of the described system [13]. Developments in microfluidics technology also include real-time PCR applications: in 2006, it was reported that a real-time PCR analyser was developed that is able to analyse 2 μ l reactions in glass capillaries [14].

Application of microfluidic technologies to PCR has also been useful for the detection of low-abundant DNA point mutations. For this purpose, flow-through biochips have been microfabricated using two different materials: a polycarbonate chip for performing the PCR reaction and the allele-specific ligation detection reaction in a continuous-flow format, and a poly(methyl methacrylate) (PMMA) chip for the detection of the allele-specific ligation detection reaction products [15].

Forensic science is a field that makes routine use of PCR-based analytical techniques. Thus, development of microfluidic PCR devices also influenced forensic medicine. The use of microfluidic devices for forensic DNA analysis has been reviewed in 2007, where advances in microfluidic separation of short-tandem-repeat (STR) fragments, and the analytical processes related to sample preparation have been discussed, along with the general advantages and challenges of microfluidic techniques [16]. In 2010, the development of an

Table 24.1 Selected reports on improvement of miniaturized PCR devices

Improvement area	References
Droplet technologies	[19–22]
Transportation mechanism	[23]
Extrusion-driving method based on microfabricated PDMS pneumatic valves	[24]
Vapor-diffusion barrier for pressure reduction	[25]
Integrated design and optimization of a cyberphysical digital-microfluidic biochip	[26]

automated instrument for human STR identification has been reported: the instrument involved a lab-on-a-chip automation system to enable DNA fingerprinting that is compatible with the CODIS standards used by forensic scientists. The miniaturized instrument aimed to improve costs, duration and simplicity of the analysis and to exclude human intervention [17]. A very recent forensic science article reported an improvement of the miniaturized PCR-STR typing method by using digital agarose droplet microfluidics: the improvement is based on minimizing inhibition by PCR inhibitors. The inhibitors were efficiently removed from the porous agarose microgel droplets that carry the DNA template through washing and significant dilution. Thus, the improved method could allow analysis of low-abundance DNA targets in the presence of inhibitors [18].

Regarding improvement of miniaturized PCR devices, there are various reports in the literature (Table 24.1). Some studies focused on droplets: the first lab-on-a-chip system for picoliter droplet generation and PCR amplification including real-time fluorescence detection enabled PCR in isolated droplets of picoliter range [19]. A droplet microfluidic chip allowing amplification-free detection of single pathogenic cells was also developed: in that chip, a peptide nucleic acid fluorescence resonance energy transfer probe was used to detect 16S rRNA in pathogenic cells [20]. The advantages of droplet microfluidics in the high-throughput analysis of single cells and molecules in cancer research, diagnostics and therapy with exceptionally high sensitivity have also been reviewed recently [21]. Similarly, novel droplet platforms for the detection of disease biomarkers have also been reviewed recently, by emphasizing the importance of microfluidic droplet technologies for personalized medicine applications [22].

Other studies on the improvement of miniaturized PCR systems include improvement of the transport of DNA samples and PCR reagents by a novel liquid transportation mechanism based on a suction-type membrane and three microvalves [23], introduction of a novel extrusion driving method based on microfabricated PDMS pneumatic valves [24], reduction of the pressure increase in fully closed microfluidic PCR systems due to high operating temperatures [25], and an integrated design and optimization of a cyberphysical digital-microfluidic biochip for the entire miniaturized PCR procedure [26].

24.3 Microfluidic DNA Microarray Systems

DNA microarray technology is a powerful tool for medicine, biomedical and basic research, as DNA microarrays can be used to search for nucleic acids and molecular signatures related to a disease or to monitor a physiological/environmental condition. Microfluidics has many advantages such as small sample volumes, minimized reagent consumption and costs, automated sample preparation and reduced sample processing time. Thus, the use of microfluidics in DNA microarray technology also reduces reagent costs and hybridization assay times, provides high-throughput sample processing, and integration and automation abilities of the front-end sample processing steps. Challenges of microfluidic DNA microarray technology include the need for low-cost manufacturing methods for the fluidic chips, providing good interfaces to the macro-world, minimizing high surface-to-volume-ratio-related problems such as non-specific analyte-wall interactions, the need for materials for the optical readout phases, and complete integration of the optical and electrical components to the microfluidics [27]. As flow is challenging in microfluidic systems, various methods developed to generate nanoliter flows in the microchannels of the microfluidic DNA microarray systems have been discussed in a more recent review. These methods include electrokinetic control, vacuum suction and syringe pumping. Of these, vacuum suction and syringe pumping are pressure-driven, and do not require consideration of the physicochemical properties of solutions, which allows flexibility. Centrifugal force has also been employed to drive liquid movement in microchannels, by using the body force from the liquid. Besides, no additional solution interface contacts from electrodes, syringes and tubing exist. When flow is provided by centrifugal force, parallel hybridizations are also made easily [28].

DNA microarrays are usually processed under one condition, which requires that all probes of the array have to function optimally under that particular assay condition. However, cross-hybridization is a common microarray problem, which results from a poor combination of the probes chosen and the assay conditions. Microfluidics-based technologies provide new array processors which are powerful tools for constructing highly specific DNA microarrays according to the needs of samples, including special samples that have DNA regions with highly varying G +C content. Thus, high sample throughput can be combined with optimal assay conditions [29].

Semiconductor nanoparticles known as quantum dots have unique optical features such as high brightness, narrow emission band, etc. which make them advantageous over traditional organic fluorophores. Thus, they have been replacing traditional organic fluorophores to be used as simple fluorescent reporters in a variety of biological techniques such as immunoassays, microarrays, fluorescent imaging, etc. Combination of advanced quantum dot probes with microfluidics and confocal fluorescence spectroscopy further enhances the detection limit to single molecule level [30].

There are various reports in the literature regarding present and future medical applications of microfluidic DNA microarray systems. An important example is the use of a microfabricated DNA microarray system for the pharmacogenomic profiling of cytochrome P450-mediated drug metabolism which can also be viewed as a work related to personalized medicine [31]. A more recent review discussed developments in enzymatically assisted detection of nucleic acids on microarrays and the potential integration of these new techniques into personalized medicine [32].

Applications of microfluidic DNA microarray systems are not limited to medicine. Basic and environmental microbiology research also makes use of microfluidic DNA microarray systems. Investigation of microbial diversity and detection of specific microorganisms (e.g. pathogens) in environmental samples such as wastewater are allowed without cultivation, by advances in molecular biology methods including microarray assays [33]. The use of chip-based microfluidic systems such as microarrays and microfluidic biochips for rapid pathogen sensing has also been reviewed previously [34].

24.4 Microfluidic Applications in Electrophoresis

Analysis of biological systems or microbial populations at the single cell level requires techniques that can rapidly detect and quantify a variety of analytes from extremely small sample volumes with very high sensitivity and selectivity. For this purpose, microelectrophoretic separations which use both traditional capillary electrophoresis and modern microfluidic methods are employed. Microelectrophoretic separation techniques allow quantification of diverse analytes in cells, such as amino acids, neurotransmitters, large proteins and subcellular organelles [35]. In a recent review on microchip electrophoresis, it has been reported that electrophoresis has been the most common separation method used in microfluidic applications for more than two decades [36].

Capillary electrophoresis is a widely used technique for water quality analysis, drug development and quality control, proteomics and DNA analysis, counter-terrorism and monitoring of corrosion. For common use of this technique in the field, portable and in situ capillary electrophoresis systems are necessary [37]. Thus, microchip capillary electrophoresis technology has been developed.

Similar to other microfluidic systems, issues related to flow have also been studied in microchip capillary electrophoresis systems. Electroosmotic flow is an important property of fluids in microchannels, and methods to measure electroosmotic flow have been reviewed for both capillary electrophoresis and microchip capillary electrophoresis [38]. Other methods discussed in the literature to improve microchip electrophoresis systems include methods to construct contactless conductivity detectors [39], pressure-driven and electrokinetic methods for reliable and

reproducible injection of representative samples in microchip electrophoresis systems [40], and methods for online sample preconcentration of all analytes in the sample reservoir before they reach the separation channel of the microfluidic capillary electrophoresis system [41].

Today, microchip capillary electrophoresis systems have several applications. The importance and clinical application potential of microchip capillary electrophoresis systems has already been foreseen and reviewed about two decades ago [42]. Rapid, sensitive and reliable detection and analysis of nerve agents by capillary electrophoresis and lab-on-a-chip technology has also been discussed [43]. As nerve agents are generally used in local conflicts or by local terroristic groups, it is important to develop such rapid and easy detection techniques as counter-terrorism applications. Considering the innovative microfluidic technologies that are applied to capillary electrophoresis, capillary electrophoresis is becoming increasingly important for green analytical separation and Green Chemistry applications including food analysis [44], where the use of solvents and reagents can be reduced, energy consumption is minimized by rapid analysis, and providing miniaturized, portable equipment [45]. The field of applied microbiology also benefits from microfluidic capillary electrophoresis applications: adaptation of capillary electrophoresis to microchip format and coupling it with mass spectrometry has been reported as a powerful approach for rapid and reliable analysis of bacterial outer membrane proteins. As these proteins are important for bacterial antibiotic sensitivity and pathogenicity, their detection is critical [46].

Other microfluidic developments in the field of electrophoresis are summarized in Table 24.2. These include two-dimensional (2-D) separation of proteins in microfluidic devices by a miniaturized 2-D gel electrophoresis system [47]; implementation of isoelectric focusing in microfluidic chips [48]; miniaturization of free-flow electrophoresis which is a method performing electrophoretic separation on a continuous analyte stream, as it is flowing through a planar flow channel where the electric field is applied perpendicularly to the flow [49]; microfluidic electrophoretic mobility shift assays [50]; and microdialysis coupled to microchip electrophoresis for in vivo and in vitro monitoring of biological events in a 'separation-based sensor' mode [51].

Table 24.2 Some examples of microfluidic developments in the field of electrophoresis

Name of the developed system	References
Miniaturized 2-D gel electrophoresis system	[47]
Isoelectric focusing system in microfluidic chips	[48]
Miniaturized free-flow electrophoresis system	[49]
Microfluidic electrophoretic mobility shift assay (EMSA)	[50]
Microdialysis coupled to microchip electrophoresis	[51]

24.5 Microfluidic Bioreactors

As a result of the miniaturization of biomedical and biochemical devices for micro-electromechanical (MEMS) systems, a variety of products have been developed, including biochips/biosensors, drug delivery systems, tissue scaffolds and bioreactors. For biochemical and biomedical applications, polymeric materials have been developed to replace the use of silicon, glass or quartz in MEMS devices, as these polymers are cheaper, easy to process and biocompatible [52]. A typical example is a microfluidic polydimethylsiloxane (PDMS) bioreactor that was used for large-scale cultivation of hepatocytes, where ten PDMS layers were stacked together as four single culture chambers, with high gas permeability of PDMS allowing sufficient oxygen supply to the microfluidic channels. With these modifications, a culture with higher cell density and sufficient oxygen and nutrient supply could be obtained, which could make the microfluidic bioreactor suitable for large-scale cultures [53]. Another example of microfluidic bioreactor systems is related to regenerative medicine process development: a novel, autoclavable, microfabricated scale-down device has been developed and inactivated mouse embryonic fibroblasts along with human embryonic stem cell colonies have been seeded to grow human embryonic stem cells [54]. In a more recent study, adherent cell cultures were grown and characterized in an online and automated fashion by using a microfabricated bioreactor. The developed bioreactor system allowed the automation of flow control, environmental control, and provided time-course data of the cell cultures for continuous monitoring of their characteristics [55]. Other microfluidic bioreactor applications include optimization of production and evaluation of retroviruses and other viral vectors for potential gene therapy applications [56], and high-throughput cytotoxicity assays allowed by long-term cultivation in microfluidic bioreactors with continuous perfusion [57].

Long-term cultivation in microfluidic bioreactors is particularly important for drug discovery and development to test the effects of drugs or toxin exposure. An early report is on the development of a 'NanoLiterBioReactor' (NBR) for long-term mammalian cell cultures. NBR prototypes have been fabricated with PDMS-based microfluidic components, and three different cell lines including fibroblasts, Chinese Hamster Ovary cells and hepatocytes were tested for cultivation, and high viabilities were achieved [58]. Periodic 'flow-stop' perfusion microchannel bioreactors have also been developed for long-term cultures of mammalian and human embryonic stem cells. These bioreactors enabled co-culturing of undifferentiated colonies of human embryonic stem cells on human foreskin fibroblast feeder cells in microchannels. It was suggested that this has been possible due to two properties of the novel method: short temporal exposure to flow followed by long static incubation [59]. Another study is on the development of a microfluidic traps system that allows prolonged culture of human embryonic stem cell aggregates. The formation of these aggregates is necessary to induce embryonic stem cell differentiation. Thus, the designed traps system facilitates cell capture and aggregation and

allows efficient gas/nutrient exchange, suggesting its future use in embryogenesis research [60].

The use of bioreactors to design three dimensional (3D) artificial cellular environments is another major field of research and development. Designing 3D biological fluidic environments for advanced cell cultures requires the use of microfluidic devices to tightly regulate fluid direction and behaviour in a geometrically limited environment. These requirements are met by the modern cultivation devices like bioreactors and multicompartiment systems [61]. The major advantage of 3D culture systems is the fact that they reveal the *in vivo* conditions much better than the 2D cultivation systems. Thus, technical systems are used to generate 3D, organ-like structures [62]. The synthesis of monodisperse water-in-oil microemulsions in microfluidic devices is important for precise bead control and high surface area, and these emulsions are used for the synthesis of highly porous polymer beads which have a high potential to be used in 3D cell culture applications [63].

A variety of studies and reviews exist in the literature on the improvement of microbioreactors, their monitoring, and control. An expert panel review article on monitoring and control of microbioreactors was published by the Working Group of the European Section of Biochemical Engineering Science (ESBES) in the European Federation of Biotechnology (EFB), by emphasizing the importance of developments in microbioreactor monitoring and control on future bioprocess development and *in vitro* toxicity testing [64]. Other studies on improvement of microbioreactors are listed in Table 24.3. These include development of a fluorescence lifetime-based optical imaging technique in microfluidic bioreactors that allows oxygen monitoring in continuous cell culture systems [65], integration of analyte-specific holographic sensors into microfluidic devices for the real-time pH monitoring in microbioreactors [66], development of a microfluidic bioreactor for continuous amperometric measurement of glucose [67], development of energy-efficient nanobubbles flexible in injection and production [68], nanoporous membrane-sealed microfluidic devices to improve cell viability and avoid perfusion requirement and related complications [69], and *in vitro* biosynthesis of metal nanoparticles in microdroplets [70].

Table 24.3 Selected reports on the improvement of microbioreactors

Improvement area	References
Oxygen monitoring	[65]
Real-time pH-monitoring	[66]
Continuous amperometric glucose measurement	[67]
Energy-efficient nanobubble development (flexibility in injection and production)	[68]
Nanoporous membrane-sealing of microbioreactors to improve cell viability and avoid perfusion and its complications	[69]
<i>In vitro</i> biosynthesis of metal nanoparticles in microdroplets	[70]

It is important to note that, apart from medicine, basic and applied microbiological research also benefits significantly from the developments in microfluidic bioreactor technology. An early example is a report on the development of a portable aerobic microbioreactor which allowed direct microscopic monitoring of the time-integrated growth response of a methanogenic bacterium, *Methanosaeta concilii*, and determination of the optimum growth conditions for the tested microorganism [71]. An integrated microbioreactor with the ability to cultivate microorganisms in distinct droplets without the need of valves, mixers or pumps has also been developed using digital microfluidics, for automated cultivation and analysis of a variety of microorganisms including bacteria, algae and yeast [72].

24.6 Monitoring Microbial Behaviour by Microfluidics

An important contribution of microfluidics to the fields of microbiology and microbial ecology is the acquired ability to observe microbial behaviour in highly controlled microenvironments, ranging from single cells to mixed microbial communities. Thus, natural microbial habitats could be mimicked at the micrometer scale. It has been suggested that this enabling technology could be exploited to control important microbial processes including antibiotic resistance and biofilm formation [73]. Some microfluidics application examples to monitor microbial behaviour are summarized in Table 24.4. Design of a microfluidic biochip by combining contactless dielectric microsensors with microfluidics for online monitoring of fungal biofilm dynamics has already been reported [74]. Additionally, secondary flow structures resulting due to semi-confined features in a microfluidic device have also been studied to investigate their effects on bacterial biofilm formation and dynamics [75].

Another important microbial behaviour is ‘chemotaxis’, the ability of some microorganisms to move towards an attractant and to move away from a repellent compound. Microfluidic technologies have also been used to study chemotaxis, as they allow observations at high spatial and temporal resolution in controlled microenvironments. Chemotaxis of free-swimming bacteria is challenging to observe by microfluidics. However, recent development of advanced microdevices generated flow-free and steady gradients. Particularly, the microfluidic gradient

Table 24.4 Microfluidics application examples to monitor microbial behaviour

Type of microbial behaviour monitored	Reference number (s)
Biofilm formation and dynamics	[74, 75]
Chemotaxis	[76, 77, 79]
Response to drugs/antibiotics	[78–80]
Quorum sensing	[81]

generators are promising to investigate the basic mechanisms of bacterial chemotaxis, and the dynamics of chemotaxis in realistic environments [76]. In a previous study, microfluidics and image analysis of swimming behaviour were combined to study chemoattraction to dimethylsulfoniopropionate, a compound produced by phytoplanktons, and the influences on marine planktonic food webs and global climate change were discussed [77].

Microfluidic technology has also been used to study microbial behaviour under the influence of drugs. The use of a microchip-based system to study the effects of the antibiotic, ampicillin, on the bacterium *Escherichia coli* has been reported [78]. Similarly, a microfluidic device was used to expose the pathogenic bacterium *Vibrio cholerae* to an antibiotic and to investigate its behavioural response [79]. In a more recent study, a 3D microfluidic culture device was used to study pharmacodynamic effects of antibiotics on bacterial cells. For this purpose, bacterial cells were encapsulated in an agarose-based gel in a commercially available microfluidics chamber, and time-lapse photography was applied to observe bacterial responses to a linear concentration gradient of antibiotics [80].

The ability of bacteria to sense other bacteria or ‘the crowd’ in their surroundings is an important phenomenon called ‘quorum sensing’ which influences many physiological characteristics such as bacterial growth rate and metabolism, antibiotic resistance and biofilm formation. In a recent study; a lithographic technique called micro-3D printing was used to organize bacteria in complex aggregates, to study the influence of spatial and environmental parameters on social behaviours of bacteria, such as quorum sensing [81].

24.7 The Use and Potential of Microfluidics in Microbial Strain Development

The aim of process development in microbial fermentations is to obtain a high production yield. Development of microbial strains used in fermentations is a common strategy to obtain high yields and more efficient bioprocesses. Microbial strain development usually requires application of high-throughput screening and optimization procedures [82].

Directed evolution is a high-throughput strategy used for microbial strain development. It is an experimental system mimicking natural evolution in a test tube at the molecular level, in a directed and rapid fashion. Thus, it is also called ‘laboratory evolution’. Directed evolution has been viewed as an aspect of nanobiotechnology, which ‘has evolved from protein engineering, via chemical biology or combinatorial chemistry, to nanobiotechnology’ [83]. Although the term ‘directed evolution’ is more commonly used to describe methods that improve protein/enzyme properties, it also describes methods to improve properties of whole

microbial cells. Those methods are more commonly referred to as ‘evolutionary engineering’, as discussed previously [84]. The basis of directed evolution methods is random mutagenesis and high-throughput screening, either at whole cell or at protein level.

The main focus of our research group is to improve stress resistance of the industrially important yeast, *Saccharomyces cerevisiae*. Our ultimate aim is to obtain robust yeast strains for more efficient yeast bioprocesses, and to understand the complex molecular mechanisms of resistance to a variety of stress types by using genomic, transcriptomic and/or proteomic approaches. We have been using evolutionary engineering strategy to obtain stress-resistant *S. cerevisiae* strains, by applying random mutagenesis with the use of a mutagenic agent called ethyl methane sulfonate, followed by selection under stress conditions. The selection procedure is applied either as successive batch cultivations in culture tubes or shake flasks, or as prolonged chemostat cultivations in a lab-scale bioreactor [85–88]. Using evolutionary engineering, we have successfully obtained *S. cerevisiae* strains with multistress resistance [86], cobalt resistance [87, 89], and nickel resistance [88]. Thus, evolutionary engineering is a powerful strategy to obtain yeast strains with desired properties, as reviewed previously [90]. However, prolonged batch cultivations in culture tubes or shake flasks cannot provide a good control over bioprocess conditions and the data are limited to end-point measurements. Alternatively, the use of long-term chemostat cultivation for selection allows better bioprocess control, but is more expensive and more labor intensive regarding sterilization, assembly, sensor calibration, operation, and cleaning of the bioreactor. Thus, the use of microfluidic bioreactors with online control over cultivation data and automation, and reduced cost and time would be beneficial for microbial bioprocess development applications, including strain development by evolutionary engineering [82, 90].

During our evolutionary engineering studies, we discovered that the final stress-resistant populations obtained from selection experiments were highly heterogeneous regarding their resistance levels against the particular stress type applied during selection. This phenomenon of ‘population heterogeneity’ is important for evolutionary engineering studies, indicating the multigenic and complex nature of stress resistance, as discussed previously [90], and it emphasizes the importance of carefully analysing many individual resistant mutants, to find those with the highest resistance levels. In fact, cell-to-cell heterogeneity in microbial bioprocesses can result from a variety of biological and environmental factors, and is known to have a severe impact on productivity. Thus, the design of innovative microfluidic systems that can analyse single cells in a microbial culture also allows detailed analysis of population heterogeneity and bioreactor inhomogeneity. Fluorescence-based imaging technologies integrated in microfluidic systems are generally used to analyse population heterogeneity. It is expected that microfluidics will become a major single-cell analytical technique in biotechnological processes and strain characterization [91].

24.8 Microfluidic Applications in Single Cell Studies

Microfluidics is suitable for single-cell characterization and analysis, because it allows precise microenvironment control, high-throughput manipulation, and multi-parameter label-free detection of individual cells [92]. Developments in microfluidic technology for chemical-content analysis of single cells have been reviewed previously, by also discussing applications at DNA, RNA and protein levels [93].

Microfluidics applications for single cell studies are found in diverse research fields. One of them is the field of drug discovery, development and delivery [94, 95]. Microfluidic drug delivery devices that make use of a concentration gradient generator for drug delivery at individual cell level are of particular interest [96]. Stem cell research also benefits from microfluidic single-cell applications: human stem cell differentiation can be regulated by using microfluidic single-cell technologies [97]. In the field of applied and environmental microbiology, single cell genomics of bacteria and archaea has been enabled by microfluidic devices [98]. In a recent review, the use of nanofabrication and microfluidic techniques to study individual bacteria and multispecies communities was discussed. It was suggested that the use of these techniques can help gain insight into complex bacterial processes like aging, electron transport, quorum sensing, and investigate microbial communities by single cell genomics [99].

24.9 Conclusions

Microfluidic systems have several advantages that make them suitable for bionanotechnological applications: rapid control, reduced size, low reagent consumption, low costs, simultaneous and independent manipulation of a large number of cells, large-scale integration and *in vivo* simulation of cellular microenvironments, etc. The present challenges of microfluidic systems such as slow diffusive mixing due to laminar flow, difficulties in adapting biological protocols to fit microfluidic experiments and other issues are being studied to improve microfluidic systems.

As presented in this chapter, there are diverse microfluidics applications in bionanotechnology, and microfluidics seems to have a great potential particularly for future applications related to single cell analysis and characterization. Today, DNA sequencing is more common and feasible, compared to sequencing of other molecules, because of the powerful PCR method for easy amplification of the DNA sample to be sequenced, and the existing high-throughput next generation DNA sequencing technologies. Development of new and miniaturized technologies for sequencing and analyzing proteins, carbohydrates and other biologically important polymers is an important challenge for microfluidics research related to bionanotechnology.

Acknowledgments Our research presented in this chapter was supported by the COST Action CM0902, Turkish Scientific and Technological Research Council (TÜBİTAK) (project no: 105T314, 107T284, 109T638, PI: ZPÇ), Istanbul Technical University (ITU) Research Funds (project no: 30108, 34200, PI: ZPÇ). BS is financially supported by the Faculty Member Training Programme (ÖYP) of the Council of Higher Education (YÖK) in Turkey.

References

1. NanoInk Inc. (2010) Introduction to nanoscale science and technology, Version 1, Chapter 5, *Introduction to Nanobiology* (Illinois, USA), pp. 162–209
2. J. Cooper McDonald, D.C. Duffy, J.R. Anderson, D.T. Chiu, H. Wu, O.J.A. Schueller, G.M. Whitesides, Fabrication of microfluidic systems in poly(dimethylsiloxane). *Electrophoresis* **21**, 27–40 (2000)
3. G. Velve-Casquillas, M. Le Berre, M. Piel, P.T. Tran, Microfluidic tools for cell biological research. *Nano Today* **5**, 28–47 (2010)
4. C.J. Campbell, B.A. Grzybowski, Microfluidic mixers: from microfabricated to self-assembling devices. *Philos. Trans. A Math. Phys. Eng. Sci.* **362**, 1069–1086 (2004)
5. E.-S. Kim, E.H. Ahn, E. Chung, D.-H. Kim, Recent advances in nanobiotechnology and high-throughput molecular techniques for systems biomedicine. *Mol. Cells* **36**, 477–484 (2013)
6. S.J. Maerkl, S.R. Quake, A systems approach to measuring the binding energy landscapes of transcription factors. *Science* **315**, 233–237 (2007)
7. E.A. Ottesen, J.W. Hong, S.R. Quake, J.R. Leadbetter, Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. *Science* **314**, 1464–1467 (2006)
8. J. Kim, M. Junkin, D.H. Kim, S. Kwon, Y.S. Shin, P.K. Wong, B.K. Gale, Applications, techniques, and microfluidic interfacing for nanoscale biosensing. *Microfluid. Nanofluid.* **7**, 149–167 (2009)
9. V. Sanchez-Freire, A.D. Ebert, T. Kalisky, S.R. Quake, J.C. Wu, Microfluidic single-cell real-time PCR for comparative analysis of gene expression patterns. *Nat. Protoc.* **7**, 829–838 (2012)
10. J.H. Tsui, W. Lee, S.H. Pun, J. Kim, D.H. Kim, Microfluidics-assisted in vitro drug screening and carrier production. *Adv. Drug Deliv. Rev.* **65**, 1575–1588 (2013)
11. Y. Zhang, Y. Tang, Y.H. Hsieh, C.Y. Hsu, J. Xi, K.J. Lin, X. Jiang, Towards a high-throughput label-free detection system combining localized-surface plasmon resonance and microfluidics. *Lab Chip* **12**, 3012–3015 (2012)
12. C.S. Zhang, J.L. Xu, W.L. Ma, W.L. Zheng, PCR microfluidic devices for DNA amplification. *Biotechnol. Adv.* **24**, 243–284 (2006)
13. G. Munchow, D. Dadic, F. Daffing, S. Hardt, K.S. Drese, Automated chip-based device for simple and fast nucleic acid amplification. *Expert Rev. Mol. Diagn.* **5**, 613–620 (2005)
14. P.N. Ngatchou, M.R. Holl, C.H. Fisher, M.S. Saini, J.C. Dong, T.T.H. Ren, W.H. Pence, D.L. Cunningham, S.E. Moody, D.A. Donaldson, D.R. Meldrum, A real-time PCR analyzer compatible with high-throughput automated processing of 2 μ L reactions in glass capillaries. *IEEE Trans. Autom. Sci. Eng.* **3**, 141–151 (2006)
15. M. Hashimoto, F. Barany, S.A. Soper, Polymerase chain reaction/ligase detection reaction/hybridization assays using flow-through microfluidic devices for the detection of low-abundant DNA point mutations. *Biosens. Bioelectron.* **21**, 1915–1923 (2006)
16. K.M. Horsman, J.M. Bienvenue, K.R. Blasier, J.P. Landers, Forensic DNA analysis on microfluidic devices: a review. *J. Forensic Sci.* **52**, 784–799 (2007)
17. C. Hurth, S.D. Smith, A.R. Nordquist, R. Lenigk, B. Duane, D. Nguyen, A. Surve, A. J. Hopwood, M.D. Estes, J.N. Yang, Z. Cai, X.J. Chen, J.G. Lee-Edghill, N. Moran, K. Elliott,

- G. Tully, F. Zenhausem, An automated instrument for human STR identification: design, characterization, and experimental validation. *Electrophoresis* **31**, 3510–3517 (2010)
18. T. Geng, R.A. Mathies, Minimizing inhibition of PCR-STR typing using digital agarose droplet microfluidics. *Forensic Sci. Int-Genetics* **14**, 203–209 (2015)
 19. N.R. Beer, B.Y. Hindson, E.K. Wheeler, S.B. Hall, K.A. Rose, I.M. Kennedy, B.W. Colston, On-chip, real-time, single-copy polymerase chain reaction in picoliter droplets. *Anal. Chem.* **79**, 8471–8475 (2007)
 20. T.D. Rane, H.C. Zec, C. Puleo, A.P. Lee, T.H. Wang, Droplet microfluidics for amplification-free genetic detection of single cells. *Lab Chip* **12**, 3341–3347 (2012)
 21. D.K. Kang, M.M. Ali, K.X. Zhang, E.J. Pone, W.A. Zhao, Droplet microfluidics for single-molecule and single-cell analysis in cancer research, diagnosis and therapy. *Trac-Trend Anal. Chem.* **58**, 145–153 (2014)
 22. H. Zec, D.J. Shin, T.H. Wang, Novel droplet platforms for the detection of disease biomarkers. *Expert Rev. Mol. Diagn.* **14**, 787–801 (2014)
 23. L.J. Chien, J.H. Wang, T.M. Hsieh, P.H. Chen, P.J. Chen, D.S. Lee, C.H. Luo, G.B. Lee, A micro circulating PCR chip using a suction-type membrane for fluidic transport. *Med. Microdevices* **11**, 359–367 (2009)
 24. Z.R. Xu, X. Wang, X.F. Fan, J.H. Wang, An extrusion fluidic driving method for continuous-flow polymerase chain reaction on a microfluidic chip. *Microchim. Acta* **168**, 71–78 (2010)
 25. G. Czilwik, I. Schwarz, M. Keller, S. Wadle, S. Zehnle, F. von Stetten, D. Mark, R. Zengerle, N. Paust, Microfluidic vapor-diffusion barrier for pressure reduction in fully closed PCR modules. *Lab Chip* **15**, 1084–1091 (2015)
 26. Y. Luo, B.B. Bhattacharya, T.Y. Ho, K. Chakrabarty, Design and optimization of a cyberphysical digital-microfluidic biochip for the polymerase chain reaction. *IEEE Trans. Comput. Aid. D* **34**, 29–42 (2015)
 27. C. Situma, M. Hashimoto, S.A. Soper, Merging microfluidics with microarray-based bioassays. *Biomol. Eng.* **23**, 213–231 (2006)
 28. L. Wang, P.C.H. Li, Microfluidic DNA microarray analysis: a review. *Anal. Chim. Acta* **687**, 12–27 (2011)
 29. M. Dufva, J. Petersen, L. Poulsen, Increasing the specificity and function of DNA microarrays by processing arrays at different stringencies. *Anal. Bioanal. Chem.* **395**, 669–677 (2009)
 30. Y. Zhang, T.H. Wang, Quantum dot enabled molecular sensing and diagnostics. *Theranostics* **2**, 631–654 (2012)
 31. L. Bissonnette, M.G. Bergeron, Next revolution in the molecular theranostics of infectious diseases: microfabricated systems for personalized medicine. *Expert Rev. Mol. Diagn.* **6**, 433–450 (2006)
 32. D.A. Khodakov, A.V. Ellis, Recent developments in nucleic acid identification using solid-phase enzymatic assays. *Microchim. Acta* **181**, 1633–1646 (2014)
 33. K.A. Gilbride, D.Y. Lee, L.A. Beaudette, Molecular techniques in wastewater: understanding microbial communities, detecting pathogens, and real-time process control. *J. Microbiol. Meth.* **66**, 1–20 (2006)
 34. J. Mairhofer, K. Roppert, P. Ertl, Microfluidic systems for pathogen sensing: a review. *Sensors* **9**, 4804–4823 (2009)
 35. L.M. Borland, S. Kottegoda, K.S. Phillips, N.L. Allbritton, Chemical analysis of single cells. *Annu. Rev. Anal. Chem.* **1**, 191–227 (2008)
 36. E.R. Castro, A. Manz, Present state of microchip electrophoresis: state of the art and routine applications. *J. Chromatogr. A* **1382**, 66–85 (2015)
 37. A.P. Lewis, A. Cranny, N.R. Harris, N.G. Green, J.A. Wharton, R.J.K. Wood, K.R. Stokes, Review on the development of truly portable and in-situ capillary electrophoresis systems. *Meas. Sci. Technol.* **24**, 042001 (2013)
 38. W. Wang, F. Zhou, L. Zhao, J.R. Zhang, J.J. Zhu, Measurement of electroosmotic flow in capillary and microchip electrophoresis. *J. Chromatogr. A* **1170**, 1–8 (2007)

39. M. Pumera, Contactless conductivity detection for microfluidics: designs and applications. *Talanta* **74**, 358–364 (2007)
40. J.M. Karlinsey, Sample introduction techniques for microchip electrophoresis: a review. *Anal. Chim. Acta* **725**, 1–13 (2012)
41. S. Yamamoto, In situ photopolymerization of polyacrylamide-based preconcentrator on a microfluidic chip for capillary electrophoresis. *Yakugaku Zasshi* **132**, 1031–1035 (2012)
42. C.L. Colyer, T. Tang, N. Chiem, D.J. Harrison, Clinical potential of microchip capillary electrophoresis systems. *Electrophoresis* **18**, 1733–1741 (1997)
43. M. Pumera, Analysis of nerve agents using capillary electrophoresis and laboratory-on-a-chip technology. *J. Chromatogr. A* **1113**, 5–13 (2006)
44. M. Koel, M. Borissova, M. Vaher, M. Kaljurand, Developments in the application of Green Chemistry principles to food analysis: capillary electrophoresis for the analysis of ingredients in food products. *Agro Food Ind Hi Tec* **22**, 27–29 (2011)
45. M. Kaljurand, M. Koel, Recent advancements on greening analytical separation. *Crit. Rev. Anal. Chem.* **41**, 2–20 (2011)
46. I. Kustos, B. Kocsis, F. Kilar, Bacterial outer membrane protein analysis by electrophoresis and microchip technology. *Expert Rev. Proteomic* **4**, 91–106 (2007)
47. H. Chen, Z.H. Fan, Two-dimensional protein separation in microfluidic devices. *Electrophoresis* **30**, 758–765 (2009)
48. G.J. Sommer, A.V. Hatch, IEF in microfluidic devices. *Electrophoresis* **30**, 742–757 (2009)
49. R.T. Turgeon, M.T. Bowser, Micro free-flow electrophoresis: theory and applications. *Anal. Bioanal. Chem.* **394**, 187–198 (2009)
50. Y.C. Pan, K. Karns, A.E. Herr, Microfluidic electrophoretic mobility shift assays for quantitative biochemical analysis. *Electrophoresis* **35**, 2078–2090 (2014)
51. R.A. Saylor, S.M. Lunte, A review of microdialysis coupled to microchip electrophoresis for monitoring biological events. *J. Chromatogr. A* **1382**, 48–64 (2015)
52. L.J. Lee, BioMEMS and micro-/nano-processing of polymers—an overview. *J. Chin. Inst. Chem. Eng.* **34**, 25–46 (2003)
53. E. Leclerc, Y. Sakai, T. Fujii, Microfluidic PDMS (polydimethylsiloxane) bioreactor for large-scale culture of hepatocytes. *Biotechnol Progr* **20**, 750–755 (2004)
54. M. Reichen, R.J. Macown, N. Jaccard, A. Super, L. Ruban, L.D. Griffin, F.S. Veraitch, N. Szita, Microfabricated modular scale-down device for regenerative medicine process development. *PLoS ONE* **7**, e22246 (2012)
55. N. Jaccard, R.J. Macown, A. Super, L.D. Griffin, F.S. Veraitch, N. Szita, Automated and online characterization of adherent cell culture growth in a microfabricated bioreactor. *JALA* **19**, 437–443 (2014)
56. H.L.N. Vu, Y.W. Li, M. Casali, D. Irimia, Z. Megeed, M.L. Yarmush, A microfluidic bioreactor for increased active retrovirus output. *Lab Chip* **8**, 75–80 (2008)
57. S.T. Yang, X.D. Zhang, Y. Wen, Microbioreactors for high-throughput cytotoxicity assays. *Curr. Opin. Drug Disc.* **11**, 111–127 (2008)
58. A. Prokop, Z. Prokop, D. Schaffer, E. Kozlov, J. Wikswo, D. Cliffl, F. Baudenbacher, NanoLiterBioReactor: long-term mammalian cell culture at nanofabricated scale. *Biomed. Microdevices* **6**, 325–339 (2004)
59. N. Korin, A. Bransky, U. Dinnar, S. Levenberg, Periodic ‘flow-stop’ perfusion microchannel bioreactors for mammalian and human embryonic stem cell long-term culture. *Biomed. Microdevices* **11**, 87–94 (2009)
60. M. Khoury, A. Bransky, N. Korin, L.C. Konak, G. Enikolopov, I. Tzchori, S. Levenberg, A microfluidic traps system supporting prolonged culture of human embryonic stem cells aggregates. *Biomed. Microdevices* **12**, 1001–1008 (2010)
61. A. Schober, U. Fernekorn, S. Singh, G. Schlingloff, M. Gebinoga, J. Hampl, A. Williamson, Mimicking the biological world: methods for the 3D structuring of artificial cellular environments. *Eng. Life Sci.* **13**, 352–367 (2013)
62. A. Schober, U. Fernekorn, B. Lubbers, J. Hampl, F. Weise, G. Schlingloff, M. Gebinoga, M. Worgull, M. Schneider, C. Augspurger, C. Hildmann, M. Kittler, M. Donahue, Applied nano

- bio systems with microfluidics and biosensors for three-dimensional cell culture. *Materialwiss Werkst* **42**, 139–146 (2011)
63. F. Lapierre, N.R. Cameron, J. Oakeshott, T. Peat, Y.G. Zhu, How to fabricate robust microfluidic systems for a dollar. *Micro/Nano Materials, Devices, and Systems*, Book Series: Proceedings of SPIE **3**, 89232Y (2013)
 64. K.V. Gernaey, F. Baganz, E. Franco-Lara, F. Kensy, U. Kruhne, M. Luebberstedt, U. Marx, E. Palmqvist, A. Schmid, F. Schubert, C.F. Mandenius, *Biotechnol. J.* **7**, 1308–1314 (2012)
 65. D. Sud, G. Mehta, K. Mehta, J. Linderman, S. Takayama, M.A. Mycek, Optical imaging in microfluidic bioreactors enables oxygen monitoring for continuous cell culture. *J. Biomed. Opt.* **11**, 050504 (2006)
 66. L.L. Bell, A.A. Seshia, C.A.B. Davidson, C.R. Lowe, Integration of holographic sensors into microfluidics for the real-time pH sensing of *L. casei*. *Euroensors XXIV Conference*, Book series: Procedia Engineering **5**, 1352–1355 (2010)
 67. S. Talaei, O. Frey, S. Psoma, P.D. van der Wal, N.F. de Rooij, Smart SU-8 pillars implemented in a microfluidic bioreactor for continuous measurement of glucose. *Euroensors XXIV Conference*, Book series: Procedia Engineering **5**, 448–451 (2010)
 68. W.B. Zimmerman, V. Tesar, H.C.H. Bandulasena, Towards energy efficient nanobubble generation with fluidic oscillation. *Curr. Opin. Colloid Interface* **16**, 350–356 (2011)
 69. S.N. Masand, L. Mignone, J.D. Zahn, D.I. Shreiber, Nanoporous membrane-sealed microfluidic devices for improved cell viability. *Biomed. Microdevices* **13**, 955–961 (2011)
 70. K.G. Lee, J. Hong, K.W. Wang, N.S. Heo, D.H. Kim, S.Y. Lee, S.J. Lee, T.J. Park, In vitro biosynthesis of metal nanoparticles in microdroplets. *ACS Nano* **6**, 6998–7008 (2012)
 71. B. Steinhaus, M.L. Garcia, A.Q. Shen, L.T. Angenent, A portable anaerobic microbioreactor reveals optimum growth conditions for the methanogen *Methanosaeta concilii*. *Appl. Environ. Microbiol.* **73**, 1653–1658 (2007)
 72. S.H. Au, S.C.C. Shih, A.R. Wheeler, Integrated microbioreactor for culture and analysis of bacteria, algae and yeast. *Biomed. Microdevices* **13**, 41–50 (2011)
 73. R. Rusconi, M. Garren, R. Stocker, Microfluidics expanding the frontiers of microbial ecology. *Ann. Rev. Biophys.* **43**, 65–91 (2014)
 74. L. Richter, C. Stepper, A. Mak, A. Reinthaler, R. Heer, M. Kast, H. Bruckl, P. Ertl, Development of a microfluidic biochip for online monitoring of fungal biofilm dynamics. *Lab Chip* **7**, 1723–1731 (2007)
 75. A. Kumar, D. Karig, R. Acharya, S. Neethirajan, P.P. Mukherjee, S. Retterer, M.J. Doktycz, Microscale confinement features can affect biofilm formation. *Microfluid. Nanofluid.* **14**, 895–902 (2013)
 76. T. Ahmed, T.S. Shimizu, R. Stocker, Microfluidics for bacterial chemotaxis. *Integr. Biol.* **2**, 604–629 (2010)
 77. J.R. Seymour, R. Simo, T. Ahmed, R. Stocker, Chemoattraction to dimethylsulfoniopropionate throughout the marine microbial food web. *Science* **329**, 342–345 (2010)
 78. S. Arora, C.S. Lim, J.Y. Foo, M.K. Sakharkar, P. Dixit, A.Q. Liu, J.M. Miao, Microchip system for monitoring microbial physiological behaviour under drug influences. *J. Eng. Med.* **223**, 777–786 (2009)
 79. J.R. Graff, S.R. Forscher-Dancause, S. Menden-Deuer, R.A. Long, D.C. Rowley, *Vibrio cholerae* exploits sub-lethal concentrations of a competitor-produced antibiotic to avoid toxic interactions. *Front Microbiol.* **4**, 8 (2013)
 80. Z.N. Hou, Y. An, K. Hjort, K. Hjort, L. Sandegren, Z.G. Wu, Time lapse investigation of antibiotic susceptibility using a microfluidic linear gradient 3D culture device. *Lab Chip* **14**, 3409–3418 (2014)
 81. J.L. Connell, J. Kim, J.B. Shear, A.J. Bard, M. Whiteley, Real-time monitoring of quorum sensing in 3D-printed bacterial aggregates using scanning electrochemical microscopy. *Proc. Natl. Acad. Sci. USA* **111**, 18255–18260 (2014)

82. H.M. Hegab, A. ElMekawy, T. Stakenborg, Review of microfluidic microbioreactor technology for high-throughput submerged microbiological cultivation. *Biomicrofluidics* **7**, 021502 (2013)
83. E.S. Kim, Directed evolution: a historical explanation into an evolutionary experimental system of nanobiotechnology, 1965-2006. *Minerva* **46**, 463-484 (2008)
84. Z.P. Çakar, Metabolic and evolutionary engineering research in Turkey and beyond. *Biotechnol. J.* **4**, 992-1002 (2009)
85. C. Alkim, B. Turanlı-Yıldız, Z.P. Çakar, Evolutionary engineering of yeast, in *Yeast metabolic engineering: methods and protocols, Method Mol Biol: 1152*, ed. by V. Mapelli (Humana Press, Dordrecht, 2014), pp. 169-183
86. Z.P. Çakar, U.O.S. Seker, C. Tamerler, M. Sonderegger, U. Sauer, Evolutionary engineering of multiple-stress resistant *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **5**, 569-578 (2005)
87. Z.P. Çakar, C. Alkim, B. Turanlı, N. Tokman, S. Akman, M. Sarikaya, C. Tamerler, L. Benbadis, J.M. Francois, Isolation of cobalt hyper-resistant mutants of *Saccharomyces cerevisiae* by *in vivo* evolutionary engineering approach. *J. Biotechnol.* **143**, 130-138 (2009)
88. G. Küçükgoze, C. Alkim, Ü. Yılmaz, H.I. Kısakesen, S. Gündüz, S. Akman, Z.P. Çakar, Evolutionary engineering and transcriptomic analysis of nickel-resistant *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **13**, 731-746 (2013)
89. C. Alkim, L. Benbadis, U. Yilmaz, Z.P. Çakar, J.M. François, Mechanisms other than activation of the iron regulon account for the hyper-resistance to cobalt of a *Saccharomyces cerevisiae* strain obtained by evolutionary engineering. *Metallomics* **5**, 1043-1060 (2013)
90. Z.P. Çakar, B. Turanlı-Yıldız, C. Alkim, Ü. Yılmaz, Evolutionary engineering of *Saccharomyces cerevisiae* for improved industrially important properties. *FEMS Yeast Res.* **12**, 171-182 (2012)
91. A. Grunberger, W. Wiechert, D. Kohlheyer, Single-cell microfluidics: opportunity for bioprocess development. *Curr. Opin. Biotechnol.* **29**, 15-23 (2014)
92. W.L. Tang, N. Xiang, D. Huang, X.J. Zhang, X.Z. Gu, Z.H. Ni, Micro-fluidics-based single-cell biophysical characterization. *Prog. Chem.* **26**, 1050-1064 (2014)
93. R.N. Zare, S. Kim, Microfluidic platforms for single-cell analysis. *Annu. Rev. Biomed. Eng.* **12**, 187-201 (2010)
94. L.F. Kang, B.G. Chung, R. Langer, A. Khademhosseini, Microfluidics for drug discovery and development: from target selection to product lifecycle management. *Drug Discov. Today* **13**, 1-13 (2008)
95. D. Lombardi, P.S. Dittrich, Advances in microfluidics for drug discovery. *Expert Opin. Drug Discov.* **5**, 1081-1094 (2010)
96. N.T. Nguyen, S.A.M. Shaegh, N. Kashaninejad, D.T. Phan, Design, fabrication and characterization of drug delivery systems based on lab-on-a-chip technology. *Adv. Drug Deliver Rev.* **65**, 1403-1419 (2013)
97. E. Cimetta, G. Vunjak-Novakovic, Microscale technologies for regulating human stem cell differentiation. *Exp. Biol. Med.* **239**, 1255-1263 (2014)
98. P.C. Blainey, The future is now: single-cell genomics of bacteria and archaea. *FEMS Microbiol. Rev.* **37**, 407-427 (2013)
99. F.J.H. Hol, C. Dekker, Zooming into see the bigger picture: microfluidic and nanofabrication tools to study bacteria. *Science* **346**, 1251821 (2014)

Chapter 25

Non-Markovian Dynamics of Qubit Systems: Quantum-State Diffusion Equations Versus Master Equations

Yusui Chen and Ting Yu

Abstract In this review we discuss recent progress in the theory of open quantum systems based on non-Markovian quantum state diffusion and master equations. In particular, we show that an exact master equation for an open quantum system consisting of a few qubits can be explicitly constructed by using the corresponding non-Markovian quantum state diffusion equation. The exact master equation arises naturally from the quantum decoherence dynamics of qubit systems collectively interacting with a colored noise. We illustrate our general theoretical formalism by the explicit construction of a three-qubit system coupled to a non-Markovian bosonic environment. This exact qubit master equation accurately characterizes the time evolution of the qubit system in various parameter domains, and paves the way for investigation of the memory effect of an open quantum system in a non-Markovian regime without any approximation.

25.1 Introduction

Recent advances in open quantum systems, quantum dissipative dynamics and quantum information science have attracted enormous interest in examining the quantum dynamics of open systems in various time domains and coupling strength ranges. Although the Lindblad master equation is a powerful theoretical tool to study an open quantum system under the Born-Markov approximation, such a Markov method will not be valid when the system is strongly coupled to an environment or the surrounding environment has a structured spectrum. In this case, it is inevitable to employ a non-Markovian quantum approach. However, unlike in

Y. Chen (✉) · T. Yu

Department of Physics and Engineering Physics, Stevens Institute of Technology,
Castle Point on Hudson, Hoboken, NJ 07030, USA
e-mail: yusui.chen@stevens.edu

T. Yu

e-mail: ting.yu@stevens.edu

the case of the standard Markov regimes, deriving the evolution equation that governs the density operator for a non-Markovian open quantum system is a long outstanding open problem. The recently developed non-Markovian quantum state diffusion (QSD) approach [1] offers an alternative way of solving the non-Markovian open quantum systems. However, from a more fundamental point of view, particularly in conjunction with the investigation of quantum decoherence and non-equilibrium quantum transport, a non-Markovian master equation approach that can be applied to both strong coupling regimes and structured medium is also highly desirable.

In this paper, we report a systematic theoretical approach that can be implemented easily for realistic quantum systems such as multiple-qubit systems [2]. The paper is organized as follows. In Sect. 25.2, we describe the principle ideas of establishing stochastic Schrödinger equations for a generic open quantum system coupled to a bosonic bath. We further present our recent work on developing a systematic non-Markovian master equation based on the stochastic non-Markovian QSD approach in Sect. 25.3. In Sect. 25.4, as examples, we study both two-qubit and three-qubit systems analytically with our new master equation approach. Some technical details are left to appendices.

25.2 Non-Markovian Quantum-State Diffusion Approach

The model under consideration is a generic open quantum system linearly coupled to a zero-temperature bosonic environment. The total Hamiltonian may be written as (setting $\hbar = 1$) [3–5]:

$$\begin{aligned} H_{tot} &= H_{sys} + H_{int} + H_{env} \\ &= H_{sys} + \sum_k \left(g_k L b_k^\dagger + g_k^* L^\dagger b_k \right) + \sum_k \omega_k b_k^\dagger b_k, \end{aligned} \quad (25.1)$$

where H_{sys} is the Hamiltonian of an arbitrary quantum system of interest, such as spins, atoms, quantum harmonic oscillators, cavities etc. The operator L is an arbitrary system operator, describing the coupling between the system of interest and its environment. $b_k(b_k^\dagger)$ is the bosonic annihilation (creation) operator of k th mode in the environment, satisfying the usual commutation relations for bosons, $[b_k, b_{k'}^\dagger] = \delta_{k,k'}$ and $[b_k, b_{k'}] = [b_k^\dagger, b_{k'}^\dagger] = 0$.

In the interaction picture with respect to the environment, the total Hamiltonian can be rewritten as (the rest of this paper is discussed in the interaction picture),

$$H_{tot} = H_{sys} + \sum_k \left(g_k L b_k^\dagger e^{i\omega_k t} + g_k^* L^\dagger b_k e^{-i\omega_k t} \right). \quad (25.2)$$

With the total Hamiltonian, the evolution for the state of the total system $|\psi_{tot}(t)\rangle$ is governed by the standard Schrödinger equation,

$$\partial_t |\psi_{tot}(t)\rangle = -i \left[H_{sys} + \sum_k \left(g_k L b_k^\dagger e^{i\omega_k t} + g_k^* L^\dagger b_k e^{-i\omega_k t} \right) \right] |\psi_{tot}(t)\rangle. \quad (25.3)$$

For a real-world problem, solving above Schrödinger equation in a non-Markovian regime is by no means an easy task due to the complexity arising from a large number of environmental variables and strong system-environment coupling. Therefore, it is desirable to develop a dynamical approach for dealing with a reduced density operator describing open quantum systems only. The quantum-state diffusion approach was developed based on a special choice of environmental basis consisting of a set of Bargmann coherent states $|z\rangle = |z_1\rangle \otimes |z_2\rangle \otimes \cdots \otimes |z_k\rangle \otimes \cdots$. For each mode, the Bargmann state is defined as

$$|z_k\rangle = e^{z_k b_k^\dagger} |0\rangle,$$

satisfying the following properties,

$$\begin{aligned} b_k |z_k\rangle &= z_k |z_k\rangle, \\ b_k^\dagger |z_k\rangle &= \frac{\partial}{\partial z_k} |z_k\rangle. \end{aligned}$$

It should be noted that the Bargmann states completeness identity is given by,

$$I = \int \frac{d^2 z}{\pi} e^{-|z|^2} |z\rangle \langle z|,$$

where $d^2 z = d^2 z_1 d^2 z_2 \cdots$. Then the state $|\psi_{tot}(t)\rangle$ for the combined total system can be expanded as,

$$\begin{aligned} |\psi_{tot}(t)\rangle &= \int \frac{d^2 z}{\pi} e^{-|z|^2} |z\rangle \langle z | \psi_{tot}(t) \rangle \\ &= \int \frac{d^2 z}{\pi} e^{-|z|^2} |\psi_t(z^*)\rangle \otimes |z\rangle, \end{aligned} \quad (25.4)$$

where

$$|\psi_t(z^*)\rangle = \langle z | \psi_{tot} \rangle.$$

Note that $|\psi_t(z^*)\rangle$ is a pure state in the system's Hilbert space, containing the complex variables z^* that will be interpreted as complex Gaussian random variables. For reasons to be explained later, $|\psi_t(z^*)\rangle$ is called a quantum trajectory

[1, 6]. Remarkably, the reduced density operator ρ_t at time point t for the system of interest can be recovered by the quantum pure state as shown below. By definition, the reduced density operator ρ_t may be obtained from ρ_{tot} by performing the partial trace over the environmental variables. For this purpose, we choose Bargmann coherent states as our basis,

$$\begin{aligned}
 \rho_t &= \text{tr}_{env}(\rho_{tot}) \\
 &= \int \frac{d^2z}{\pi} e^{-|z|^2} \langle z | \psi_{tot} \rangle \langle \psi_{tot} | z \rangle \\
 &= \int \frac{d^2z}{\pi} e^{-|z|^2} |\psi_t(z^*)\rangle \langle \psi_t(z)| \\
 &= \mathcal{M}(|\psi_t(z^*)\rangle \langle \psi_t(z)|),
 \end{aligned} \tag{25.5}$$

where the symbol

$$\mathcal{M}(\cdot) = \int \frac{d^2z}{\pi} e^{-|z|^2} (\cdot) \tag{25.6}$$

stands for the statistical average over the random variables z^* [1, 6, 7].

From (25.3), one can derive a stochastic differential equation for a quantum trajectory when the environmental bath is in a vacuum state [1],

$$\begin{aligned}
 \partial_t |\psi_t(z^*)\rangle &= -i \langle z | \left[H_{sys} + \sum_k \left(g_k L b_k^\dagger e^{i\omega_k t} + h.c. \right) \right] |\psi_{tot}(t)\rangle \\
 &= \left[-i H_{sys} + L z_t^* - i L^\dagger \sum_k g_k^* \frac{\partial}{\partial z_k^*} e^{-i\omega_k t} \right] |\psi_t(z^*)\rangle,
 \end{aligned} \tag{25.7}$$

where

$$z_t^* = -i \sum_k g_k z_k^* e^{i\omega_k t} \tag{25.8}$$

is a complex Gaussian process.

In a more general situation where the environment is in a thermal equilibrium state

$$\rho_{env}(0) = \frac{e^{-\beta \sum_k \omega_k b_k^\dagger b_k}}{Z},$$

where $\beta = \frac{1}{k_B T}$ and Z is the partition function $Z = \text{tr}(e^{-\beta \sum_k \omega_k b_k^\dagger b_k})$, the bath correlation function can be written in the following form

$$\alpha(t, s) = \sum_k |g_k|^2 \left[\coth \frac{\omega_k}{2k_B T} \cos \omega_k(t-s) - i \sin \omega_k(t-s) \right].$$

For the zero temperature case, the correlation function reduces to

$$\alpha(t, s)|_{T=0} = \sum_k |g_k|^2 e^{-i\omega_k(t-s)}. \quad (25.9)$$

It is interesting to note that the stochastic process defined in (25.8) satisfies,

$$\begin{aligned} \mathcal{M}(z_t) &= 0, \\ \mathcal{M}(z_t z_s) &= 0, \\ \mathcal{M}(z_t^* z_s) &= \alpha(t, s). \end{aligned} \quad (25.10)$$

Equation (25.10) shows that z_t^* typically represents a non-Markovian Gaussian process characterised by the correlation $\alpha(t, s)$. Taking the Lorenz spectrum as an example,

$$J(\omega) = \frac{\Gamma}{2\pi} \frac{1}{(\omega - \omega_s + \Omega_c)^2 + \gamma^2},$$

we can explicitly show that the correlation function takes a very simple form,

$$\alpha(t, s) = \frac{\Gamma\gamma}{2} e^{(-\gamma + i\Omega_c)|t-s|}, \quad (25.11)$$

which is commonly called the Ornstein-Uhlenbeck type correlation function. Ω_c represents the central frequency of the environment and $\frac{1}{\gamma}$ is the correlation-time of the environment. When the parameter $\gamma \rightarrow \infty$, the Ornstein-Uhlenbeck correlation function recovers the well-known Markov approximation described by a Dirac delta function,

$$\alpha(t, s) \approx \Gamma \delta(t, s).$$

In (25.7), the term $\frac{\partial}{\partial z_k^*} |\psi_t(z^*)\rangle$ can be cast as a functional derivative by using the chain rule,

$$\begin{aligned} -i \sum_k g_k^* L^\dagger e^{-i\omega_k t} \frac{\partial}{\partial z_k^*} |\psi_t(z^*)\rangle &= -i \sum_k g_k^* L^\dagger e^{-i\omega_k t} \int_0^t ds \frac{\partial z_s^*}{\partial z_k^*} \frac{\delta}{\delta z_s^*} |\psi_t(z^*)\rangle \\ &= -L^\dagger \int_0^t ds \alpha(t, s) \frac{\delta}{\delta z_s^*} |\psi_t(z^*)\rangle. \end{aligned}$$

By defining the O operator,

$$O(t, s, z^*)|\psi_t(z^*)\rangle = \frac{\delta}{\delta z_s^*}|\psi_t(z^*)\rangle, \quad (25.12)$$

the non-Markovian quantum-state diffusion (QSD) equation driven by the complex Gaussian process z_t^* is written as,

$$\partial_t|\psi_t(z^*)\rangle = \left(-iH_{\text{sys}} + Lz_t^* - L^\dagger \bar{O}(t, z^*)\right)|\psi_t(z^*)\rangle, \quad (25.13)$$

where $\bar{O}(t, z^*) = \int_0^t ds \alpha(t, s) O(t, s, z^*)$.

The exact non-Markovian QSD equations are generic for open quantum system models represented by (25.1). Note that these non-Markovian stochastic equations are derived from the generic microscopic Hamiltonian (25.1) or (25.2) without any approximation. For practical numerical simulations, it is useful to recast the QSD equation into a time convolutionless form by introducing a time-local operator O . The dynamical equation of the O operator can be determined by its consistency condition,

$$\frac{\partial}{\partial t} \frac{\delta}{\delta z_s^*} |\psi_t(z^*)\rangle \equiv \frac{\delta}{\delta z_s^*} \frac{\partial}{\partial t} |\psi_t(z^*)\rangle.$$

Putting the definition of O operator (25.12) and the QSD (25.13) into above equation, the dynamical equation of O operator is given by,

$$\partial_t O(t, s, z^*) = [-iH_{\text{sys}} + Lz_t^* - L^\dagger \bar{O}(t, z^*), O(t, s, z^*)] - L^\dagger \frac{\delta \bar{O}}{\delta z_s^*}. \quad (25.14)$$

with the initial condition

$$O(t, s = t, z^*) = L. \quad (25.15)$$

For many interesting models, such as dephasing models [8], multiple-qubit dissipative systems [9–12], and quantum Brownian motion [13], the exact non-Markovian QSD equations have been established [5, 14–17]. Consequently, one can study the non-Markovian behaviors of quantum decoherence and quantum entanglement, based on the numerically recovered reduced density operator ρ_t . However, from a more fundamental point of view, it is known that the corresponding non-Markovian master equations are very useful in describing quantum dissipative dynamics, quantum transport processes, and quantum decoherence. Therefore, it is of great interest to establish a generic relation between the stochastic QSD equations and their master equation counterparts.

25.3 Non-Markovian Master Equation Approach

After discussing the non-Markovian QSD approach, we will study the relationship between the non-Markovian QSD and master equation approaches in this section. As a fundamental tool, the master equation governs the evolution of the reduced density operator for an open quantum system. However, deriving a systematic non-Markovian master equation for a generic open quantum system is a rather difficult problem. Up to now, exact master equations are available only for some specific models, such as the dephasing model, qubit dissipative model, and Brownian motion model [5, 7, 13–26]. Traditionally in quantum optics, in the case of weak coupling and broadband approximation, one can adequately describe the dynamics of atoms coupled to a quantized radiation field by a Lindblad master equation [27],

$$\partial_t \rho_t = [-iH_{\text{sys}}, \rho_t] - \frac{\Gamma}{2} (L^\dagger L \rho_t + \rho_t L^\dagger L - 2L \rho_t L^\dagger), \quad (25.16)$$

where ρ_t is the reduce density operator of the system of interest, L is the Lindblad operator and Γ represents a decay rate. However, when the Born-Markov approximation ceases to be valid as shown in many cases involving strong couplings and structured spectrum distributions, non-Markovian dynamics has to be invoked. It is shown that the non-Markovian dynamics can bring new interesting physical phenomena, such as a regeneration of quantum entanglement, slow quantum coherence decay and so on. In this section, we show a systematic way of deriving the non-Markovian master equations from stochastic QSD equations.

As shown in (25.5), the reduced density matrix ρ_t can be formally recovered by taking the statistical average over all the quantum trajectories,

$$\rho_t = \mathcal{M}[|\psi_t(z^*)\rangle\langle\psi_t(z)|].$$

From this starting point, we can write down the formal master equation as,

$$\partial_t \rho_t = [-iH_{\text{sys}}, \rho_t] + L \mathcal{M}[z_t^* P_t] - L^\dagger \mathcal{M}[\bar{O}(t, z^*) P_t] + \mathcal{M}[z_t P_t] L^\dagger - \mathcal{M}[P_t \bar{O}^\dagger(t, z)] L, \quad (25.17)$$

where P_t is the stochastic projection operator $P_t(z, z^*) = |\psi_t(z^*)\rangle\langle\psi_t(z)|$.

By applying the Novikov's theorem [8],

$$\mathcal{M}[z_t^* P_t] = \int_0^t ds \mathcal{M}[z_t^* z_s] \mathcal{M}\left[\frac{\delta P_t}{\delta z_s}\right],$$

it is easy to obtain the following results,

$$\begin{aligned}\mathcal{M}[z_t^* P_t] &= \mathcal{M}[P_t \bar{O}^\dagger], \\ \mathcal{M}[z_t P_t] &= \mathcal{M}[\bar{O} P_t].\end{aligned}\quad (25.18)$$

The detailed proof of the above results can be found in the Appendix 1. Therefore, the formal master equations can be written as

$$\partial_t \rho_t = [-iH_{\text{sys}}, \rho_t] + [L, R(t)] - [L^\dagger, R^\dagger(t)]. \quad (25.19)$$

where,

$$R(t) = \mathcal{M}(P_t \bar{O}^\dagger).$$

As a note, we point out that non-Markovian master equations may provide a possibility to find an exact analytical solution. Even in numerical simulations, in some cases, such as small quantum systems, a master equation can significantly reduce computational complexity. Generally, the O operator contains noise terms, therefore, the term $\mathcal{M}(\bar{O} P_t)$ is still hard to derive analytically.

Example Here we consider the one qubit dissipative model as an example to show how to use Novikov's theorem to derive an exact master equation. The total Hamiltonian in this case is given by [1, 8],

$$H_{\text{tot}} = \frac{\omega}{2} \sigma_z + \sigma_- \sum_k g_k b_k^\dagger e^{i\omega_k t} + \sigma_+ \sum_k g_k^* b_k e^{-i\omega_k t}.$$

Then, the non-Markovian QSD (25.13) can be explicitly written as,

$$\partial_t |\psi_t(z^*)\rangle = (-i\frac{\omega}{2} \sigma_z + \sigma_- z_t^* - \sigma_+ \bar{O}) |\psi_t(z^*)\rangle. \quad (25.20)$$

And the O operator takes the form of

$$O(t, s) = f(t, s) \sigma_-, \quad (25.21)$$

where the coefficient function $f(t, s)$ satisfies the initial condition $f(t, t) = 1$ and it obeys the equation of motion,

$$\begin{aligned}\partial_t f(t, s) &= i\omega f + Ff, \\ F(t) &= \int_0^t ds \alpha(t, s) f(t, s).\end{aligned}$$

Here we choose the Ornstein-Uhlenbeck type correlation function (25.11) as an example, such that the coefficient function $F(t)$ satisfies

$$\begin{aligned}\frac{d}{dt}F(t) &= \frac{\Gamma\gamma}{2} - \gamma F + i\omega F + F^2, \\ F(0) &= 0.\end{aligned}$$

Using Novikov's theorem (25.18), we have

$$\begin{aligned}\mathcal{M}(\bar{O}P_t) &= F(t)\sigma_-\rho_t, \\ \mathcal{M}(P_t\bar{O}^\dagger) &= F^*(t)\rho_t\sigma_+.\end{aligned}$$

Then the exact master equation can be shown explicitly as

$$\partial_t\rho_t = \left[-i\frac{\omega}{2}\sigma_z, \rho_t\right] - (F\sigma_+\sigma_-\rho_t + F^*\rho_t\sigma_+\sigma_- - (F + F^*)\sigma_-\rho_t\sigma_+). \quad (25.22)$$

Next, we check its Markov limit: writing the correlation function in the form

$$\alpha(t, s) = \Gamma\delta(t, s), \quad (25.23)$$

then $F(t)$ can be calculated as

$$F(t) = \int_0^t ds \Gamma\delta(t, s)f(t, s) = \frac{\Gamma}{2}.$$

The master equation in the Markov limit is easily obtained from (25.22),

$$\partial_t\rho_t = \left[-i\frac{\omega}{2}\sigma_z, \rho_t\right] - \frac{\Gamma}{2}(\sigma_+\sigma_-\rho_t + \rho_t\sigma_+\sigma_- - 2\sigma_-\rho_t\sigma_+), \quad (25.24)$$

which clearly takes the standard Lindblad form.

25.4 Multiple-Qubit Systems

In this section, we discuss a multiple-qubit system coupled to a common bosonic environment. The multiple-qubit model is of interest in quantum information as it represents a quantum memory realised by two-level systems such as spins or atoms [28–34]. Studies of dissipation and decoherence for multiple qubit systems are useful to understand quantum decoherence control and quantum disentanglement processes. Such studies can help us to develop new theoretical and experimental

strategies to control quantum decoherence [35–37]. Here, we consider a generic N-qubit model,

$$H_{tot} = H_{sys} + L \sum_k g_k b_k^\dagger e^{i\omega_k t} + L^\dagger \sum_k g_k^* b_k e^{-i\omega_k t},$$

$$H_{sys} = \sum_j \frac{\omega_j}{2} \sigma_z^j + J_{xy} \sum_j \left(\sigma_x^j \sigma_x^{j+1} + \sigma_y^j \sigma_y^{j+1} \right),$$

where $L = \sum_j \kappa_j \sigma_-^j$ is the dissipative coupling operator of the system, κ_j is the coupling constant for j^{th} qubit. The non-Markovian QSD equation is written as

$$\partial_t |\psi_t(z^*)\rangle = \left(-iH_{sys} + Lz_t^* - L^\dagger \bar{O} \right) |\psi_t(z^*)\rangle, \quad (25.25)$$

where O operator is determined by the following equation,

$$\partial_t O(t, s, z^*) = [-iH_{sys} + Lz_t^* - L^\dagger \bar{O}(t), O(t, s)] - L^\dagger \frac{\delta \bar{O}(t)}{\delta z_s^*}, \quad (25.26)$$

together with the initial condition $O(t, s = t) = \sum_j \kappa_j \sigma_-^j$.

Differing from the previous simple example, O operator is no longer free of noise when the size of the system increases. In general, the O operator is typically involved with noise z^* . Note that O operator can be formally written in the functional expansion of noise [8],

$$O(t, s, z^*) = O_0(t, s) + \int_0^t ds_1 z_{s_1}^* O_1(t, s, s_1) + \int_0^t ds_1 \int_0^t ds_2 z_{s_1}^* z_{s_2}^* O_2(t, s, s_1, s_2) + \cdots, \quad (25.27)$$

where O_0 is the zeroth order, which does not contain noise z^* ; also, operators O_n by definition do not contain noise. For a simple example, the one qubit case, $O = f(t, s)\sigma_-$ is a special case in which O operator only contains the O_0 term. The initial conditions for each term of the O operator are [13],

$$O_0(t, s = t) = L,$$

$$O_n(t, s = t) = 0.$$

Substituting (25.27) into (25.26), we have a set of coupled differential equations for each term O_n in the O operator (Appendix 2),

$$\begin{aligned}
\partial_t O_0(t, s) &= [-iH_{\text{sys}} - L^\dagger \bar{O}_0(t), O_0(t, s)] - L^\dagger \bar{O}_1(t, s), \\
\partial_t O_1(t, s, s_1) &= [-iH_{\text{sys}} - L^\dagger \bar{O}_0(t), O_1(t, s, s_1)] - [L^\dagger \bar{O}_1(t, s_1), O_0(t, s)] \\
&\quad - L^\dagger (\bar{O}_2(t, s_1, s) + \bar{O}_2(t, s, s_1)), \\
&\quad \text{etc.},
\end{aligned} \tag{25.28}$$

together with the boundary conditions

$$\begin{aligned}
O_1(t, s, t) &= [L, O_0(t, s)], \\
O_2(t, s, t, s_1) + O_2(t, s, s_1, t) &= [L, O_1(t, s, s_1)], \\
&\quad \text{etc.}
\end{aligned}$$

As we have shown in (25.19), explicitly finding $R(t)$ is the key to determine the exact master equation. In the next section, we will exhibit the detail of deriving $R(t)$ for some important qubit systems.

25.4.1 Two-Qubit Systems

For simplicity, we take the two-qubit system as our first example to show the details of our analytical derivation. The two-qubit system has generated enormous interest due to its relevance in quantum computing and quantum information. For example, the entanglement measure for a qubit system takes a particular simple form for the two-qubit system known as concurrence [38]. The Hamiltonian for the two-qubit model is given by,

$$\begin{aligned}
H_{\text{tot}} &= H_{\text{sys}} + L \sum_k g_k b_k^\dagger e^{i\omega_k t} + L^\dagger \sum_k g_k^* b_k e^{-i\omega_k t}, \\
H_{\text{sys}} &= \frac{\omega_1}{2} \sigma_z^1 + \frac{\omega_2}{2} \sigma_z^2 + J_{xy} (\sigma_x^1 \sigma_x^2 + \sigma_y^1 \sigma_y^2), \\
L &= \kappa_1 \sigma_-^1 + \kappa_2 \sigma_-^2.
\end{aligned}$$

As discussed above, the non-Markovian QSD equation is given by,

$$\partial_t |\psi_t(z^*)\rangle = \left(-iH_{\text{sys}} + L z_t^* - L^\dagger \bar{O} \right) |\psi_t(z^*)\rangle, \tag{25.29}$$

where the O operator can be written as

$$O(t, s, z^*) = O_0(t, s) + \int_0^t ds_1 z_{s_1}^* O_1(t, s, s_1), \tag{25.30}$$

where,

$$O_0(t, s) = f_1(t, s)\sigma_-^1 + f_2(t, s)\sigma_-^2 + f_3(t, s)\sigma_z^1\sigma_-^2 + f_4(t, s)\sigma_-^1\sigma_z^2, \quad (25.31)$$

$$O_1(t, s, s_1) = f_5(t, s, s_1)\sigma_-^1\sigma_-^2. \quad (25.32)$$

Inserting the explicit form of O operator into the equation of motion (25.26),

$$\partial_t O_0 = [-iH_{\text{sys}} - L^\dagger \bar{O}_0, O_0] - L^\dagger \int_0^t d\tau \alpha(t, \tau) f_5(t, \tau, s) \sigma_-^1 \sigma_-^2, \quad (25.33)$$

$$\partial_t O_1 = [-iH_{\text{sys}}, O_1] - [L^\dagger \bar{O}_0, O_1] - [L^\dagger \bar{O}_1, O_0], \quad (25.34)$$

we have the evolution equations for the coefficient functions as

$$\begin{aligned} \partial_t f_1 &= i\omega_1 f_1 - 2iJ_{xy} f_3 + (\kappa_1 F_1 + \kappa_2 F_3) f_1 + \kappa_2 (F_4 - F_1) f_3 \\ &\quad + (\kappa_1 F_4 + \kappa_2 F_3) f_4 - \frac{\kappa_2}{2} F_5, \end{aligned} \quad (25.35)$$

$$\begin{aligned} \partial_t f_2 &= i\omega_2 f_2 - 2iJ_{xy} f_4 + (\kappa_1 F_4 + \kappa_2 F_2) f_2 + (\kappa_1 F_4 + \kappa_2 F_3) f_3 \\ &\quad + \kappa_1 (F_3 - F_2) f_4 - \frac{\kappa_1}{2} F_5, \end{aligned} \quad (25.36)$$

$$\begin{aligned} \partial_t f_3 &= i\omega_2 f_3 - 2iJ_{xy} f_1 + \kappa_1 (F_3 - F_2) f_1 + (\kappa_1 F_4 + \kappa_2 F_3) f_2 \\ &\quad + (\kappa_1 F_4 + \kappa_2 F_2) f_3 - \frac{\kappa_1}{2} F_5, \end{aligned} \quad (25.37)$$

$$\begin{aligned} \partial_t f_4 &= i\omega_1 f_4 - 2iJ_{xy} f_2 + (\kappa_1 F_4 + \kappa_2 F_3) f_1 + \kappa_2 (F_4 - F_1) f_2 \\ &\quad + (\kappa_1 F_1 + \kappa_2 F_3) f_4 - \frac{\kappa_2}{2} F_5, \end{aligned} \quad (25.38)$$

$$\begin{aligned} \partial_t f_5 &= i(\omega_1 + \omega_2) f_5 + (\kappa_1 F_1 + \kappa_1 F_4 + \kappa_2 F_2 + \kappa_2 F_3) f_5 \\ &\quad + (\kappa_1 f_1 - \kappa_1 f_4 + \kappa_2 f_2 - \kappa_2 f_3) F_5, \end{aligned} \quad (25.39)$$

where $F_j(t) = \int_0^t d\tau \alpha(t, \tau) f_j(t, \tau)$ ($j = 1, 2, 3, 4$) and $F_5(t, s) = \int_0^t d\tau \alpha(t, \tau) f_5(t, \tau, s)$. Based on the previous discussion, we have the initial conditions as

$$f_1(t, t) = \kappa_1, f_2(t, t) = \kappa_2, \quad (25.40)$$

$$f_3(t, t) = 0, f_4(t, t) = 0, \quad (25.41)$$

$$f_5(t, t, s_1) = 0, \quad (25.42)$$

and the boundary condition,

$$f_5(t, s, t) = 2(\kappa_1 f_3(t, s) + \kappa_2 f_4(t, s)). \quad (25.43)$$

For this two-qubit model, $R(t) = \mathcal{M}(P_t \bar{O}^\dagger)$ can be evaluated explicitly. By the ansatz of O operator,

$$\begin{aligned} R(t) &= \mathcal{M}(P_t \bar{O}^\dagger) \\ &= \mathcal{M}(P_t \bar{O}_0^\dagger + P_t \int_0^t ds_1 z_{s_1} \bar{O}_1^\dagger(t, s_1)). \end{aligned}$$

Since both O_0 and O_1 are free of noise, therefore, we have,

$$R(t) = \rho_t \bar{O}_0^\dagger + \int_0^t ds_1 \mathcal{M}(z_{s_1} P_t) \bar{O}_1^\dagger(t, s_1). \quad (25.44)$$

Applying Novikov's theorem (25.18), we obtain,

$$\begin{aligned} \mathcal{M}(z_{s_1} P_t) &= \int_0^t ds_2 \alpha(s_1, s_2) \mathcal{M}(O(t, s_2) P_t) \\ &= \int_0^t ds_2 \alpha(s_1, s_2) \left[O_0(t, s_2) \rho_t + \int_0^t ds_3 O_1(t, s_2, s_3) \mathcal{M}(z_{s_3}^* P_t) \right]. \end{aligned}$$

Repeatedly applying Novikov's theorem, we get,

$$\begin{aligned} \mathcal{M}(z_{s_3}^* P_t) &= \int_0^t ds_4 \alpha(s_3, s_4) \mathcal{M}(P_t O_1^\dagger(t, s_4)) \\ &= \int_0^t ds_4 \alpha(s_3, s_4) \left[\mathcal{M}(P_t O_0^\dagger(t, s_4)) + \int_0^t ds_5 \mathcal{M}(P_t z_{s_5}) O_1^\dagger(t, s_4, s_5) \right]. \end{aligned}$$

In general, repeating the Novikov theorem may generate an infinite number of terms. However, as shown below, for our two-qubit model, we can get a closed equation in a finite number of steps. Note that, if we put all the results back into $R(t)$, we have

$$\begin{aligned}
R(t) = & \rho_t \bar{O}_0^\dagger + \int_0^t ds_1 \int_0^t ds_2 \alpha(s_1, s_2) O_0(t, s_2) \rho_t \bar{O}_1^\dagger(t, s_1) \\
& + \int_0^t ds_1 \int_0^t ds_2 \alpha(s_1, s_2) \int_0^t ds_3 O_1(t, s_2, s_3) \mathcal{M}(z_{s_3}^* P_t) \bar{O}_1^\dagger(t, s_1). \quad (25.45)
\end{aligned}$$

It is easy to check that,

$$\mathcal{M}(z_{s_3}^* P_t) O_1^\dagger(t, s, s_1) = 0,$$

since

$$\begin{aligned}
O_0^\dagger O_1^\dagger &= 0, \\
O_1^\dagger O_1^\dagger &= 0.
\end{aligned}$$

The two identities are called “forbidden conditions” [2], which result in a closed noise-free $R(t)$ operator,

$$R(t) = \rho_t \bar{O}_0^\dagger + \int_0^t ds_1 \int_0^t ds_2 \alpha(s_1, s_2) O_0(t, s_2) \rho_t \bar{O}_1^\dagger(t, s_1). \quad (25.46)$$

Finally, we determine the exact non-Markovian master equation for the two-qubit system in a bosonic environment. Here, we explicitly exhibit $R(t)$ with coefficient functions:

$$\begin{aligned}
R^\dagger(t) = & (F_1 \sigma_-^1 + F_2 \sigma_-^2 + F_3 \sigma_z^1 \sigma_-^2 + F_4 \sigma_-^1 \sigma_z^2) \rho_t \\
& + \sigma_-^1 \sigma_-^2 \rho_t [r_1(t) \sigma_+^1 + r_2(t) \sigma_+^2 + r_3(t) \sigma_z^1 \sigma_+^2 + r_4(t) \sigma_+^1 \sigma_z^2], \quad (25.47)
\end{aligned}$$

where $r_j(t) = \int_0^t ds_1 \int_0^t ds_2 \alpha(s_1, s_2) f_j^*(t, s_2) F_5(t, s_1)$, ($j = 1, 2, 3, 4$).

In Fig. 25.1, we show the dynamics of quantum entanglement in the two-qubit system. For calculational simplicity, we choose the Ornstein-Uhlenbeck type of correlation function (25.11) in our numerical simulation. Figure 25.1a shows a few single-trajectory paths, numerically simulated by the non-Markovian QSD equation. In Fig. 25.1b, we use 100-trajectory averaged (dash-dotted curve) and 1000-trajectory (dashed curve) averaged reduced density operators ρ_t to simulate the entanglement dynamics. Also we show the result simulated by using the non-Markovian master equation (solid line). The non-Markovian dynamics for 1000 quantum trajectories shows a high degree of agreement with the master equation approach.

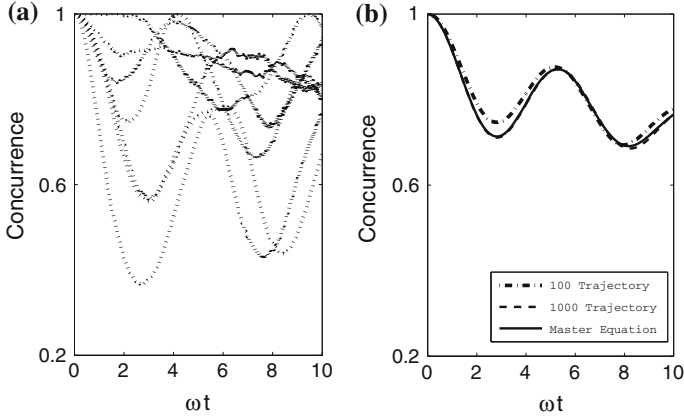


Fig. 25.1 Quantum entanglement in two-qubit system, initially prepared in a Bell state $\frac{1}{\sqrt{2}}(|10\rangle + |01\rangle)$. We show the results: **a** a set of single-trajectory evolution (*dashed*), and **b** 100 trajectories averaged (*dash-dotted*), 1000 trajectories averaged (*dashed*) and master equation (*solid*). The parameters are set as: $\omega_1 = \omega_2 = \omega$, $\kappa_1 = \kappa_2 = 1$, $J_{xy} = 0$ and $\gamma = 0.1$.

25.4.2 Three-Qubit Systems

As another interesting example, in this section, we extend our derivation for the two-qubit system to the case of a three-qubit model. With the derived non-Markovian master equation, we study quantum decoherence and quantum disentanglement in a multiple-qubit system. Although there is no convenient computable measure of entanglement for multipartite systems, we can still investigate the entanglement transfer between two qubits in a multiple-qubit system. The total Hamiltonian for the three-qubit system (shown in Fig. 25.2) is,

$$H_{tot} = H_{sys} + L \sum_k g_k b_k^\dagger e^{i\omega_k t} + L^\dagger \sum_k g_k^* b_k e^{-i\omega_k t},$$

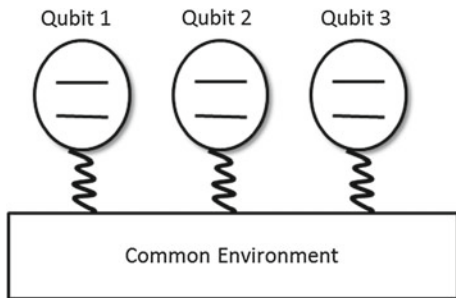
$$H_{sys} = \sum_{j=1}^3 \frac{\omega_j}{2} \sigma_z^j + J_{xy} \sum_{j=1}^2 \left(\sigma_x^j \sigma_x^{j+1} + \sigma_y^j \sigma_y^{j+1} \right),$$

where $L = \sum_{j=1}^3 \kappa_j \sigma_-^j$ is the Lindblad operator coupling the system to the environment. The non-Markovian QSD equation in this case is given by,

$$\partial_t |\psi_t(z^*)\rangle = \left(-iH_{sys} + L z_t^* - L^\dagger \bar{O} \right) |\psi_t(z^*)\rangle. \quad (25.48)$$

For the three-qubit dissipative model, the O operator contains up to the second-order of noise and can be written in a functional expansion as

Fig. 25.2 Schematic of the 3-qubit system coupled to a common environment



$$O(t, s, z^*) = O_0(t, s) + \int_0^t ds_1 z_{s_1}^* O_1(t, s, s_1) + \int_0^t ds_1 \int_0^t ds_2 z_{s_1}^* z_{s_2}^* O_2(t, s, s_1, s_2).$$

Here, we do not explicitly show the form of O operators. However, we still have the boundary conditions and initial conditions from the O operator evolution equation. The evolution equations for O_0 , O_1 and O_2 are

$$\partial_t O_0(t, s) = [-iH_{\text{sys}} - L^\dagger \bar{O}_0, O_0] - L^\dagger \bar{O}_1(t, s), \quad (25.49)$$

$$\begin{aligned} \partial_t O_1(t, s, s_1) = & [-iH_{\text{sys}}, O_1] - [L^\dagger \bar{O}_0, O_1] - [L^\dagger \bar{O}_1, O_0] \\ & - L^\dagger (\bar{O}_2(t, s, s_1) + \bar{O}_2(t, s_1, s)), \end{aligned} \quad (25.50)$$

$$\begin{aligned} \partial_t O_2(t, s, s_1, s_2) = & [-iH_{\text{sys}}, O_2] - [L^\dagger \bar{O}_0, O_2] - [L^\dagger \bar{O}_2, O_0] \\ & - [L^\dagger \bar{O}_1(t, s_1), O_1(t, s, s_2)] - [L^\dagger \bar{O}_1(t, s_2), O_1(t, s, s_1)]. \end{aligned} \quad (25.51)$$

The boundary conditions are

$$\begin{aligned} O_1(t, s, t) &= [L, O_0(t, s)], \\ O_2(t, s, t, s_1) + O_2(t, s, s_1, t) &= [L, O_1(t, s, s_1)]. \end{aligned}$$

The initial conditions are

$$\begin{aligned} O_0(t, s = t) &= L, \\ O_1(t, s = t, s_1) &= 0, \\ O_2(t, s = t, s_1, s_2) &= 0. \end{aligned}$$

And the “forbidden conditions” are

$$\begin{aligned} O_0 O_2 &= 0, \\ O_1 O_2 &= 0, \\ O_2 O_2 &= 0. \end{aligned} \quad (25.52)$$

Equations (25.49–25.51), together with their initial conditions fully determine the O operator. In order to derive the exact master equation, the last step is to evaluate $R(t) = \mathcal{M}[P_t \bar{O}^\dagger]$ in the form

$$R(t) = \rho_t \bar{O}_0^\dagger + \int_0^t ds_1 \mathcal{M}(z_{s_1} P_t) \bar{O}_1^\dagger(t, s_1) + \int_0^t ds_1 \int_0^t ds_3 \mathcal{M}(z_{s_1} z_{s_3} P_t) \bar{O}_2^\dagger(t, s_1, s_3). \quad (25.53)$$

Similar to the two-qubit case, employment of the Novikov theorem (25.18) and the forbidden conditions (25.52) leads to $R(t)$ of (25.53) in the form

$$\begin{aligned} R &= \rho_t \bar{O}_0^\dagger + \int_0^t ds_1 \int_0^t ds_2 \alpha(s_1, s_2) O_0(t, s_2) \rho_t \bar{O}_1^\dagger(t, s_1) \\ &+ \int_0^t ds_1 \int_0^t ds_2 \int_0^t ds_3 \int_0^t ds_4 \alpha(s_1, s_2) \alpha(s_3, s_4) O_1(t, s_2, s_3) \rho_t \bar{O}_0^\dagger(t, s_4) \bar{O}_1^\dagger(t, s_1) \\ &+ \int_0^t ds_1 \int_0^t ds_2 \int_0^t ds_3 \int_0^t ds_4 \alpha(s_1, s_3) \alpha(s_2, s_4) O_0(t, s_3) O_0(t, s_4) \rho_t \bar{O}_2^\dagger(t, s_1, s_2) \\ &+ \int_0^t ds_1 \int_0^t ds_2 \int_0^t ds_3 \int_0^t ds_4 \alpha(s_1, s_3) \alpha(s_2, s_4) O_1(t, s_3, s_4) \rho_t \bar{O}_2^\dagger(t, s_1, s_2). \end{aligned} \quad (25.54)$$

The detailed derivation of this can be found in Appendix 3. With the exact form of $R(t)$, the exact non-Markovian master equation may be explicitly obtained,

$$\partial_t \rho_t = [-iH_{\text{sys}}, \rho_t] + [L, R] - [L^\dagger, R^\dagger]. \quad (25.55)$$

It should be noted that in the above derivation, the correlation function $\alpha(t, s)$ can have an arbitrary form. Therefore our derivation of the exact master equation is completely general.

In Fig. 25.3, we plot the entanglement dynamics of a pair of qubits in the 3-qubit model with four different initial states, including a separate state and three

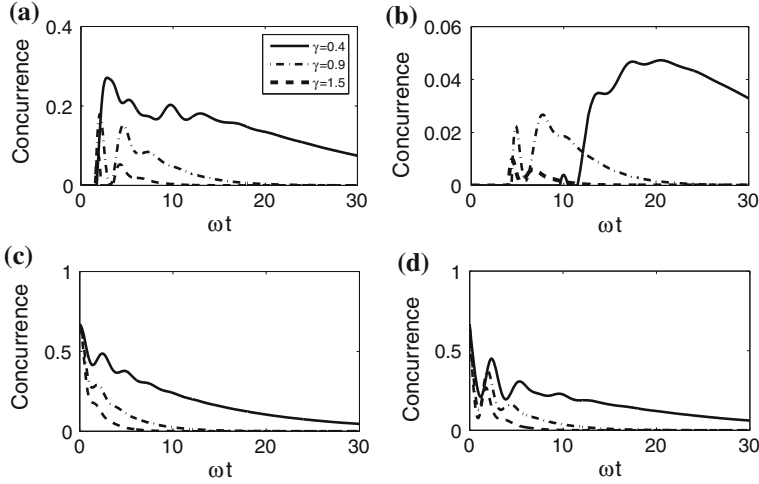


Fig. 25.3 The dynamics of entanglement between qubit 1 and qubit 2 (see Fig. 25.2) with different initial states. **a** $|111\rangle$, **b** $(|111\rangle + |000\rangle)/\sqrt{2}$, **c** $(|100\rangle + |010\rangle + |001\rangle)/\sqrt{3}$, **d** $(|110\rangle + |101\rangle + |011\rangle)/\sqrt{3}$

maximally entangled states (GHZ state and W state). Without loss of generality, the concurrence between qubit 1 and qubit 2 is studied. In Fig. 25.3a, b, since the initial three-qubit states are $|111\rangle$ and $\frac{1}{\sqrt{2}}(|111\rangle + |000\rangle)$ respectively, there is no entanglement between the qubit-pair considered. When we choose different memory times $1/\gamma$ (taking Ornstein-Uhlenbeck noise as an example again (25.11)), the degrees of the generated quantum entanglement are different. When $\gamma = 0.4$, a typical non-Markovian regime, the maximally generated entanglement is much higher than that in the case with $\gamma = 1.5$ representing the Markov limit. In Fig. 25.3c, d, the initial GHZ state of the three-qubit system is maximally entangled, and the reduced density matrices for qubits 1 and 2 are also entangled. When $\gamma = 0.4$, the early revival of entanglement in both cases is a typical non-Markovian feature.

Furthermore, we consider the entanglement transfer between two pairs of qubits. In Fig. 25.4, we prepare a Bell state for the qubit-pair 1 and 2. The idea is to observe the way entanglement transfers from qubits 1 and 2 to qubits 2 and 3. Because of the symmetry of the model, the behaviors of quantum entanglements C_{13} and C_{23} are identical. In Fig. 25.4a, c, the correlation parameter $\gamma = 0.4$ is fixed, therefore these two graphs show the short-time behavior of non-Markovian entanglement evolution. For different initial states, the speed of generating quantum entanglement is also different. In Fig. 25.4b, d, with the environment close to the Markov limit with $\gamma = 1.5$, we see that the entanglement drops to its final steady state quickly, as expected. It is interesting to note that the quantum entanglement between a pair of qubits does not actually vanish for a long time. Contrary to the two-qubit system

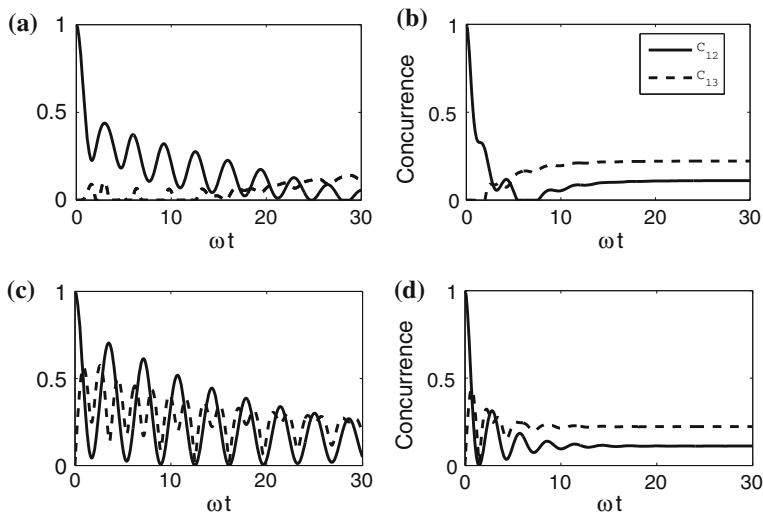


Fig. 25.4 The dynamics of quantum entanglement between three qubit-pairs C_{12} (qubit 1 and 2, solid) and C_{13} (qubit 1 and 3, dashed). Left column shows a non-Markovian regime with $\gamma = 0.4$. Right column shows a regime close to Markov limit (we choose $\gamma = 1.5$). **a** and **b** use the same initial state $(|11\rangle + |00\rangle) \otimes |0\rangle/\sqrt{2}$; while **c** and **d** use the initial state $(|10\rangle + |01\rangle) \otimes |0\rangle/\sqrt{2}$

dissipatively coupled to a bosonic environment, most two-qubit entangled states will be disentangled eventually, except for the Bell state $(|10\rangle - |01\rangle)/\sqrt{2}$, which preserves the quantum information due to the decoherence-free subspace. However, as shown in Fig. 25.4b, d, quantum entanglement can be stored in a pair of qubits robustly. This result can be naturally extended to N-qubit systems; the capacity of storing quantum information will increase as the size of quantum system is enlarged.

25.4.3 A Note on General N-Qubit Systems

We remark that the previous derivations for the two-qubit and the three-qubit systems can be extended to the more general case of N-qubit systems, with the Lindblad operator $L = \sum_j \kappa_j \sigma_j^-$. The general procedure for generalizing our results to N-qubit systems is highlighted as follows. First, we need to determine the maximum order of noise in the O operator. It is easy to prove that $L^{N+1} = 0$ for a N-qubit system, and the last term of the O operator, O_{N-1} , must be in the form of L^N . And the highest order of noise in the O operator is $N - 1$ [11]. For example, the O operator contains the first-order noise in the two-qubit model, and up to the second order of noise in the three-qubit models. Similarly, there is at most $N - 1$ order of noise for the N-qubit models. Second, we need to determine the “forbidden conditions”.

The close condition for qubit is $(\sigma_-^j)^2 = 0$. One can see that if two O operator components O_j and O_k satisfy this condition $j + k > N - 2$, then $O_j O_k = 0$. Therefore, generally, one can obtain the explicit form of $R(t) = \mathcal{M}(P_t \bar{O}^\dagger)$, by calculating

$$\mathcal{M}(z_{s_1} \cdots z_{s_{2j-1}} P_t) = \int_0^t \cdots \int_0^t ds_2 \cdots ds_{2j} \left(\prod_j \alpha(s_{2j-1}, s_{2j}) \right) \mathcal{M} \left[\left(\prod_j \frac{\delta}{\delta z_{s_{2j}}^*} \right) P_t \right]. \quad (25.56)$$

Once the closed form of the $R(t)$ operator is obtained, the exact master equation is determined.

25.5 Conclusion

In this paper, based on the non-Markovian QSD approach, we analytically and numerically investigate multiple-qubit systems dissipatively coupled to a non-Markovian zero-temperature bosonic environment. We have explicitly demonstrated how to establish an exact non-Markovian master equation from the corresponding quantum state diffusion equation. Our approach is very flexible in the sense that it can be readily modified to solve many other types of models such as hybrid systems consisting of qubits, qutrits, continuous variable systems and multiple-environment systems, to name a few [39]. The time-local exact master equation approach studied in this paper represents a new advance in our investigations of non-Markovian quantum dynamics and non-equilibrium quantum dynamics. We expect that our newly developed theoretical approach will be useful in attacking many real-world problems.

Appendix 1

Here we supply a proof of the Novikov theorem. To make the proof more generic, we calculate the term $\mathcal{M}(z_\tau P_t)$, where τ and t are two independent time indexes. In this, $\mathcal{M}(z_t P_t)$ is the limit case in which $\tau = t$. By the definition of ensemble average in (25.6), we have [8]

$$\mathcal{M}(z_\tau P_t) = \int \frac{d^2 z}{\pi} e^{-|z|^2} z_\tau P_t. \quad (25.57)$$

where $|z|^2 = \sum_k |z_k|^2$ and $d^2z = d^2z_1 d^2z_2 \dots$. With the definition of $z_\tau = i \sum_k g_k^* z_k e^{-i\omega_k \tau}$, we have

$$\mathcal{M}(z_\tau P_t) = \int \frac{d^2z_1}{\pi} \frac{d^2z_2}{\pi} \dots \prod_n e^{-|z_n|^2} \left(i \sum_k g_k^* z_k e^{-i\omega_k \tau} \right) P_t.$$

Since all z_k are independent to each other, the above integration can be simplified as

$$\mathcal{M}(z_\tau P_t) = i \sum_k g_k^* e^{-i\omega_k \tau} \left(\prod_{n \neq k} \int \frac{d^2z_n}{\pi} e^{-|z_n|^2} \right) \int \frac{dz_k dz_k^*}{\pi} e^{-|z_k|^2} z_k P_t.$$

Integrating by parts, then we have,

$$\begin{aligned} & \int \frac{dz_k dz_k^*}{\pi} e^{-|z_k|^2} z_k P_t \\ &= \int \frac{dz_k dz_k^*}{\pi} \left(-\frac{\partial}{\partial z_k^*} e^{-|z_k|^2} \right) P_t \\ &= \int \frac{dz_k dz_k^*}{\pi} \left[\left(-\frac{\partial}{\partial z_k^*} e^{-|z_k|^2} P_t \right) + e^{-|z_k|^2} \frac{\partial}{\partial z_k^*} P_t \right] \\ &= \int \frac{dz_k dz_k^*}{\pi} e^{-|z_k|^2} \frac{\partial}{\partial z_k^*} P_t. \end{aligned}$$

Then

$$\mathcal{M}(z_\tau P_t) = i \sum_k g_k^* e^{-i\omega_k \tau} \int \frac{d^2z}{\pi} e^{-|z|^2} \frac{\partial}{\partial z_k^*} P_t.$$

Using the functional derivative chain rule,

$$\begin{aligned} \mathcal{M}(z_\tau P_t) &= i \sum_k g_k^* e^{-i\omega_k \tau} \int \frac{d^2z}{\pi} e^{-|z|^2} \int_0^t ds \frac{\partial z_s^*}{\partial z_k^*} \frac{\delta}{\delta z_s^*} P_t \\ &= \int \frac{d^2z}{\pi} e^{-|z|^2} \int_0^t ds \alpha(\tau, s) O(t, s, z^*) P_t, \\ \mathcal{M}(z_\tau P_t) &= \int_0^t ds \alpha(\tau, s) \mathcal{M}[O(t, s, z^*) P_t]. \end{aligned}$$

Now we have the Novikov theorem,

$$\begin{aligned}\mathcal{M}(z_\tau P_t) &= \int_0^t ds \mathcal{M}(z_\tau z_s^*) \mathcal{M}[O(t, s, z^*) P_t], \\ \mathcal{M}(z_t^* P_t) &= \int_0^t ds \mathcal{M}(z_\tau^* z_s) \mathcal{M}[P_t O^\dagger(t, s, z)].\end{aligned}$$

In the limit $\tau = t$, we obtain

$$\begin{aligned}\mathcal{M}(z_t P_t) &= \mathcal{M}(\bar{O}(t, z^*) P_t), \\ \mathcal{M}(z_t^* P_t) &= \mathcal{M}(P_t \bar{O}^\dagger(t, z)).\end{aligned}\tag{25.58}$$

Appendix 2

Inserting the expansion series of O operator (25.27) into the O operator evolution (25.26), we have

$$\partial_t O(t, s) = \partial_t O_0(t, s) + z_t^* O_1(t, s, t) + \int_0^t ds_1 z_{s_1}^* \partial_t O_1(t, s, s_1) + \cdots, \tag{25.59}$$

for the left hand side. Furthermore, the right hand side of (25.26) can be expanded as

$$\begin{aligned}& [-iH_{\text{sys}} + Lz_t^* - L^\dagger \bar{O}, O] - L^\dagger \frac{\delta \bar{O}}{\delta z_s^*} \\ &= [-iH_{\text{sys}} + Lz_t^* - L^\dagger \bar{O}_0, O_0] - L^\dagger \frac{\delta}{\delta z_s^*} \int_0^t d\tau \alpha(t, \tau) \int_0^t ds_1 z_{s_1}^* O_1(t, \tau, s_1) \\ &\quad + [-iH_{\text{sys}} + Lz_t^*, \int_0^t ds_1 z_{s_1}^* O_1] - [L^\dagger \bar{O}_0, \int_0^t ds_1 z_{s_1}^* O_1] - [L^\dagger \int_0^t ds_1 z_{s_1}^* \bar{O}_1, \bar{O}_0] \\ &\quad - L^\dagger \frac{\delta}{\delta z_s^*} \int_0^t d\tau \alpha(t, \tau) \int_0^t ds_1 \int_0^t ds_2 z_{s_1}^* z_{s_2}^* O_2(t, \tau, s_1, s_2) \\ &\quad + \cdots.\end{aligned}\tag{25.60}$$

By the definition $\bar{O} = \int_0^t ds \alpha(t, s) O(t, s, z^*)$, we can calculate the terms

$$\begin{aligned} & \frac{\delta}{\delta z_s^*} \int_0^t d\tau \alpha(t, \tau) \int_0^t ds_1 z_{s_1}^* O_1(t, \tau, s_1) \\ &= \int_0^t d\tau \alpha(t, \tau) \int_0^t ds_1 \delta(s, s_1) O_1(t, \tau, s_1) \\ &= \int_0^t d\tau \alpha(t, \tau) O_1(t, \tau, s) = \bar{O}_1(t, s), \end{aligned}$$

and

$$\begin{aligned} & \frac{\delta}{\delta z_s^*} \int_0^t d\tau \alpha(t, \tau) \int_0^t ds_1 \int_0^t ds_2 z_{s_1}^* z_{s_2}^* O_2(t, \tau, s_1, s_2) \\ &= \int_0^t d\tau \alpha(t, \tau) \int_0^t ds_1 \int_0^t ds_2 z_{s_1}^* \delta(s, s_2) O_2(t, \tau, s_1, s_2) \\ &+ \int_0^t d\tau \alpha(t, \tau) \int_0^t ds_1 \int_0^t ds_2 z_{s_2}^* \delta(s, s_1) O_2(t, \tau, s_1, s_2) \\ &= \int_0^t d\tau \alpha(t, \tau) \int_0^t ds_1 z_{s_1}^* O_2(t, \tau, s_1, s) + \int_0^t d\tau \alpha(t, \tau) \int_0^t ds_2 z_{s_2}^* O_2(t, \tau, s, s_2) \\ &= \int_0^t d\tau \alpha(t, \tau) \int_0^t ds_1 z_{s_1}^* (O_2(t, \tau, s_1, s) + O_2(t, \tau, s, s_1)) \\ &= \int_0^t ds_1 z_{s_1}^* (\bar{O}_2(t, s_1, s) + \bar{O}_2(t, s, s_1)). \end{aligned}$$

Equating the two sides for each order of noise z^* , we obtain a set of dynamical equations for the O_n ($n = 1, 2, \dots$). For the non-noise term, we have

$$\partial_t O_0 = [-iH_{sys} - L^\dagger \bar{O}_0, O_0] - L^\dagger \bar{O}_1.$$

For the first-order noise terms, we have

$$\begin{aligned} & \int_0^t ds_1 z_{s_1}^* \partial_t O_1 \\ &= \int_0^t ds_1 z_{s_1}^* \left\{ [-iH_{\text{sys}} L^\dagger \bar{O}_0, O_1] - [L^\dagger \bar{O}_1, O_0] - L^\dagger (\bar{O}_2(t, s_1, s) + \bar{O}_2(t, s, s_1)) \right\}, \end{aligned}$$

and the evolution equation for O_1 is obtained as

$$\partial_t O_1 = [-iH_{\text{sys}} - L^\dagger \bar{O}_0, O_1] - [L^\dagger \bar{O}_1, O_0] - L^\dagger (\bar{O}_2(t, s_1, s) + \bar{O}_2(t, s, s_1)).$$

Similarly, the set of coupled dynamical equations for all O_n can be determined sequentially. For the terms containing z_t^* , the boundary conditions can be obtained as

$$\begin{aligned} O_1(t, s, t) &= [L, O_0(t, s)], \\ O_2(t, s, s_1, t) + O_2(t, s, t, s_1) &= [L, O_1(t, s, s_1)], \\ &\text{etc.} \end{aligned}$$

Appendix 3

In order to explicitly derive the $R(t)$ for the three-qubit system model, we need to calculate two terms $\mathcal{M}\{z_{s_1} P_t\}$ and $\mathcal{M}\{z_{s_1} z_{s_3} P_t\}$. Since the term $\mathcal{M}\{z_{s_1} z_{s_3} P_t\}$ contains second order of noise, it can be evaluated by using Novikov's theorem twice (25.18).

$$\begin{aligned} \mathcal{M}\{z_{s_1} P_t\} &= \int_0^t ds_2 \alpha(s_1, s_2) \mathcal{M}\{O(t, s_2) P_t\} \\ &= \int_0^t ds_2 \alpha(s_1, s_2) \left[O_0(t, s_2) \rho_t + \int_0^t ds_3 O_1(t, s_2, s_3) \mathcal{M}\{z_{s_3}^* P_t\} \right] \\ &\quad + \int_0^t ds_2 \alpha(s_1, s_2) \int_0^t ds_3 \int_0^t ds_5 O_2(t, s_2, s_3, s_5) \mathcal{M}\{z_{s_3}^* z_{s_5}^* P_t\}, \\ \mathcal{M}\{z_{s_1} z_{s_3} P_t\} &= \int_0^t ds_2 \alpha(s_1, s_2) \mathcal{M}\{z_{s_3} O(t, s_2) P_t\} \\ &= \int_0^t ds_2 \int_0^t ds_4 \alpha(s_1, s_2) \alpha(s_3, s_4) \mathcal{M}\left\{ \frac{\delta O(t, s_2)}{\delta z_{s_4}^*} P_t \right\} \\ &\quad + \int_0^t ds_2 \int_0^t ds_4 \alpha(s_1, s_2) \alpha(s_3, s_4) \mathcal{M}\{O(t, s_2) O(t, s_4) P_t\}. \end{aligned}$$

After eliminating the zero terms by the “forbidden conditions”, $R(t)$ can be explicitly shown as (25.54).

References

1. L. Diósi, N. Gisin, W.T. Strunz, Non-Markovian quantum state diffusion. *Phys. Rev. A* **58**, 1699–1712 (1998)
2. Y. Chen, J.Q. You, T. Yu, Exact non-Markovian master equations for multiple qubit systems: quantum-trajectory approach. *Phys. Rev. A* **90**, 052104 (2014)
3. C. Gardiner, P. Zoller, *Quantum Noise* (Springer-Verlag, Berlin Heidelberg, 2004)
4. Caldeira Ao, A.J. Leggett, Path integral approach to quantum Brownian motion. *Phys. A* **121**, 587–616 (1983)
5. B.L. Hu, J.P. Paz, Y. Zhang, Quantum Brownian motion in a general environment: exact master equation with nonlocal dissipation and colored noise. *Phys. Rev. D* **45**, 2843–2861 (1992)
6. W.T. Strunz, L. Diósi, N. Gisin, Open system dynamics with non-Markovian quantum trajectories. *Phys. Rev. Lett.* **82**, 1801–1805 (1999)
7. W.T. Strunz, L. Diósi, N. Gisin, T. Yu, Quantum trajectories for Brownian motion. *Phys. Rev. Lett.* **83**, 4909–4913 (1999)
8. T. Yu, L. Diósi, N. Gisin, W.T. Strunz, Non-Markovian quantum-state diffusion: perturbation approach. *Phys. Rev. A* **60**, 91–103 (1999)
9. J. Jing, T. Yu, Non-Markovian relaxation of a three-level system: quantum trajectory approach. *Phys. Rev. Lett.* **105**, 240403 (2010)
10. X. Zhao, J. Jing, B. Corn, T. Yu, Dynamics of interacting qubits coupled to a common bath: non-Markovian quantum-state-diffusion approach. *Phys. Rev. A* **84**, 032101 (2011)
11. J. Jing, X. Zhao, J.Q. You, T. Yu, Time-local quantum-state-diffusion equation for multilevel quantum systems. *Phys. Rev. A* **85**, 042106 (2012)
12. J. Jing, X. Zhao, J.Q. You, W.T. Strunz, T. Yu, Many-body quantum trajectories of non-Markovian open systems. *Phys. Rev. A* **88**, 052122 (2013)
13. T. Yu, Non-Markovian quantum trajectories versus master equations: finite-temperature heat bath. *Phys. Rev. A* **69**, 062107 (2004)
14. C. Anastopoulos, B.L. Hu, Two-level atom-field interaction: exact master equations for non-Markovian dynamics, decoherence, and relaxation. *Phys. Rev. A* **62**, 033821 (2000)
15. C. Chou, T. Yu, B.L. Hu, Exact master equation and quantum decoherence of two coupled harmonic oscillators in a general environment. *Phys. Rev. E* **77**, 011112 (2008)
16. C. Anastopoulos, S. Shresta, B.L. Hu, Non-Markovian entanglement dynamics of two qubits interacting with a common electromagnetic field. *Quant. Inf. Process.* **8**, 549–563 (2009)
17. C.H. Fleming, A. Roura, B.L. Hu, Exact analytical solutions to the master equation of quantum Brownian motion for a general environment. *Ann. Phys.* **326**, 1207–1258 (2011)
18. H.M. Wiseman, Stochastic quantum dynamics of a continuously monitored laser. *Phys. Rev. A* **47**, 5180–5192 (1993)
19. H.P. Breuer, W. Huber, F. Petruccione, Fast Monte Carlo algorithm for nonequilibrium systems. *Phys. Rev. E* **53**, 4232–4235 (1996)
20. N. Gisin, I.C. Percival, The quantum-state diffusion model applied to open systems. *J. Phys. A: Math. Gen.* **25**, 5677–5691 (1992)
21. B.L. Hu, J.P. Paz, Y. Zhang, Quantum Brownian motion in a general environment II: nonlinear coupling and perturbative approach. *Phys. Rev. D* **47**, 1576–1594 (1993)
22. S. Lin, C. Chou, B.L. Hu, Disentanglement of two harmonic oscillators in relativistic motion. *Phys. Rev. D* **78**, 125025 (2008)

23. C.H. Fleming, B.L. Hu, Non-Markovian dynamics of open quantum systems: stochastic equations and their perturbative solutions. *Ann. Phys.* **327**, 1238–1276 (2012)
24. J.P. Paz, A.J. Roncaglia, Dynamics of the entanglement between two oscillators in the same environment. *Phys. Rev. Lett.* **100**, 220401 (2008)
25. W. Zhang, P. Lo, H. Xiong, M. Tu, F. Nori, General non-Markovian dynamics of open quantum systems. *Phys. Rev. Lett.* **109**, 170402 (2012)
26. W.T. Strunz, T. Yu, Convolutionless non-Markovian master equations and quantum trajectories: Brownian motion. *Phys. Rev. A* **69**, 052115 (2004)
27. G. Lindblad, Entropy, information and quantum measurements. *Commun. Math. Phys.* **33**, 305–322 (1973)
28. J. Gea-Banacloche, Qubit-qubit interaction in quantum computers. *Phys. Rev. A* **57**, R1 (1998)
29. F. Setiawan, H. Hui, J.P. Kestner, X. Wang, S.D. Sarma, Robust two-qubit gates for exchange-coupled qubits. *Phys. Rev. B* **89**, 085314 (2014)
30. A.C. Doherty, M.P. Wardrop, Two-qubit gates for resonant exchange qubits. *Phys. Rev. Lett.* **111**, 050503 (2013)
31. L.A. Wu, D.A. Lidar, Dressed qubits. *Phys. Rev. Lett.* **91**, 097904 (2003)
32. A.M. Childs, W. van Dam, Quantum algorithms for algebraic problems. *Rev. Mod. Phys.* **82**, 1–52 (2010)
33. P. Kok, W.J. Munro, K. Nemoto, T.C. Ralph, J.P. Dowling, G.J. Milburn, Linear optical quantum computing with photonic qubits. *Rev. Mod. Phys.* **79**, 135–174 (2007)
34. P. van Loock, S.L. Braunstein, Multipartite entanglement for continuous variables: a quantum teleportation network. *Phys. Rev. Lett.* **84**, 3482–3485 (2000)
35. F. Xue, S.X. Yu, C.P. Sun, Quantum control limited by quantum decoherence. *Phys. Rev. A* **73**, 013403 (2006)
36. A.M. Brańczyk, P.E.M.F. Mendonça, A. Gilchrist, A.C. Doherty, S.D. Bartlett, Quantum control of a single qubit. *Phys. Rev. A* **75**, 012329 (2007)
37. F. Delgado, Quantum control on entangled bipartite qubits. *Phys. Rev. A* **81**, 042317 (2010)
38. W.K. Wootters, Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.* **80**, 2245–2248 (1998)
39. Y. Chen, J.Q. You, T. Yu, Generic non-Markovian master equations for multilevel systems. Submitted to *Phys. Rev. A* (2015)

Chapter 26

Computing with Emerging Nanotechnologies

M. Altun

Abstract As current CMOS based technologies are approaching their anticipated limits, emerging nanotechnologies start to replace their role in electronic circuits. New computing models have been proposed. This chapter overviews both deterministic and stochastic computing models targeting nano-crossbar switching arrays and emerging low-density circuits. These models are demonstrated with implementations using Boolean and arithmetic logic. Performance parameters of the models such as area, speed, and accuracy, are also evaluated in comparison with those of conventional circuits.

26.1 Introduction

In 1965, Gordon Moore made an influential prediction about CMOS size shrinking, formulated as the Moore Law stating that the number of transistors on a chip doubles every 18–24 months [1]. His prediction has kept its validity for decades. Nowadays this trend has reached a critical point and it is widely accepted that the trend will end in the next decade. Even Gordon accepted that his prediction will lose its validity in near future [2]. At this point, research is shifting to novel forms of nanoarchitectures including nano-crossbar arrays and probabilistic/stochastic circuits and systems [3–5]. Such technologies have apparent advantages over conventional CMOS technologies, such as high performance capacity and easy manufacturability.

Nano-crossbar arrays are regular and dense structures that are generally fabricated by self-assembly as opposed to lithography based conventional and relatively costly CMOS fabrication techniques [6]. Conventional lithographic techniques face severe challenges for emerging nanotechnologies due to their need for directed

M. Altun (✉)

Electronics and Communication Engineering Department, Istanbul Technical University,
Ayazaga Campus, 34469 Maslak, Istanbul, Turkey
e-mail: altunmus@itu.edu.tr

manipulation of molecules which is quite costly in nanoscale. Along with these advantages in terms of circuit size and fabrication, nano-crossbar arrays have drawbacks including reliability issues and CMOS integration problems, standing against commercial production. While reliability of nano-crossbars have been satisfactorily improved using reconfigurable architectures and new defect tolerance techniques [5, 7, 8], discussed later in this chapter, CMOS integration is still a major problem; CMOL is the strongest candidate for this problem [9]. Indeed, if all parts of a computing system can be successfully realized with nano-crossbar arrays then there will be no need for integration, but current state-of-the art has not reached this point, hopefully in the next decade.

The concept of using stochastic computing models is not new; it dates back to a seminal paper by John von Neumann in 1956 [10]. With the advent of a variety of types of emerging nanoscale technologies, the model has found renewed interest [3, 11]. Unlike conventional CMOS that is solely based on deterministic operations, stochastic circuits use probabilities as inputs and outputs. This feature is invaluable to cope with uncertainties seen in emerging nanotechnologies in the form of variability, reliability, and noise problems [4]. Stochastic computing also offers smaller circuit implementations for arithmetic functions using much fewer transistors compared to conventional CMOS circuits. This feature attracts low density obligated technologies such as printed/flexible electronics [12]. Here, the main drawback is high error rates seen in stochastic computing by nature. Methods, as discussed later in this chapter, have been proposed to improve it.

In this chapter, we focus on computing models for nano-crossbar arrays and stochastic circuits. These models are demonstrated with implementations using Boolean and arithmetic logic. Performance parameters of the models such as area, speed, and accuracy, are also evaluated in comparison with those of conventional circuits. This chapter is organized as follows. In Sect. 26.2, we investigate nano-crossbar array based computing models. We present Boolean function implementation and defect tolerance techniques in Sects. 26.2.1 and 26.2.2, respectively. We evaluate the techniques on standard benchmark circuits in Sect. 26.2.3. In Sect. 26.3, we introduce stochastic computing models and their application areas. We present techniques to reduce error rates and to achieve error free stochastic computing in Sects. 26.3.1 and 26.3.2, respectively. In Sect. 26.4, we present conclusions.

26.2 Computing with Nano-crossbar Arrays

Unlike conventional CMOS that can be patterned in complex ways with lithography, self-assembled nanoscale systems generally consist of regular structures. Logical functions and memory elements are achieved with arrays of crossbar-type switches. In this study, we target this type of switching arrays where each cross-point behaves as a switch, either two-terminal or four-terminal. This is illustrated in Fig. 26.1. Depending on the used technology, a two-terminal switch based

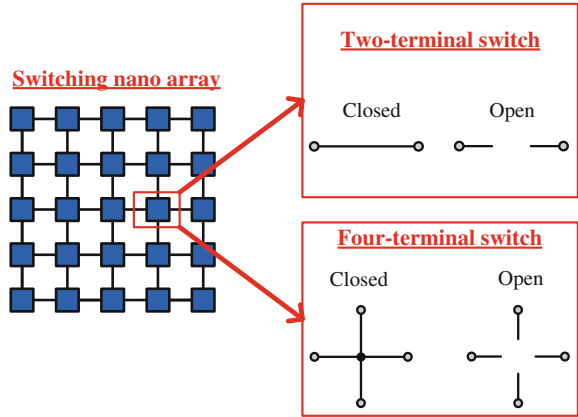


Fig. 26.1 A switching crossbar nanoarray modeled with two-terminal and four-terminal switches

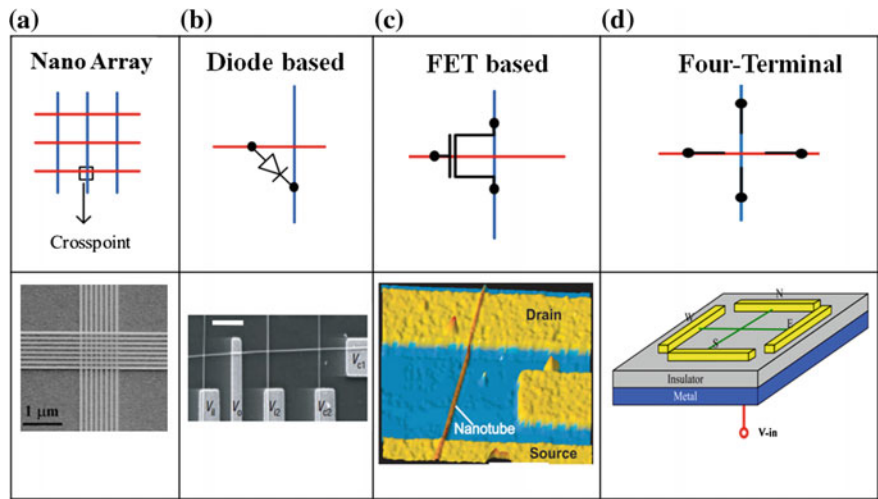


Fig. 26.2 a Nano-crossbar [14] based on b diode [14], c FET [15], d four-terminal switch [23] crosspoints

crosspoint can be modeled as a diode [13, 14] or a FET [15, 16]. This is illustrated in Fig. 26.2. Note that both diode and FET based crosspoints conduct current in one direction. However, four-terminal switches conduct current in multiple directions.

We implement Boolean functions by considering array sizes. Table 26.1 compares different implementation methodologies for few XOR functions (Parity functions) regarding array sizes. The columns “diode based” and “FET based” represent two-terminal switch based implementation methodologies. These methodologies have been proposed to implement simple logic functions [17, 18].

Table 26.1 Array sizes for nanoarray computing models; $\text{XOR2} = x_1 \oplus x_2$, $\text{XOR3} = x_1 \oplus x_2 \oplus x_3$, and $\text{XOR4} = x_1 \oplus x_2 \oplus x_3 \oplus x_4$

	Two-terminal switch based nanoarray models		Four-terminal switch based nanoarray models	
	Diode based (Optimal) [13]	FET based (Optimal) [16]	Four-terminal based [23]	Four-terminal based (Optimal) [36]
XOR2	2×5 array 10 switches	4×4 array 16 switches	2×2 array 4 switches	2×2 array 4 switches
XOR3	4×7 array 28 switches	6×8 array 48 switches	4×4 array 16 switches	3×3 array 9 switches
XOR4	8×9 array 72 switches	8×16 array 128 switches	8×8 array 64 switches	3×5 array 15 switches

In this study, we generalize them to be applicable for any given Boolean function with offering optimal array size formulations. The last two columns represent four-terminal switch based implementation methodologies that offer favorably better results.

Defect rates are much higher for nano-crossbars compared to conventional CMOS circuits [19]. Therefore developing new defect tolerance techniques for nano-crossbars is a must, especially for high defect rates up to 20 %. Tolerating such high defect rates necessitates using reconfigurable crossbar architectures and redundancy [5, 20]. A predetermined design with static crossbars is not capable for defect tolerance because it is not possible to create alternative routes for defective regions. On the contrary, reconfigurable designs can be manipulated for defect tolerance. In this study, we assess and compare defect tolerance performances of the three different reconfigurable nano-crossbar architectures/technologies, represented in Fig. 26.2, that is conducted through finding a valid mapping in accordance with the proposed algorithm and defect maps in case of randomly distributed defects. We consider randomly occurred stuck-open and stuck-closed crosspoint defects causing permanently open and closed defective crosspoint devices or switches, respectively.

This study is at the technology-independent level. The presented synthesis and optimization methods are applicable to variety of nanoarray based emerging technologies including nanowire and nanotube crossbar arrays [13, 15–18], magnetic switch-based structures [8], arrays of single-electron transistors [21], and memristive arrays [22].

26.2.1 Implementing Boolean Logic Functions

We investigate three major implementation methodologies developed for switching nanoarrays. We classify them as two-terminal or four-terminal switch based.

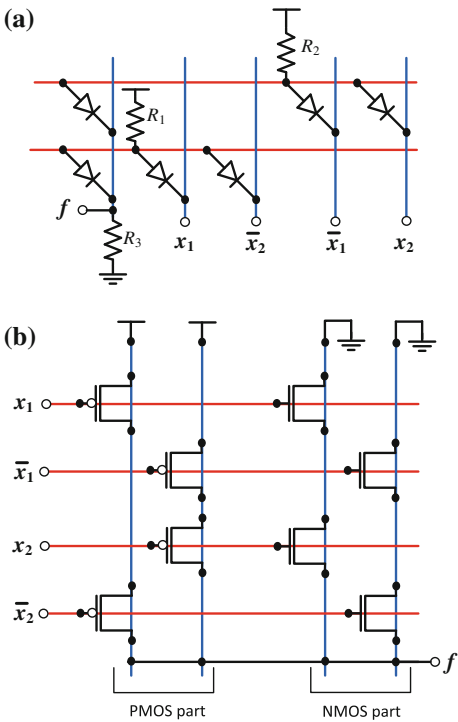
26.2.1.1 Two-Terminal Switch Based Methodologies

These methodologies consider each crosspoint of an array as a two-terminal switch that behaves like a diode or a FET. Since diodes and FETs conduct current through their two terminals that are anode and cathode for diodes and source and drain for FETs, they are fundamentally two-terminal switches.

Boolean functions are implemented by using conventional techniques from diode-resistor logic and CMOS logic with an important constraint regarding nanoarray structures. Boolean functions should be implemented in their sum-of-products (SOP) forms; other forms such as factored or BDD (Binary Decision Diagram) cannot be used since these forms require manipulation/wiring of switches that is not applicable for self-assembled nanoarrays. Figure 26.3 shows implementations of a Boolean function XOR2 with diode and with FET based nanoarrays.

Array size formulations: Given a target Boolean function f , we derive formulas of the array sizes required to implement f . This is shown in Table 26.2. For diode based implementations, each product of f requires a row (horizontal line), and each literal of f requires a column (vertical line) in an array. Additionally, one extra column is needed to obtain the output. For FET based implementations, each product of f and its dual, f^D , requires a column, and each literal of f requires a row in

Fig. 26.3 **a** Diode and **b** FET based nanoarrays implementing XOR2 = $x_1 \oplus x_2$ with 2×5 and 4×4 arrays, respectively



an array. As an example shown in Fig. 26.3: $f = \text{XOR2} = x_1\overline{x_2} + \overline{x_1}x_2$ has 4 literals and 2 products; $f^D = x_1x_2 + \overline{x_1}\overline{x_2}$ has 2 products. This results in array sizes of 2×5 and 4×4 for diode and FET based implementations, respectively. Note that both formulas, for diode and FET, always result in optimal array sizes; no further reduction is possible.

26.2.1.2 Four-Terminal Switch Based Methodology

This methodology considers each crosspoint of an array as a four-terminal switch. An example is shown in Fig. 26.4. The four terminals of the switch are all either mutually connected (ON) or disconnected (OFF). Boolean functions are implemented with top-to-bottom paths in an array by taking the sum (OR) of the product (AND) of literals along each path. This makes Boolean functions implemented in their sum-of-products (SOP) forms. Figure 26.5a, b show the implementations of a Boolean function XOR2 in an array and lattice representations, respectively. Figure 26.5c shows a lattice of four-terminal switches implementing a Boolean function $x_1x_2x_3 + x_1x_2x_5x_6 + x_2x_3x_4x_5 + x_4x_5x_6$. The function is computed by taking the sum of the products of the literals along each path. These products are $x_1x_2x_3$, $x_1x_2x_5x_6$, $x_2x_3x_4x_5$, and $x_4x_5x_6$.

Array size formulation: Given a target Boolean function f , the array size formula was proposed by Altun and Riedel [23] that is shown in Table 26.3. In their implementation, each product of f and its dual, f^D , require a column and a row, respectively, in an array. As an example shown in Fig. 26.5a, $f = \text{XOR2} = x_1\overline{x_2} + \overline{x_1}x_2$ and $f^D = x_1x_2 + \overline{x_1}\overline{x_2}$ have both 2 products. This results in an array size of 2×2 .

Examining the array size formulas in Tables 26.2 and 26.3, we see that while the formulas in Table 26.2 always result in optimal sizes, the sizes derived from the formula in Table 26.3, that is, for four-terminal switch based arrays, are not

Table 26.2 Array size formulas for diode and FET based implementations

Type	Array size formulas (Optimal)
Diode	(Number of products in f) x (“number of literals in f ” + 1)
FET	(Number of literals in f) x (“number of products in f ” + “number of products in f^D ”)

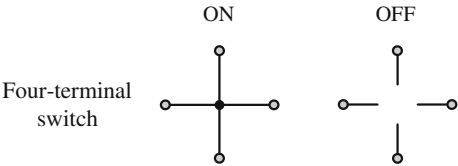


Fig. 26.4 A four-terminal switch with its two states

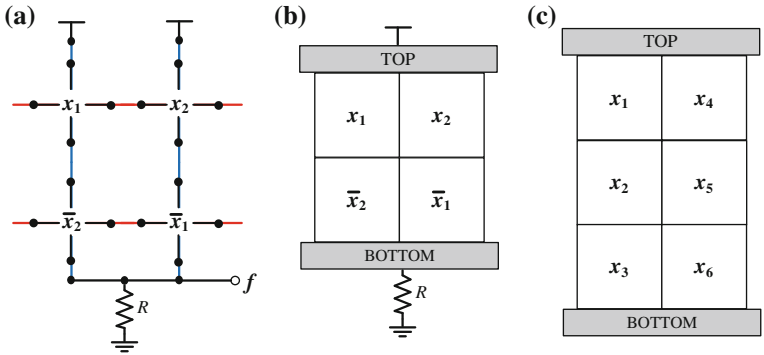


Fig. 26.5 **a** Four-terminal switch based nanoarray and **b** its lattice representation implementing $\text{XOR2} = x_1 \oplus x_2$ with a size of 2×2 . **c** Four-terminal switch based lattice implementing $x_1x_2x_3 + x_1x_2x_5x_6 + x_2x_3x_4x_5 + x_4x_5x_6$

Table 26.3 Array size formula for four-terminal switch based implementation

Type	Array size formula (Non-optimal)
Four-terminal	(Number of products in f) x (number of products in f^D)

necessarily optimal. In the following part we present an algorithm that finds an optimal size implementation of any given target Boolean function.

Finding whether a certain array with assigned literals to its switches implements a target function is the main problem in finding optimal sizes. This problem requires to check if each assignment of 0's and 1's to the switches, corresponding to a row of the target function's truth table, results in logic 1 (a top-to-bottom path of 1's exists). To check this we have to enumerate all top-to-bottom paths; the size of this task grows exponentially with the array size. This is a general statement that holds also for our algorithm described below.

Our algorithm finds optimal array sizes to implement given target Boolean functions with arrays of four-terminal switches in four steps:

- (1) Obtain irredundant sum-of-products (ISOP) expressions of a given-target function f_T and its dual f_T^D . Determine the upper bound on the array size using the formula in Table 26.3.
Upper Bound (UB) = (number of products in f_T) $\times$ (number of products in f_T^D). The implementable **lower bound (LB)** values are taken from the lower bound table proposed in [23].
- (2) List the array shapes (RxC) (which are in between **LB** and **UB**) into the 'List of Implementable Nanoarray Shapes' and sort them regarding of array sizes, in ascending order. While ordering, first take the array shape which has lower number of rows (e.g. if the kth shape is " 3×4 ", then the (k + 1)th shape can be " 4×3 "). Suppose that there are total of **N** different shapes in the list. For step 3, start with n = 1 ($1 \leq n \leq N$).

- (3) Compute the value of the following statement for the n th shape.
 The Statement: An array which has the shape in the n th line of the list is implementable for f_T .
 If the statement is **TRUE**
 Change **UB** to the **R** $\times$ **C** (save the design);
 Go to the step 4;
 If the statement is **FALSE**
 Increase the number “ n ” by 1 ($n = n+1$);
 Repeat step 3.
- (4) Declare that **UB** is **optimal size** for given-target function f_T can be realized in.

Our algorithm is mainly based on finding a design in a certain sized array such that the design implements f_T . Our algorithm does not check every possible design. If it did then it would be intractable even for small sized arrays. For example, if a target function f_T having 6 variables, 8 literals, is tested on a 3×4 array then there are 12^{10} possible designs and 2^6 truth table rows. Note that for each of the 12 switches in the array there are 10 different options; it might be one of the 8 literals, 0, or 1. In this scenario, the algorithm would have to check $12^{10} \times 2^6$ truth table rows. To overcome this problem, we discard a significant portion of designs to be checked. For this purpose, we offer 3 major improvements:

- I. We create a library of reduced number of $R \times 2$ sized sub-designs. We use them to achieve $R \times C$ sized designs. While creating sub-designs we exploit the following simple lemmas. First lemma allows us to discard designs implementing a product (s) that does not imply f_T . The second lemma allows us to discard designs with “0” assignments to the switches if f_T has a product having a single literal.

Lemma 1 *If a design has a path realizing a product p for which $f_T \neq f_T + p$, then the design can not implement f_T .*

Proof Since p is not an implicant of f_T , then a design including p implements a different function.

Lemma 2 *If a function f_T has a single variable product term $p = x$ then the algorithm does not need to assign “0” to the switches.*

Proof All the “0” assignments can be replaced with x ’s without a loss of generality.

- II. If there is a product of f_T such that the number of literals of the product equals to the number of switches in the longest top-to-bottom path in the array, then we settle that particular product onto that particular path.
- III. We discard designs having fewer number of total literals than the total number literals of f_T .

These improvements make our algorithm much faster. As an example, suppose that XOR3 is given as a target function for which the improved algorithm runs roughly 400 times faster. For 3×2 sized sub-designs, there are $8^6 = 262,144$ designs.

With applying the proposed improvements, this number is reduced to 12,114, roughly 20 times smaller than the unimproved one. Since we use two sub-arrays for XOR3, for the optimal array size of 3×4 , the improved algorithm works 400 times faster.

26.2.2 Defect Tolerance

In this section, defect tolerance performance of switching nano-crossbar arrays is extensively studied. Three types of nanoarrays where each crosspoint behaves as a diode, FET, and four-terminal switch, are considered. For each crosspoint, both stuck-open and stuck-closed defect probabilities are independently taken into consideration. A fast heuristic algorithm using indexing and mapping techniques is presented [20]. The algorithm measures defect tolerance performances of the crossbar arrays that are expected to implement a certain given function. The algorithm's effectiveness is demonstrated on standard benchmark circuits that shows 99 % accuracy compared with an exhaustive search. The benchmark results, presented in the next section, also show that not only the used technology, the nanoarray type, but more significantly the specifics of given functions affect defect tolerance performances.

Mapping a target logic function on a defective crossbar is an NP-complete problem [24]. In the worst-case scenario, an $N \times M$ crossbar has $N!M!$ permutations that is intractable for a reasonable computing time. Different algorithms and heuristics are presented to tackle this issue. Graph based models are proposed in [25, 26] that use a fan-out embedding heuristic and a maximum flow algorithm, respectively. In addition to graph based approaches, "Integer Linear Programming" is used in [27] that employs a pruning centered approach with certain constraints. It is shown in [27] that defect tolerance results might dramatically depend on the chosen algorithm correlated to the algorithm's accuracy. We test and compare our algorithm with an exhaustive search to establish an accuracy of 99 %. However, it should be noted that large crossbars are computationally intractable to be included in exhaustive search so we only consider crossbars up to 7×7 size for comparison. In addition, the approaches mentioned so far are using pre-determined crossbar sizes to find a mapping for a chosen logic function. We use an optimal crossbar to realize a logic function which means that both the function and the crossbar matrices have the same size. We test logic functions in Irredundant Sum-of-Products (ISOP) form that is consistent with using optimal crossbar sizes. Also we include three different logic families for comparison which departs from the mentioned studies in the literature.

We present a heuristic algorithm that creates and compares index representations of a given function and a defective certain sized crossbar to be used to implement the function. We show that if the index representations are not matched then the defective crossbar cannot be used to implement the function; otherwise, it can be used. We prove that this is a necessary and sufficient condition. Our algorithm

eliminates considerable amount of crossbar mapping permutations that is the main headache for the mentioned studies in the literature. Furthermore, our viewpoint is comprehensive, that is, concerned both with the types of the crossbar technologies and the characteristics of given functions. The presented indexing based algorithm is directly used for diode and CMOS based logic of nano-crossbars. For four-terminal switch based logic, the presented algorithm is partially used; a conventional matrix-based matching is mainly performed. Also we include three different logic families for comparison which departs from stated studies. It is important to determine features of different families due to post-production selection according to inherent defect types.

As follows, we first explain the algorithm used for diode and CMOS based logic of nano-crossbars [20]. Then, we briefly explain the defect tolerance technique used for four-terminal switch based logic. In the next section, we present experimental results and elaborate on them.

26.2.2.1 The Algorithm for Diode and CMOS Based Logic

The outline of our four-step algorithm is shown below. We will then explain each step in details. The algorithm will be demonstrated with an example in Fig. 26.6. It should be noted that the example and the following explanations are for stuck-open defects, nevertheless they can be applied easily for stuck-closed defects by considering defects as 1s (as opposed to 0s) to be matched with 1s (as opposed to 0s) in the function matrix.

Input: Function matrix and crossbar (defective) matrix
Output: If there is a matching, “YES”; otherwise “NO”

- Step 1: If the number of defective switches is greater than the corresponding elements in the function matrix, then return “NO”.
- Step 2: Sort matrices according to the row and column index, if the crossbar matrix has at least one row or column index greater than a row or column index of the function matrix, return “NO”.
- Step 3: If the number of defects is equal to or smaller than the worst-case limit W_C , return “YES”.
- Step 4: Find the reduced matrix and find set of double indices. Start subarray search. If a subarray is found with the equal set of double indices as the reduced matrix with 20,000 trials, return “YES”; otherwise return “NO”.

The explanations of the algorithm steps for stuck-open defects:

Step 1: *If the number of defective switches is greater than the corresponding elements in the function matrix, then return “NO”.*

We consider stuck-open defects so algorithm checks 0s in the crossbar matrix to match 0s in the function matrix. If 0s in the crossbar matrix is greater than 0s in the function matrix, then it is not possible to find mapping.

Step 2: *Sort matrices according to the row and column index, if the crossbar matrix has at least one row or column index greater than a row or column index of the function matrix, return “NO”.*

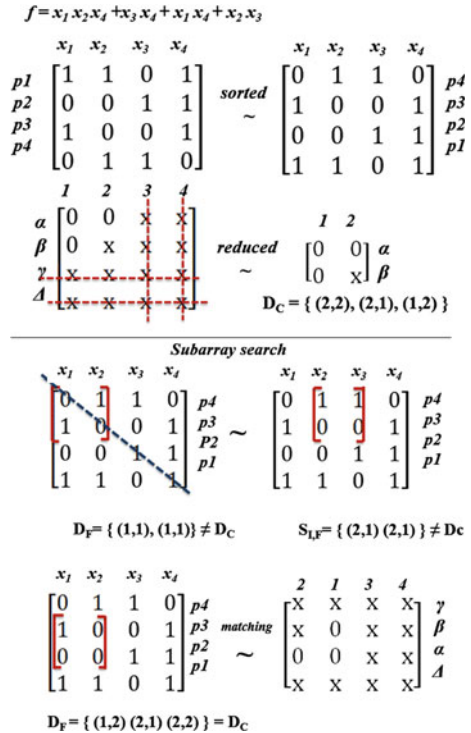
We sort matrices according to the row and column indices. For example, in Fig. 26.6 index sets of the sorted function and crossbar matrices for 0 are:

$$I_{R,F} = \{2, 2, 2, 1\} \quad I_{C,F} = \{2, 2, 2, 1\}$$

$$I_{R,C} = \{2, 1, 0, 0\} \quad I_{C,C} = \{2, 1, 0, 0\}.$$

There is a perfect matching between sets. However, if a member of a crossbar set would be greater than corresponding member in a function set, there would be no matching. This would mean there are excessive

Fig. 26.6 Fourth step of the presented algorithm: sorting the function matrix, crossbar reducing, and subarray search



defective elements for matching. For the example in Fig. 26.6, this is not an issue so we proceed to the step 3.

- Step 3: *If the number of defects is equal to or smaller than the worst-case limit W_C , return “YES”.*

The worst-case limit of a function matrix is the maximum number of tolerable defects in any defect distribution related to the row and column index. We find W_C with using sets of row and column indices. First we choose minimum members in sets of row and column indices, separately. After that we choose the minimum between two members obtained in the first step. W_C gives us minimum row and column index. If crossbar matrix has defective elements less than or equal to W_C , then defects can be matched with any row and column in a function matrix. For illustration, let us now follow the process of finding the W_C of the function matrix shown in Fig. 26.6.

$$I_{R,F} = \{2, 2, 2, 1\} I_{C,F} = \{2, 2, 2, 1\}$$

$$\min\{I_{R,F}\} = 1 \text{ and } \min\{I_{C,F}\} = 1 \text{ so } W_C = 1$$

In Fig. 26.7, the crossbar matrix has three defective elements, so we cannot conclude if there is a matching without performing the fourth step.

- Step 4: *Find the reduced matrix and find set of double indices. Start subarray search. If a subarray is found with the equal set of double indices as the reduced matrix with 20,000 trials, return “YES”; otherwise return “NO”.*

In a crossbar matrix, Xs corresponding to functional switches do not change the double index of a matrix element. For this reason, we erase the columns and rows consisting of only such elements (Xs) for compactness. The acquired matrix keeps the same set of double indices. Figure 26.7 shows an example for this.

In the next step, we use a subarray search to find a matching between a reduced matrix and a subarray. Then, the function and crossbar matrices are sorted according to the sets of indices. We use this to increase the chance of finding matrices with the same set of double indices. Since matching elements are collected to the one side, the search progresses diagonally. It can be seen from Fig. 26.6 that searching a subarray checks only the set of double indices of a chosen subarray, due to the Double Index Theorem, presented below. As long as they have the same set of double indices, permutation of matrices is not necessary. Once two matrices are found with the same set, it means there is a mapping, so the algorithm returns “YES”.

$$\begin{array}{c} \begin{array}{c} 1 \quad 2 \quad 3 \quad 4 \\ \alpha \begin{bmatrix} 0 & x & 0 & x \\ \beta \begin{bmatrix} x & x & x & x \\ \gamma \begin{bmatrix} 0 & x & x & 0 \end{bmatrix} \end{array} \end{array} \quad \begin{array}{c} 1 \quad 2 \quad 3 \quad 4 \\ \alpha \begin{bmatrix} 0 & x & 0 & x \\ \beta \begin{bmatrix} x & x & x & x \\ \gamma \begin{bmatrix} 0 & x & x & 0 \end{bmatrix} \end{array} \end{array} \end{array}$$

Fig. 26.7 Reduced matrix for subarray search

Double Index Theorem *There is a one-to-one matching between two matrices if and only if their set of double indices are equivalent.*

Lemma 1 *Row and column permutations do not alter the double index of a matrix element.*

Numbers of the double index are defined with the row and column indices in which the element is found. Therefore even though permutations changes the position of a row or a column, element is still in the same row and column with the same row and column indices which defines the double index.

Lemma 2 *Set of double indices is unique for a given matrix.*

Proof The proof is by contradiction. Let's assume such a matrix that has two different set of double indices. Therefore sets should have different double indices for one element or more. Double indices in a matrix are determined according to a row and column index. Since matrix is not altered and has same number of elements in rows and columns for a chosen value of 0 or 1, it is not possible to have different row and column indices that comprises double indices. This contradicts with the assumption of having different double index for an element or more.

Lemma 3 *Set of double indices for a matrix does not change with row and column permutations.*

Proof Rows and columns of a matrix is consisted of matrix elements which have the same double index after permutations according to Lemma 1. Therefore set stays the same since its members are not changed.

Proof of the Double Index Theorem Sufficiency If there is a one-to-one matching between two matrices, then they are identical by definition. Therefore their set of double indices is equivalent according to Lemma 2.

Necessity Lemma 2 states that set of indices is unique for a certain matrix and Lemma 3 states that set of double indices stays the same after row and column permutation. Therefore, if two matrices have the same set of double indices, then they are either identical or permutation of one another; in both case they can be matched one-to-one.

Subarray search has a trial limit of 20,000. We choose this value because when compared with exhaustive search, it gives 99 % accuracy. When we run the algorithm by removing this limitation, we see that there is almost no change in the values. This validates the heuristic algorithm presented above interpreted as a negative result meaning that there is no mapping at all. In subarray search, the double index theorem is only valid for matrices with the same number of elements to be matched. In case of unequal number of elements, new defective elements should be introduced to the reduced matrix to equalize the number of elements. If there are N missing elements for matching, 2^N possibilities are to be considered. Instead of doing this, functional switches denoted with x are marked with 0 and defective switches with 1. Same is applied to the subarray in the function matrix. Next, element by element multiplication of matrices is executed. Since functional

switches can be matched with 0s and 1s in the function matrix, the resulting matrix can be compared with the reduced matrix. If they are equal, there is a matching.

In the CMOS based design two matrices are used to model a logic function. All principles used for diode based design are valid, with the exception of input permutations. In the diode based design both inputs (columns) and products (rows) can be permuted with respect to the crossbar which are checked for a mapping. In the CMOS based design, the first matrix is for the function itself and the second matrix for its complement. Important distinction is that inputs are in the same order for both matrices. For this reason while a mapping is searched for a crossbar, rows of matrices can be permuted independently; however inputs must be in the same order for both matrices.

26.2.2.2 Defect Tolerance for Four-Terminal Switch Based Logic

For four-terminal switch based design, we partially use the above presented algorithm. We use matrix based defect maps similar to those used there. Our defect tolerance technique mainly depends on permutation trials. Therefore, we mention it only briefly.

Defect Map Due to its layout method, the four-terminal switch based design [17] uses every switch on the crossbar. Therefore there is no unused switch like those in diode and CMOS based designs. However, certain functions yield redundant paths or extra connections between top and bottom plates. Figure 26.8 shows occurrence of extra connections. If a defect existing in a crossbar appears only on one of the connections, then it can be compensated with the other connection and the correct result can still be achieved. A defect map of a function implemented with four-terminal switch based design displays these type of connections.

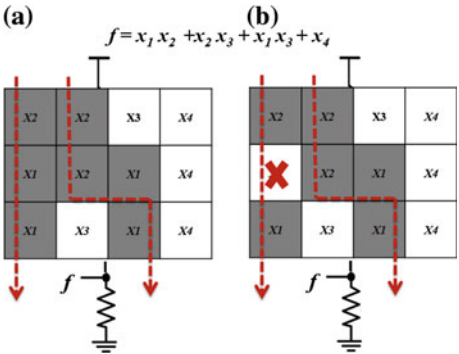


Fig. 26.8 a Four-terminal implementation of f with two connections between *top* and *bottom* plates; b in case of a stuck-open defect, first connection is broken down, however since there is a second connection f evaluates correctly

26.2.3 *Simulation Results*

In Table 26.4, we report synthesis results for standard benchmark circuits [28]. We treat each output of a benchmark circuit as a separate target function. The number of products for each target function f_T and its dual f_T^D are obtained through sum-of-products minimization using the program Espresso [29]. The array size values for “Diode”, “CMOS”, and “4-terminal” are calculated by using the formulas in Tables 26.2 and 26.3. The array size values for “Optimal 4-terminal” are obtained using the presented optimization algorithm in Sect. 26.2.1: Implementing Boolean logic functions.

Examining the numbers in Table 26.4, we always see the same sequence from the worst to the best result as “CMOS”, “Diode”, “4-terminal”, and “Optimal 4-terminal”. This proves that models based on four-terminal switches overwhelm those based on two-terminal switches regarding the array size. Further, the numbers obtained by our optimal synthesis method compares very favorably to the numbers obtained by previous methods.

In Table 26.5, we use standard benchmark circuits to measure defect tolerance performances of different nano-crossbar technologies. We consider stuck-open and stuck-closed defect probabilities/rates of 10 and 20 % for each crosspoint independently. Simulations are conducted in Matlab. Crossbars with random defects are produced with Matlab’s predetermined function generator. To obtain defect tolerance values, a sample size of around 600 is used. At this level the defect tolerance fluctuation stabilizes. All experiments run on a 1.70-GHz Intel Core i5 CPU (only single core used) with 4.00 GB memory. It takes 0.2 s for each sample in average to check a valid mapping that satisfies an accuracy of 99 % compared with an exhaustive search.

Table 26.5 shows the results of benchmark functions with respect to defect rates and defect types as well as the crossbar technologies. Considering the technologies and the related logic synthesis methodologies, the diode based logic always has a better defect tolerance performance compared with the CMOS based one. The reason behind this is directly connected to the number of matchings necessitated for valid mapping. Since the CMOS based logic uses two different planes for function realization, it needs to satisfy two matchings instead of one. Another important conclusion is that the four-terminal switch based design yields better results for stuck-closed defects than for stuck-open ones since the design generally requires to assign the same literals to multiple switches on the same conduction path. Characteristics of the functions also play an important role in the defect tolerance. Since stuck-open defects are tolerated with zeros in the functions’ matrices, functions with relatively higher number of products compared to their number of literals have a better chance for tolerating these defects. On the contrary, functions with relatively higher number of literals compared to their number of products have a better chance for tolerating stuck-closed defects. For example, Ex33 in Table 26.1 has a 14 % defect tolerance for stuck-closed and 99 % for stuck-open type defects.

Table 26.4 Array sizes of three different nano-crossbar based logic families

Benchmark	CMOS	Diode	4-terminal	Optimal 4-terminal
Alu 0	30	18	6	6
Alu 1	30	18	6	6
Alu 2	30	18	6	6
Alu J	30	18	6	6
B12 0	80	32	24	12
B12 1	120	70	35	16
B12 3	30	20	8	8
B12 4	42	28	8	3
B12 6	132	77	35	18
B12 7	110	66	24	18
B12 8	90	70	14	14
C17 0	36	18	9	6
C17 1	30	20	8	8
Clpl 0	64	32	16	12
Clpl 1	36	18	9	9
Clpl 2	16	8	4	4
Clpl 3	144	72	36	18
Clpl 4	100	50	25	15
Del 1	25	10	6	6
Del 2	72	36	16	12
Del 5	35	15	12	6
Del 6	36	18	9	6
Ex5 31	156	104	32	24
Ex5 33	110	77	21	21
Ex5 46	81	54	18	18
Ex5 49	72	54	12	12
Ex5 50	81	63	14	14
Ex5 61	64	48	12	12
Ex5 62	49	35	10	10
Misexl 1	48	16	8	8
Misexl 2	132	55	35	15
Misexl 3	156	60	40	24
MLsexl 4	121	44	28	16
Misexl 5	90	45	25	15
Misexl 6	143	66	42	18
Misexl 7	81	36	20	15
Mp2d 4	345	75	90	24
Newrag	108	72	32	18

Table 26.5 Defect tolerance performances of three different nano-crossbar based logic families

Circuit name	Diode				CMOS				Four-terminal			
	Stuck-closed		Stuck-open		Stuck-closed		Stuck-open		Stuck-closed		Stuck-open	
	10 %	20 %	10 %	20 %	10 %	20 %	10 %	20 %	10 %	20 %	10 %	20 %
Alu 0	85	54	99	96	63	50	93	72	86	64	52	26
Alu 1	85	54	99	96	83	53	97	73	86	64	52	26
Alu 2	85	54	99	96	83	53	97	73	86	64	52	26
Alu 3	85	54	99	96	83	53	94	73	86	64	52	26
B12 0	98	80	98	74	46	5	95	59	28	7	28	7
B12 1	92	33	99	75	58	1	99	91	36	9	58	19
B12 3	96	77	96	78	90	58	93	68	42	17	43	16
B12 4	84	40	99	96	79	25	93	74	42	17	43	16
B12 6	68	2	95	47	14	1	99	85	27	7	21	11
B12 7	44	2	99	95	34	1	82	75	42	11	22	3
B12 8	32	1	99	97	24	1	99	99	82	40	22	4
C17 0	95	78	99	94	92	70	98	87	53	26	53	26
CI7 1	96	77	96	78	91	64	92	69	43	16	43	16
ClplO	97	69	99	98	78	23	99	92	62	29	53	20
Clpl 1	98	84	99	95	98	83	98	82	39	14	48	20
Clpl 2	97	82	99	94	93	79	98	92	67	40	65	42
Clpl 3	87	53	99	81	49	1	50	21	18	3	41	10
Clpl 4	91	41	99	97	74	6	63	50	18	3	41	12
Del 1	99	97	95	75	84	52	93	73	52	25	52	25
Del 2	93	55	99	96	68	9	99	96	28	6	28	6
Del 5	99	95	97	85	96	84	84	53	65	38	53	26
Del 6	95	79	99	88	94	70	98	86	53	25	53	25
Ex5 31	56	5	99	95	30	1	83	64	35	7	22	2
Ex5 33	14	1	99	98	9	1	60	43	66	28	25	4
Ex5 46	45	5	99	99	38	1	84	65	17	1	28	6
Ex5 49	3	0	99	99	1	1	84	65	90	32	28	6
ExS 50	23	1	99	99	22	1	87	45	93	75	22	4
Ex5 61	29	2	99	99	25	1	98	78	90	32	43	16
Ex5 62	28	1	98	85	23	1	96	74	95	74	37	12
Misexl 1	99	96	92	66	65	17	52	9	44	16	43	16
Mlsexl 2	78	18	99	92	30	1	98	87	29	6	36	10
Misexl 3	94	38	99	86	10	1	96	67	8	1	38	11
Misexl 4	93	44	99	94	8	1	99	89	27	5	38	10
Misexl 5	86	45	97	80	63	3	95	64	26	4	42	14
Misexl 6	89	28	99	86	26	1	93	73	12	2	29	5
Misexl 7	95	57	99	92	50	1	99	93	21	3	49	15
Newtag	59	4	99	98	52	1	96	52	62	22	30	7

Significant difference between values occurs because of the mentioned literal related properties. Using Ex33 is more favorable for a crossbar with higher probability of stuck-open type defects.

26.3 Stochastic Computing

Stochastic computing (SC), represents values in time domain by random bit streams [29] and was first presented in a paper written by John von Neumann in 1956 [10]. These values are interpreted as probabilities. Therefore, the available range is $[0,1]$ interval. For instance, if a 16 bit length stream contains 12 1s, it represents $p = 0.75$ independently from the position of 1s in the stream. For example, both $(0,0,0,0,1,1,1,1,1,1,1,1,1,1,1,1)$ and $(0,1,0,1,1,1,1,1,1,0,1,1,1,1,1,0)$ represent the same probability $p = 0.75$. The same probability values can be obtained with different bit stream lengths. Complicated stream operations can be performed with simple circuits in SC. Multiplication and scaled addition are basic arithmetic operations of SC.

Stochastic multiplication can be realized with one AND gate independently of the input bit stream length. In contrary to that, conventional (digital) CMOS multipliers need about 1000 gates for 16 bit multipliers and this number is exponentially increasing with the number of bits. Due to the interval $[0,1]$, addition cannot be performed in SC. Instead, scaled addition that guarantees to remain in the interval is described. MUX is used for scaled addition. Figure 26.9 shows stochastic multiplication and scaled addition operations.

Error rates

Different combination of streams represent same values and this yields error rates in SC. These error rates are extremely high for short bit streams. To reduce error rates, longer bit streams should be used. Figure 26.10 illustrates this; in order have error rates below 1 %, streams having more than 1000 bits are needed. These extremely long streams extend the operation time. Hence, classical SC can not be effectively used in mathematical operations.

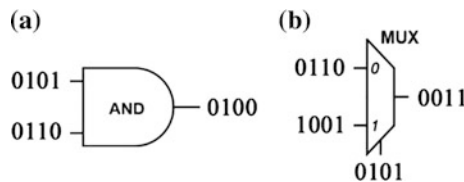


Fig. 26.9 Stochastic implementation of arithmetic operations: **a** Multiplication. **b** Scaled addition

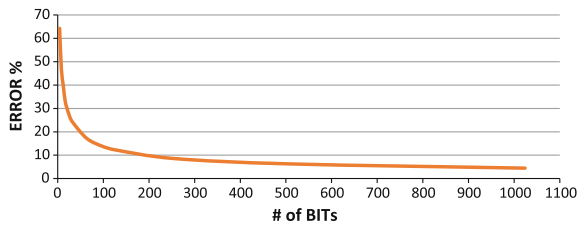


Fig. 26.10 Error rates versus number of bits in the stream for AND gate while $p_1 = 1/2$ and $p_2 = 1/2$

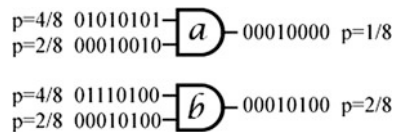


Fig. 26.11 Stochastic multiplication with and without an error

We calculate this error as

$$\frac{|z - z_{expected}|}{z_{expected}} \quad (26.1)$$

where z is the probability value of the output. And expected value varies for different logic gates.

In order to get “1” at the output of an AND gate, both of the inputs have to be “1”. Assume that the probability of being “1” is p_1 for the 1st input and p_2 for the 2nd input. Hence expected value of multiplication/AND gate is $z = p_1 p_2$. Following the same logic, expected value at the output of a MUX is $z = p_1 p_s + p_2 (1 - p_s)$.

Let two input probability values be p_1 and p_2 respectively depending on the number of 1s in streams. The expected value at the output is $p_1 \times p_2$. However, the same probability obtained with different permutations of input may not always yield the expected results. At this point error rates should be considered. This is illustrated in Fig. 26.11. The expected value can be obtained from the AND gate symbolized with a . However, output value with an error rate is obtained from the AND gate symbolized with b in Fig. 26.11.

Effect of dependency in stochastic computing

In conventional SC, bit streams which have Bernoulli distribution are independent from each other. If input streams become dependent, which is not desired, expected result for classical SC will change. If the inverse of the first input is applied as the second input to an AND gate, “0” will be obtained at the output. It is expected that $z = p(1 - p)$, but due to dependency this expectation fails. Moreover, if the same inputs are applied, the output will be same as the input. It acts like a buffer. It is

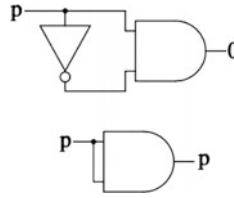


Fig. 26.12 Effect of dependency in SC

expected that $z = p \cdot p$, but due to dependency this expectation fails. These examples are shown in Fig. 26.12.

Application areas

SC needs less area than conventional hardware implementations [4]. On the contrary, error rates are worse. Hence, the commercial application fields are limited (image processing, LDPC codes) [30]. Recently, printed/flexible electronics becomes a major candidate for stochastic computing. Printed electronics strictly requires low density circuits that is a perfect fit with stochastic computing if error rates can be reduced to proper values [31]. This is our motivation. In this study, we offer a method based on improving the generation of bit streams. In Sect. 26.3.1, a detailed information about our method and the conventional one is given. Comparative analysis made between two methods demonstrate our methods success. We also propose a method to achieve error free stochastic computing in Sect. 26.3.2

26.3.1 Reducing Error Rates

We analyze and make comparison between two SC methods, namely random bit assigning and shuffling methods.

26.3.1.1 Random Bit Assigning Method

In order to generate bit streams in SC, conventional random bit generators are used [30]. We name this technique as random bit assigning method (RBAM) [32]. Let desired input probability value for RBAM be p_1 and let random number generator generate values between [0–1] with binomial distribution. If the generated value is smaller than p_1 , “1” is added to the stream. If the generated value is bigger than p_1 “0” is added to the stream. Because this is a stochastic procedure, there exists difference between the desired and the obtained input values. There are three 1s and five 0s in the generated 8 bit stream in Table 26.6. For this reason, generated input bit stream is physically $p_e = 0.375$. Consequently, the result at the output will be faulty.

Table 26.6 8 bit length stream generated with RBAM for $p_1 = 0.5$

Generated number from generator	Bit stream
0.89967	0
0.45847	1
0.81537	0
0.20818	1
0.71289	0
0.43615	1
0.64214	0

26.3.1.2 Random Bit Shuffling Method

Let the input desired value for random bit shuffling method (RBSM) be p_1 . In this method, depending on the desired probability value and the length of the bit stream, a bit stream containing necessary and sufficient amounts of 1 is generated. Fisher-Yates shuffling algorithm [33], adapted to computers by Durstenfeld [34], is applied. In this way, it is guaranteed that the desired and the generated input values are probabilistically identical. With this approach output error rates will be decreased. We use basic logic gates to compare these two methods. The results are obtained by fixing the two inputs, changing the inputs between [0, 1] interval and using different bit lengths.

RBSM–RBAM Comparison for $p_1 = 1/2$, $p_2 = 1/2$ Bit Streams

As shown in Table 26.7, shuffling method gives better results for AND and OR gates. The longer bit length reduces the error rates as expected and the difference of error rates between two methods also is reduced. The expectation is that error rates of two methods overlap at infinity. Two methods give similar results for XOR gate between 4 and 128 bits, however for longer bit streams shuffling method gives better results.

RBSM–RBAM Comparison While 64-Bit Inputs p_1 and p_2 Changing

The graphics in Fig. 26.13 are symmetrical with respect to the diagonal line. Two inputs for the ideal operating range of AND gate should be 0.5 and/or greater. Average error rates for RBAM and RBSM are 30.11 and 19.31 %, respectively. The results are better for RBSM.

The graphics in Fig. 26.14 are symmetrical with respect to the diagonal line. The error rates of an OR gate are better than those for an AND gate. Two inputs for the ideal operating range of an OR gate should be 0.3 and/or greater. Average error rates for RBAM and RBSM are 6.07 and 2.24 %, respectively. The results are better for RBSM.

Table 26.7 Obtained error rates with different bit lengths and three logic gates (%)

Number of bit	RBAM			RBSM		
	AND	OR	XOR	AND	OR	XOR
4	64.3	20.82	37.55	32.55	11.2	33.65
8	46.58	15.38	27.1	26.44	8.77	25.76
12	39.92	12.98	22.52	21.22	7.24	21.85
16	33.38	11.05	19.4	19.04	6.42	19.09
20	30.28	10.22	17.68	17.58	5.85	17.24
24	27.75	9.38	16.35	16.01	5.2	15.66
32	24.28	8.2	14.02	13.82	4.67	13.18
64	17.1	5.68	9.65	8.32	3.05	10.24
100	13.62	4.58	7.88	7.93	2.58	7.82
128	12.12	4.05	7.08	6.21	2.23	7.1
256	8.52	2.8	4.98	3.98	1.6	3.57
512	6.22	2	3.55	2.6	1.34	2.03
1024	4.42	1.4	2.45	1.95	0.35	1.56

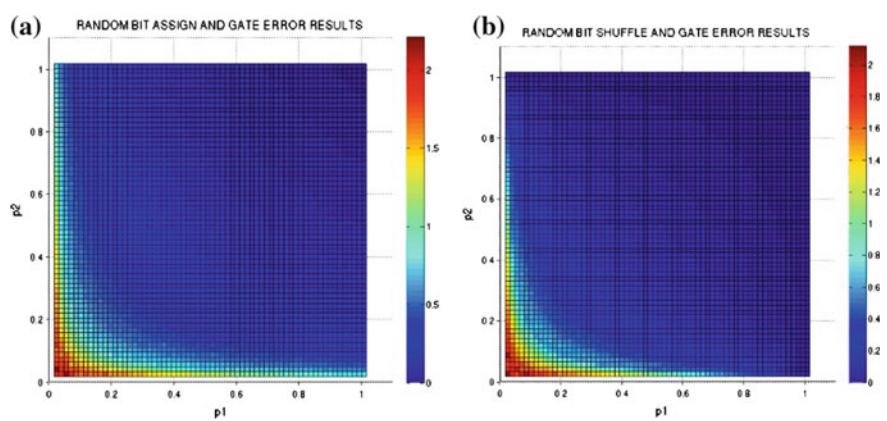


Fig. 26.13 p_1 - p_2 -error graph for **a** RBAM and **b** RBSM AND gate

The graphics in Fig. 26.15 are symmetrical with respect to both diagonal lines. The error rates of an XOR gate are better than those for an AND gate but worse than those for an OR gate. Two inputs for the ideal operating range of XOR gate should be 0.3 and/or greater. However, the two inputs should not take place on the diagonal line at the same time. Average error rates for RBAM and RBSM are 10.94 and 6.58 %, respectively. The results are better for RBSM.

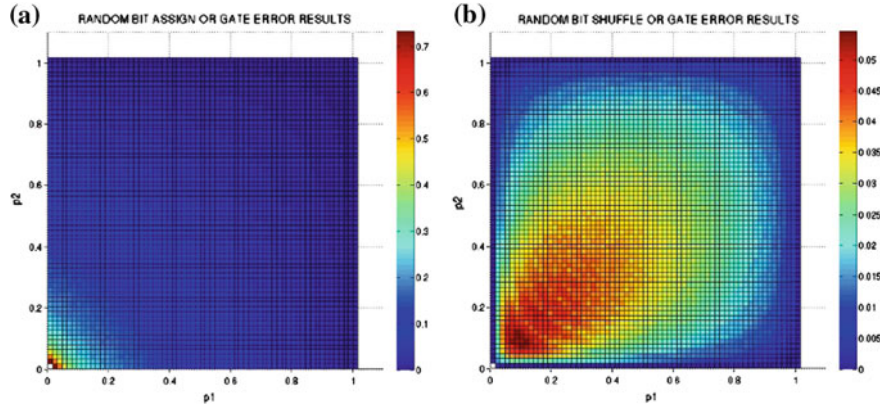


Fig. 26.14 p_1 - p_2 -error graph for **a** RBAM and **b** RBSM OR gate

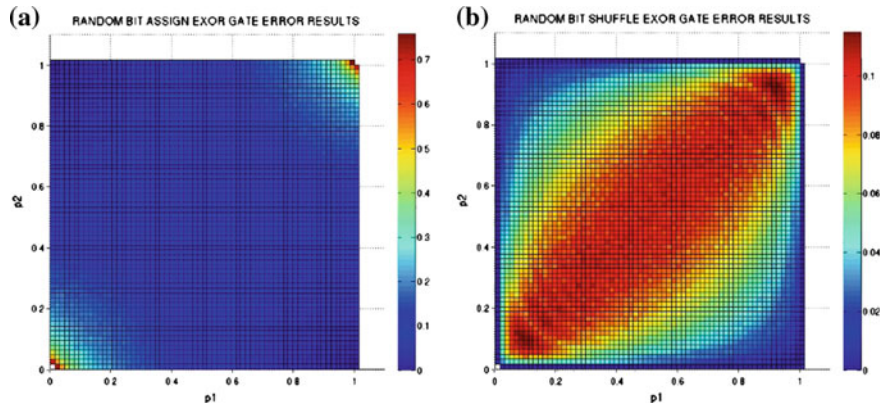


Fig. 26.15 p_1 - p_2 -error graph for **a** RBAM and **b** RBSM XOR gate

26.3.2 Error Free Stochastic Computing

Conventional wisdom in stochastic computing is that input bit streams should be independent; we previously discussed potential accuracy problems related to dependencies. In this work, however, we use dependent inputs to improve accuracy; we achieve error free outputs.

26.3.2.1 Realizing Error Free Multiplication with 0.5

We suggest an error free block for multiplication by 0.5 shown in Fig. 26.16. With this block, the input stream is multiplied by 0.5 and zero error at the output stream

is achieved. However if the length of the input stream is n , the length of the output stream will be $2n$ because of the merging operation. The merge block can be realized with a delay switch. Thus, the stream obtained from the second AND gate waits the stream from the first AND gate to be finished.

26.3.2.2 Generating Any Probability Value Without an Error

We adapt the block in Fig. 26.16 to switch circuits [35] in order to obtain any probability value. It is illustrated in Fig. 26.17. Inputs are always 0.5 ($0.12 = 1/2$). In order to shift the binary value one digit to the right and add “0” to the most significant bit, the circuit block based on AND gate is used ($0.012 = 1/4$). In order

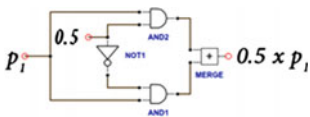


Fig. 26.16 Error free multiplication for $0.5 p_i$

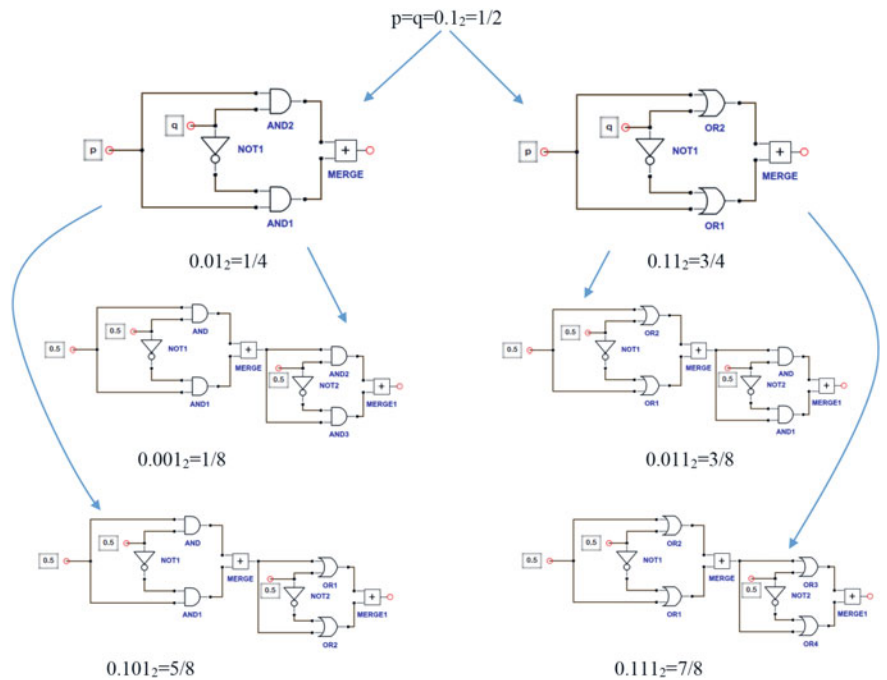


Fig. 26.17 Probability generation tree

to shift the binary value one digit to right and add “1” to the most significant bit, the circuit block based on OR gate is used ($0.112 = 3/4$). By cascading the blocks, any desired probability value can be accurately obtained without an error.

26.4 Conclusions

This chapter overviews both deterministic and stochastic computing models targeting nano-crossbar switching arrays and emerging low-density circuits. These models are demonstrated with implementations using Boolean and arithmetic logic. Performance parameters of the models such as area, reliability, and accuracy, are also evaluated.

Acknowledgments This work is supported by TUBITAK (The Scientific and Technological Council of Turkey) Career Program #113E760 and EU (European Union) H2020-MCSA Research and Innovation Staff Exchange Program #691178.

References

1. L. Wilson, International technology roadmap for semiconductors (ITRS). Semiconductor Industry Association (2013)
2. M. Dubash, Moore’s Law is dead, says Gordon Moore. Techworld (April 2005)
3. M. Haselman, S. Hauck, The future of integrated circuits: a survey of nanoelectronics. *Proc. IEEE* **98**(1), 11–38 (2010)
4. W. Qian, X. Li, M.D. Riedel, K. Bazargan, D.J. Lilja, An architecture for fault-tolerant computation with stochastic logic. *IEEE Trans. Comput.* **60**(1), 93–105 (2011)
5. D.B. Strukov, K.K. Likharev, Reconfigurable nano-crossbar architectures. *Nanoelectronics Inf. Technol.* 543–562 (2012)
6. G.M. Whitesides, B. Grzybowski, Self-assembly at all scales. *Science* **295**(5564), 2418–2421 (2002)
7. M. Altun, M.D. Riedel, Robust computation through percolation: synthesizing logic with percolation in nanoscale lattices. *Int. J. Nanotechnol. Mol. Comput. (IJNMC)* **3**(2), 12–30 (2011)
8. A. Khitun, M. Bao, K.L. Wang, Spin wave magnetic nanofabric: a new approach to spin-based logic circuitry. *IEEE Trans. Magn.* **44**(9), 2141–2152 (2008)
9. Z. Abid, M. Liu, W. Wang, 3D integration of CMOL structures for FPGA applications. *IEEE Trans. Comput.* **60**(4), 463–471 (2011)
10. J. Von Neumann, Probabilistic logics and the synthesis of reliable organisms from unreliable components. *Automata Stud.* **34**, 43–98 (1956)
11. S. Gaba, P. Knag, Z. Zhang, W. Lu, Circuits and Systems (ISCAS), in *2014 IEEE International Symposium on Memristive devices for stochastic computing*. IEEE (2014, June), pp. 2592–2595
12. A.C. Arias, J.D. MacKenzie, I. McCulloch, J. Rivnay, A. Salleo, Materials and applications for large area electronics: solution-based approaches. *Chem. Rev.* **110**(1), 3–24 (2010)
13. Y. Huang, X. Duan, Y. Cui, L.J. Lauhon, K.H. Kim, C.M. Lieber, Logic gates and computation from assembled nanowire building blocks. *Science* **294**(5545), 1313–1317 (2001)

14. W. Lu, C.M. Lieber, Nanoelectronics from the bottom up. *Nat. Mater.* **6**(11), 841–850 (2007)
15. P. Avouris, Molecular electronics with carbon nanotubes. *Acc. Chem. Res.* **35**(12), 1026–1034 (2002)
16. G. Snider, U.S. Patent No. 6,919,740. Washington, DC: U.S. Patent and Trademark Office (2005)
17. Z. Chen, et al., An integrated logic circuit assembled on a single carbon nanotube. *Science* 311.5768, 1735–1735 (2006)
18. H. Yan, H.S. Choe, S. Nam, Y. Hu, S. Das, J.F. Klemic, C.M. Lieber, Programmable nanowire circuits for nanoprocessors. *Nature* **470**(7333), 240–244 (2011)
19. C.P. Collier, E.W. Wong, M. Belohradský, F.M. Raymo, J.F. Stoddart, P.J. Kuekes, J.R. Heath, Electronically configurable molecular-based logic gates. *Science* **285**(5426), 391–394 (1999)
20. O. Tunali, M. Altun, Defect tolerance in diode, FET, and four-terminal switch based nano-crossbar arrays, in *IEEE International Symposium on Nanoscale Architectures, 2015. NANOARCH 2015*. pp. 82–87. IEEE (2015, July)
21. Y.C. Chen, S. Eachempati, C.Y. Wang, S. Datta, Y. Xie, V. Narayanan, Automated mapping for reconfigurable single-electron transistor arrays, in *Proceedings of the 48th Design Automation Conference*, pp. 878–883. ACM (2011, June)
22. Y. Levy, J. Bruck, Y. Cassuto, E.G. Friedman, A. Kolodny, E. Yaakobi, S. Kvatinsky, Logic operations in memory using a memristive Akers array. *Microelectron. J.* **45**(11), 1429–1437 (2014)
23. M. Altun, M.D. Riedel, Logic synthesis for switching lattices. *IEEE Trans. Comput.* **61**(11), 1588–1600 (2012)
24. A.M.S. Shrestha, S. Tayu, S. Ueno, Orthogonal ray graphs and nano-PLA design. In *ISCAS* (2009, May), pp. 2930–2933
25. W. Rao, A. Orailoglu, R. Karri, Topology aware mapping of logic functions onto nanowire-based crossbar architectures, in *Proceedings of the 43rd Annual Design Automation Conference*. ACM (2006, July), pp. 723–726
26. J. Huang, M.B. Tahoori, F. Lombardi, On the defect tolerance of nano-scale two-dimensional crossbars, in *Proceedings of 19th IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems*, 2004. DFT 2004. IEEE (2004, October), pp. 96–104
27. M. Zamani, H. Mirzaei, M.B. Tahoori, ILP formulations for variation/defect-tolerant logic mapping on crossbar nano-architectures. *ACM J. Emerg. Technol. Comput. Syst. (JETC)* **9**(3), 21 (2013)
28. K. McElvain, in *Distributed as part of the MCNC International Workshop on Logic Synthesis IWLS'93 benchmark set: Version 4.0* (vol. 93) (1993, May)
29. B.R. Gaines (1967, April). Stochastic computing. In *Proceedings of the April 18–20, 1967, Spring Joint Computer Conference*. ACM, pp. 149–156
30. A. Alaghi, J.P. Hayes, Survey of stochastic computing. *ACM Trans. Embed. Comput. Syst. (TECS)* **12**(2s), 92 (2013)
31. W.S. Wong, A. Salleo (eds.), *Flexible electronics: materials and applications* (vol. 11). Springer Science & Business Media (2009)
32. S. Yavuz, M. Altun, Stokastik Hesaplama Hata Oranlarını Azaltmak için Rastgele Bit Karıştırma Yöntemi—Random Bit Shuffling Method for Reducing Error Rates in Stochastic Computing, in *Elektrik—Elektronik—Bilgisayar ve Biyomedikal Mühendisliği Sempozyumu*, pp. 728–732. ELECO (2014)
33. R.A. Fisher, F. Yates, *Statistical tables for biological, agricultural and medical research*, 3rd edn. (1949)
34. R. Durstenfeld, Algorithm 235: random permutation. *Commun. ACM* **7**(7), 420 (1964)
35. H. Zhou, P.L. Loh, J. Bruck, The synthesis and analysis of stochastic switching circuits. (2012). arXiv preprint [arXiv:1209.0715](https://arxiv.org/abs/1209.0715)
36. C. Morgul, M. Altun, Synthesis and optimization of switching nanoarrays, in *2015 IEEE 18th International Symposium on Design and Diagnostics of Electronic Circuits & Systems (DDECS)*. pp. 161–164. IEEE (2015, April)

Index

A

- Ab initio, 201
- Ab initio modeling, 181
- Ablation, 174
- Ablation threshold, 174
- Absorbing light, 248
- Absorption (ABS), 332
- Absorption coefficient, 364
- Absorption edge, 341
- Absorption properties, 443
- Absorption spectrum, 378
- Acceptor, 189, 253, 566
- Accuracy, 647
- ACN, 470
- Acoustic plasmon, 205, 225
- Acoustic plasmon (lower), 217
- Acoustic surface plasmon (ASP), 222, 223
- Activation states, 143
- Active nanostructures, 108
- Active pharmaceutical ingredients, 110
- Ad-atom, 184–190, 199, 201, 275
- Adsorption, 183–185, 187–189, 478, 481, 483, 492, 500
- Adsorption ability, 428
- Adsorption capacity, 495
- Advantages, 424, 433, 436
- AFM, 113
- Ag modification, 448
- AgGa_{0.5}In_{0.5}Se₂, 372
- Agglomeration, 439
- Aggregates, 438
- AlGaAs/GaAs heterostructure, 19
- AlGaIn/GaN, 24, 43
- AlGaIn, 161
- Algorithm, 641
- Alkalinity, 487
- Ammonium, 477, 478, 487, 492, 495, 498
- Amorphous, 292, 294
- Amorphous phase, 295
- Amorphous silicon dioxide, 115
- Amphiphilic molecules, 510
- Anatase, 426, 430, 431, 436, 437
- Anatase and rutile, 434
- Anatase/brookite, 432, 441, 445, 446
- Anatase/rutile, 433, 434, 436, 439, 442, 447
- Anatase TiO₂, 384
- Angle of incidence, 103
- Angle resolved X-ray photoemission spectroscopy, 249
- Anharmonic contribution, 153
- Anisotropic growth, 361
- Antenna mechanism, 453, 454
- Antibody conjugate, 520
- Antibody conjugated QDs, 215, 569
- AO, 262
- Aperture, 73, 74, 76, 78
- Applications, 246, 441
- Applications of QDs, 506
 - bionanotechnology and nanomedicine, 565
- Applied potential, 208
- Aptamer, 519
- Aquaporins, 546
- Aqueous synthesis, 558
- AR-XPS, 249
- Arc discharge, 244
- Armored protection systems, 547
- ARPES, 312, 322
- Array size formulas, 640
- Arrays of single-electron transistors, 638
- Artificial preparation, 427
- Artificial retinas, 113
- Asialoglycoprotein receptor, 582
- Asymptotic absorption edge, 336
- ATR prism, 142
- Attenuated total reflection, 133
- Au modified, 442
- Autofluorescent, 557
- Average error rates, 656

- Axial, 353
- Axicons, 176
- Axis of central maximum, 103
- B**
- Bacterial biofilm, 599
- π band, 227
- d band, 228
- Band bending, 311
- Band of frequencies, 64
- Band offsets, 40
- Band structure, 453
- Bandgap, 205, 221, 229, 364, 427, 455
- Bands (conduction, valence), 206
- Bargmann coherent state, 611
- Barrier layer, 5
- Batch, 495
- Beer-Lambert law, 151
- Benchmark circuits, 649
- Benchmark functions, 649
- Bernoulli distribution, 653
- Bessel beam, 175, 176
- Bi-doped ZnO, 402, 403, 405
- Bi-doped ZnO films, 402
- Bilayer, 229
- Bilayer graphene, 217, 219, 229
- Binding, 269
- Biocompatibility, 543, 557, 559
 - in QDs, 562
- Biocompatible, 536
- Biocompatible silica, 560
- Bioconjugation, 514, 562
- Bioconjunction, 514
- Biodistribution, 520
- Bioimaging, 557
- Biointerfaces, 541
- Biological, 500
- Biological labelling, 514
- Biomacromolecules, 512
- Biomarkers, 583
- Biomaterials, 543
- Biomedical applications of QDs, 573
- Biomeolecular engineering, 147
- Biomimetic, 540
- Biomimicry, 548
- Biomineralization, 536
- Biomolecule, 131
- Bionanotechnology, 565, 589, 591, 602
- Bionsensor, 583
- Biophotonics, 539
- Biosensing, 246
- Biosensors, 539
- Biotechnology, 556
- Biotin–streptavidin recognition, 563
- Biphasic TiO₂, 448
- Bismuth-doped zinc oxide, 401, 402
- Bit lengths, 656
- Bit streams, 652, 655
- Blend, 301
- Bloch functions, 14
- Block copolymers, 510
- Blue shift, 227, 336
- Bode phase, 468
- Boron, 190
- Boron oxide, 182, 188, 189
- Bottom-up approach, 356
- Boundary condition of O operator, 619, 624
- Bound state-one dimensional, 56
- Bovine Rhodopsin, 139
- Breakthrough, 482, 483, 487
- Brookite, 426, 428, 430, 431, 445
- Brookite/anatase, 448
- Bulk dielectric function, 212
- Bulk heterojunction, 380
- Bulk plasma, 213
- Bulks, 292
- C**
- C₄H₆O₆/TiCl₃ molar ratio, 432
- Calcination, 301, 387
- Calcium phosphate, 537
- 3D calibration artefact, 117
- Capacitance, 467
- Capacitive, 466
- Capacitor, 467
- Capacity of storing quantum information, 627
- Carbodiimide chemistry, 518
- Carbohydrates, 513
- Carbon, 181–184, 187–193, 197, 199–202
- Carbon cluster, 188, 190, 199
- Carbon computer, 113
- Carbon dimer, 184, 185
- Carbon dots, 564
- Carbon fibers, 462
- Carbon monoxide, 190–192, 201
- Carbon nanotubes, 117, 239, 380, 387, 388
- Carbon trimer, 184, 188, 201
- Cd-free RGD peptides, 568
- Cd_{0.08}Se_{0.92} quantum dots, 333
- CdCl₂ treatment, 380
- CdS_{0.08}Se_{0.92}, 334
- CdSe nanocrystals, 329
- CdSe QDs, 572
- CdSexS_{1-x} Nanocrystals, 328
- CdSe/ZnS core shell, 507
- CdSO₄ structure, 339
- CdSTe/CdTe, 44
- CdS thin film, 328, 336, 338–340

- CdSxSe_{1-x}, 332, 333
 - CdTe/CdS, 328, 367, 378, 380, 390
 - CdTe thin film, 329–331
 - CdX and ZnX (X = S, Se, Te) binaries, 38
 - 3D cell culture, 598
 - Cell labeling, 565
 - Cell membrane mimics, 544
 - Certification of reference materials, 115
 - CFME, 462
 - Chalcopyrites, 362
 - Challenges, 424, 456
 - Change the band structure, 253
 - I–V characteristic, 369
 - Charge collection, 376
 - Charge neutrality level, 311
 - Charge separation, 453
 - Charge separation mechanism, 452, 454
 - Charge transfer (CT), 252, 571
 - Charging, 263
 - ChemFETs, 254
 - Chemical bath deposition (CBD), 335
 - Chemical vapor deposition, 182, 200, 202, 245
 - Chemisorption, 509
 - Chemotaxis, 599
 - Chiral vector, 240
 - Chromism, 293
 - Chronic diseases, 581
 - Circuit, 472
 - Click chemistry, 513
 - Clinical trials, 583
 - Clinoptilolite, 488, 491, 492, 497
 - Closed noise-free operator, 622
 - Cluster, 172
 - CMOS, 181, 182, 193, 196
 - CNT chemiresistors, 254
 - CO₂ photoreduction, 445
 - Coadsorbed, 284
 - Colored and bleached tungsten oxide films, 305
 - Combined uncertainty, 106
 - Compact QDs, 561
 - Complex Gaussian random variable, 611
 - Composite system, 215
 - Composition dependent bond lengths, 26
 - IV–VI, III–V and II–VI compounds, 10
 - Concentration of SO₄²⁻, 434
 - Conclusion, 456
 - Conducting polymers, 389
 - Conducting screen, 84
 - Conducting substrate, 229
 - Conduction band offset, 42
 - Conductivity, 380
 - Conductor, 220
 - Conformational changes, 140
 - Consistency condition of *O* operator, 614
 - Contaminants, 252
 - Contamination, 182, 196
 - Control, surface electron accumulation
in InN nanowires, 314, 320, 321, 323
 - Copolymer, 520
 - Copper, 182
 - Copper-free click chemistry, 513
 - Core-shell, 353, 379
 - Core-shell solar cell, 353
 - Correlation function, 613
 - Coulomb coupled, 214, 216, 220
 - Coulomb potential, 214
 - Counter electrode, 384, 387, 388
 - Coupled plasmons, 215
 - Coupled plasmon spectrum, 217
 - Covalent attachment, 513
 - Cr-TiO₂, 451
 - Cr³⁺doping, 450
 - CRET/BRET, 567
 - Critical distance, 205, 229
 - Critical wavenumber, 213
 - Cross-correlation, 121
 - Crossbar-type switches, 636
 - 3D culture systems, 598
 - Crystal growth, 360
 - Crystal phases, 425
 - Crystalline, 295
 - Crystallite sizes, 455
 - Crystallization, 428, 487
 - Crystallographic phases, 293
 - CT using QDs
as donors/acceptors, 567
 - CuInSe₂, 367, 390
 - Cumulant expansion, 153
 - Current density-voltage (J–V) characteristics, 250
 - Cut-off frequency, 86, 196
 - Cyclic voltammetric measurements, 389
 - Cyclic voltammograms, 464
- D**
- Damping, 213, 228
 - Data, 472
 - D-band, 243
 - Debye-Waller factor, 152
 - Decoherence-free space, 627
 - Defect calculations, 401, 403, 404, 410, 414, 419
 - Defect distribution, 646
 - Defect formation energy, 405, 406
 - Defect maps, 648
 - Defect probabilities/rates, 643, 649
 - Defect tolerance, 636, 649
 - Defect transition energy, 406, 407, 418

- Degenerate case, 213
- Degradation of organic pollutants, 446
- Delta-dot, 52
- Delta-prime, 49, 50
- Density functional theory (DFT), 1, 183, 184, 192, 193
- Density functional theory modelling, 32
- Density of states, 241
- Density perturbation, 208
- Density perturbation response, 61
- Density perturbation response function, 209
- Density plot, 89, 91–95
- Deoxidized, 280
- Deposition rate, 184, 185
- Desorption, 184–189, 201, 492
- Deterministic, 659
- DFT+U method, 408, 410
- DFT with MBJLDA functional, 37, 46
- Diamondoid, 118
- 2,4-dichlorophenol, 450
- Dielectric, 206
- Diffraction, 73, 103
- Diffraction zones, 88–90
- Diffusion, 183–186, 190, 192, 193, 199, 201, 481
- Dimensional metrology, 109
- Dimensional nanometrology, 118
- 3,4-(2,2-dimethylpropylenedioxy)thiophene, 462
- Diode, 638
- Diode behaviour, 374
- Dirac cones, 228
- Dirac delta function, 49, 74, 77, 207, 210
- Dirac-delta-function-derivative, 64
- Dirac-like energy spectrum, 229
- Dirac point, 227
- Direct attachment, 518
- Directed evolution, 600
- Dispersion, 205
- Dispersion relation, 54
- Distorted bond length, 11
- Divergent integral, 78
- DMF, 470
- DNA microarrays, 594
- DNA sequencing, 602
- Doctor blade technique, 388
- Donor, 189, 253, 567
- Doped, 218, 219
- Doping, 227, 228, 319, 321
- Double indices, 646
- Double layer, 463
- Double quantum dot, 50, 53
- Drawbacks, 433
- Drinking water, 495
- Drug delivery, 516, 581
- Drugs, 580
- D-states, 29
- Dual, 640
- Dual-modality imaging, 571
- Dye sensitized solar cells, 380, 388
- Dynamic, 212
- Dynamics of quantum entanglement, 622
- Dyson integral equation, 50
- E**
- EDS, 337
- Effective potential, 60, 208
- Efficiency, 480
- Eigenenergy, 53
- Eigenenergy spectrum, 59
- Einstein frequency, 153
- EIS, 468
- Elastic stiffness constants, 7
- Electric field, 263
- Electric micro-field, 342, 344
- Electrical, 305
- Electrical conductivity, 248
- Electrical measurements, 318
- Electrochemical, 298, 356, 464
- Electrochromic, 298
- Electrodeposited, 465
- Electrolytes, 461
- Electron-beam evaporation technique, 355
- Electron-hole continuum, 225
- Electron-hole excitations, 225
- Electron acceptors, 428
- Electron accumulation, 311
- Electron beam energy, 224
- Electron density, 319
- Electron energy loss spectroscopy (EELS), 221, 226
- Electron propagation along wire, 58
- Electron transmission, 64
- Electron yield mode, 155
- Electronic properties, 309, 314
- Electrophoresis, 595, 596
- Electropolymerisation, 464
- Electrospinning, 298
- Electrospun, 301
- Electrospun tungsten oxide nanofibers, 299
- Emerging low-density circuits, 659
- Empirical pseudopotential bandgaps, 17
- Empirical pseudopotential bands, 31
- Empirical pseudopotential calculations, 36
- Encapsulation, 512
- Energy, 345
- Energy barrier, 184, 185, 187, 192, 195, 199
- Energy donors, 566

- Energy loss spectrum, 221
 - Energy resources, 424
 - Engineering metrology, 109
 - Enhanced ferromagnetism, 114
 - Enhanced photocatalytic activity, 451
 - Ensemble method, 120
 - Entanglement transfer, 626
 - Environmental problems, 424
 - Epilayer, 8
 - Epitaxial Graphene, 221
 - EPR, 453
 - Equation of motion of O operator, 620
 - Equilibrium, 480
 - Equivalent, 472
 - Error free multiplication, 657
 - Error free stochastic computing, 657
 - Error rates, 653, 655
 - Etchant concentration, 363
 - Evanescence wave, 133
 - Evolution equation of O operator, 624, 632
 - Evolutionary engineering, 601
 - Exact master equation, 616, 628
 - Exchange-correlation functional, 33
 - Excited state, 27
 - Exciton, 384
 - Exciton binding energy, 371
 - Exhaustion, 483
 - Exhaustive search, 647
 - Expanded, 116
 - Experimental, 473
 - Extended XAFS (EXAFS), 152
 - External quantum efficiency, 158
 - Extreme UV nanolithography, 110
 - Extrinsic graphene, 218, 219
- F**
- FDA, 585
 - Feature size in microchips, 110
 - Femtosecond, 173
 - Femtosecond laser, 132
 - Fermi energy, 213, 227
 - Fermi level, 5, 227, 382
 - Fermi level pinning, 311, 314, 323
 - Fictitious excited s^* orbital, 21
 - Field amplitude distributions, 89
 - Field distribution, 85, 86, 88
 - Field effect transistor (FET), 182, 194, 638
 - Field of full sheet, 82
 - Fill factor, 365, 381, 387
 - Filtration, 498
 - First Brillouin zone, 18
 - First-principles, 266
 - First principle ab initio methods, 12
 - Fixed-bed, 482, 495
 - Fluidized bed, 498
 - Fluorescence imaging, 516
 - Fluorescence mode, 154
 - Fluorine-doped thin oxide (FTO), 377
 - Fluorochromes, 557
 - Fly ash, 494
 - Foldable photoelectronics, 114
 - Forbidden condition, 622, 625
 - Formal master equation, 615
 - Formation, 282
 - Four-terminal, 638
 - Free-standing, 206
 - Free-standing graphene, 221, 228, 229
 - Frequency dependence, 96, 97
 - FRET, 523, 566
 - applications, 525, 526
 - donors, 523
 - pair, 523
 - Freundlich, 480
 - FTIR-ATR, 466
 - Fullerenes, 113
 - Functional expansion of O operator, 618
 - Functional materials, 114
 - Functional switches, 646
 - Fundamental constants of nature, 107
 - Fusion tags, 570
- G**
- GaAsN/GaAs, 24, 43, 95–97, 102, 132, 355
 - Gallium arsenide, 206
 - Gap, 206
 - Gapless graphene, 217, 229
 - Gapless graphene polarizability, 217
 - Gapped graphene, 217, 220
 - Gas sensing, 250
 - Gate potential, 227
 - Gauge dependence, 57
 - GAUSSIAN, 136
 - Gaussian, 174
 - G-band, 242, 243
 - Gene technology, 525
 - Generalized functions, 49
 - Germanium, 182, 183, 195–197, 200
 - Germanium ad-atom, 202
 - Germanium dimer, 198, 199
 - Germline cells, 585
 - GGA, 270
 - Glass substrates, 372
 - Globally accepted standards, 122
 - GLU122, 146
 - Glycoconjugated polymer system, 513
 - G-Protein coupled receptors, 139
 - Grüneisen parameter, 345
 - Grain radius, 345

- Graphene, 121, 178, 181, 182, 184, 186, 187, 194, 196, 205–208, 213, 221, 228, 229, 239, 262, 385
 Graphene band structure, 225
 Graphene bandgap, 217
 Graphene base transistor, 182, 186, 194–196
 Graphene/metal interface, 223, 226
 Graphene nonlocality, 217
 Graphene oxide (GOX), 262
 Green's function, 49, 51, 53, 56, 64, 67, 73, 85
 Green synthesis methods, 507
 GROMACS, 136
 Group II–VI compounds, 25
 Group velocity, 223, 228
 Growth process, 319
 Growth time, 363, 374
- ## H
- Halloysite, 494
 Hamiltonian, 14
 Hamiltonian matrix, 22
 Harmonic oscillator, 51
 Harrison Scaling Law, 20
 Hartree Coulomb term, 33
 Hazard, 255
 HBAC, 137
 Healthcare, 585
 Healthcare Industry, 110
 Heavy-hole, 8
 α -helices, 144
 Helicon Wave excited Plasma Sputtering Epitaxy (HWPSE), 164
 Helium treatment, 445
 Helmholtz, 73, 74, 79
 Helmholtz Green's function, 103
 Helmholtz scalar wave, 86
 Heteroepitaxial growth, 7
 Heterointerface, 2, 161
 Heterojunction bipolar transistor (HBT), 4
 Heterostructures, 1, 161
 - Type I, 2
 - Type II, 2
 - Type III, 3
 II–VI heterostructures, 42
 III–V heterostructures, 42
 H₂ evolution efficiency, 443
 Hexagonal boron nitride, 182, 188, 188–190, 195, 201
 Hexamethylenetetramine, 372
 HF exposure, 249
 Hierarchical structural architectures, 547
 High-Affinity Fusion Tag Targeting Approaches, 569
- High-resolution electron energy loss spectroscopy (HREELS), 205
 High temperature, 439
 High-temperature calcination method, 438
 H₂ molecule adsorbed, 252
 H₂O, 267
 H₂O/Ti mole ratio, 433
 Homogenous thin fibers, 302
 Hopping terms, 22
 Hot injection, 508
 H₂ production, 442
 HREELS, 223, 224
 HRTEM, 452
 HSE+U* calculations, 409, 410, 418
 HSE+U approach, 409
 Hubbard U, 408, 411
 Huisgen-type click chemistry, 513
 Hybrid-functional+U approach, 408
 Hybrid heterostructure, 370
 Hybrid solar cell, 368, 369
 Hybridization, 228
 Hydrazine, 358
 Hydrogen and energy storage, 246
 Hydrogen bond dynamics, 134
 Hydrogen bonds, 138
 Hydrogen production, 441
 Hydrophilic QDs, 508
 Hydrophobic nanocrystals, 560
 Hydrostatic strain, 9
 Hydrothermal, 429, 432, 487
 Hydrothermal synthesis, 558
 Hydrothermal technique, 357, 371, 374
 Hydroxyl groups, 278
 Hyperfine splitting constant, 107
- ## I
- Image processing, 654
 Imaginary part, 213
 Imaging, 514
 Immunostaining specificity, 562
 Impedance, 461
 Impenetrable barrier, 49
 In-Situ Low Temperature Optical Absorption Spectroscopy, 346
 In ternary and quaternary alloys, 2
 In vitro imaging, 566
 In vitro targeting, 515, 518, 526
 In vivo, 526
 in vivo applications, 520
 In vivo imaging, 568
 in vivo vascular imaging, 522
 Incident field, 80
 Incident wave, 79, 88
 Incident wave vector, 83, 87

- Infinite barrier model, 210
- Inflammation, 521
- Influence factors, 429
- InGaN/GaN, 24, 43
- Initial condition of O operator, 618, 624
- InN, 309, 310
- InN thin films, 312, 314
- Integral equation, 73–76, 103
- Integrated nanosystems, 108
- Inter-laboratory comparison, 116
- Interaction picture, 610
- Interband, 222
- Interband plasmons, 222
- Interband transitions, 225
- Intercalated, 228
- Intercalated graphene, 228
- Interface strain, 6
- Interference fringes, 103
- Internal quantum efficiency, 159
- International Avogadro project, 107
- International comparisons, 115
- Interstitial atom, 199
- Intraband excitations, 225
- Intraband plasmons, 223
- Intrinsic band bending, 317
- Introduction, 239
- Inverse dielectric function, 50, 59, 206, 208, 211, 213, 214
- Inverse dispersion factor, 215
- Inverse dispersion function, 218–220
- Inverted band structure, 314
- $\text{In}_x\text{Ga}_{1-x}\text{N}$, 155
- Ion-insertion, 298
- ITRS roadmap, 124
- K**
- Kinetic, 481
- Kohn-Hohenberg-Sham studies, 33
- Kohn-Sham, 263
- k .p method, 13
- k -points, 225, 228
- L**
- Lab-on-a-chip, 593
- Landau damped, 218–220
- Landau damping, 217, 225, 229
- Landau quantization, 58
- Langmuir, 480
- Langmuir-Blodgett deposition, 382
- Laser-assisted, 172
- Laser ablation, 244
- Laser calibration, 117
- Laser diodes (LDs), 150
- Lateral wave vector, 75
- Lattice matched AlGaAs/GaAs, 42
- Lattice mismatch, 6, 19, 362
- Layer-surface separation, 229
- Layered titanate, 430
- Layers, single and double, 205
- LCNAO, 267
- LDPC codes, 654
- Leachate, 478, 497, 498
- Ligand conjugated QDs, 569
- Ligand exchange, 508, 510
- Light absorption, 354, 444
- Light amplification/photronics, 112
- Light emitting diodes (LEDs), 150
- Light-hole, 8
- Light trapping, 366
- Lindblad master equation, 615
- Linear dispersion, 222
- Linear medium, 207
- Linear plasmon, 228
- Lipid bilayer, 140
- Liposomes, 584
- Local, 206, 213
- Local defects, 376
- Local density approximation (LDA), 34, 269
- Long and short wavelengths, 221
- Long wavelength limit, 229
- Lotus effect, 542
- Low dimensional, 49
- Löwdin orbitals, 13
- Low frequency dynamics, 139
- Lower bound, 641
- M**
- M_{11} , 242
- Magnetic field, 56
- Magnetic nanoparticles, 110
- Magnetic switch-based structures, 638
- Magnetite, 537
- Magnitude of Z , 468
- Markers, 583
- Markov limit, 617
- Massless graphene, 213
- Master equation, 609
- Matching, 646
- Maximally entangled state, 626
- MBJLDA functional, 11, 36
- Mechanism, 451
- Mechanism model, 454
- Mechanosynthesis, 112
- Memory time, 626
- Memristive arrays, 638
- Mesoporous, 296, 297
- Metal-insulator-semiconductor (MIS) devices, 249

- Metal, 213
- Metal assisted chemical etching, 356
- Metal based toxicity, 559
- Metal nanoclusters, 564
- Metal nanoparticles, 357
- Metal organic chemical vapor deposition (MOCVD), 2
- Metallic, 241
- Metallic tungsten precursor, 299
- Methods, 429
- Methyl orange, 448
- Metrological AFMs, 115
- Metrology, 105
- MgZnO, 164
- Mica, 181–183, 190, 192–194, 201
- Microbial behaviour, 599
- Microbial strain development, 600
- Microbioreactors, 598
- Microchannels, 597
- Microchip capillary electrophoresis, 596
- Microelectrophoretic separation, 595
- Microemulsion-mediated solvothermal method, 434
- Microfluidic bioreactors, 597, 598
- Microfluidic channels, 597
- Microfluidic DNA microarray, 595
- Microfluidic DNA microarray technology, 594
- Microfluidic PCR systems, 593
- Microfluidic polydimethylsiloxane (PDMS) bioreactor, 597
- Microfluidics, 589, 591, 592, 594, 599, 600, 602
- Micropore, 485, 491, 501
- Microwave-assisted irradiation, 558
- Migration behavior, 456
- Migration direction, 452
- Miniaturized 2-D gel electrophoresis, 596
- Minimization, 649
- Mixed-phase TiO₂, 423, 424, 440, 441, 444
- MLG, 221, 224–226, 228
- Mobility, 206
- Mobility direction, 451
- Model, 472
- Model potential, 210
- Modelling, 49, 333
- Modelling of electronic structure, 5
- Modification, 484, 493
- Modification of TiO₂, 424
- Modified, 474
- Modified Becke-Johnson density functional (MBJLDA), 34
- Modified virtual crystal approximation (MVCA), 10, 17, 24
- Molecular beam epitaxy, 2, 182, 183, 186–191, 197, 200, 202
- Molecular dynamic modelling, 123
- Molecular imprinting, 538
- Molecular metrology, 109
- Monoclinic crystalline, 294
- Monolayer, 220, 229
- Monolayer graphene (MLG), 218, 220, 222, 223, 226, 229
- Monte-Carlo-Simulations, 123
- Monte Carlo modelling, 123
- Montmorillonite, 491
- Morphology, 474
- MOVPE, 160
- MQW, 158
- mRNA, 525
- Muffin-tin approximation, 152
- Multi-walled nanotubes, 241
- Multimodal approach, 516
- Multiphoton, 174
- Multiple-qubit system, 617, 628
- Multiplexed fluorescent assay, 518
- Multiplexing, 523, 557
- Multiwalled CNTs, 388
- N**
- Na⁺, 430
- NaCl, 430
- Nano-biopsy, 584
- Nano-crossbar arrays, 635
- Nano-crossbar switching arrays, 659
- Nano-hole, 73, 74, 80, 82, 83, 85, 86, 95–97, 102, 103
- Nano-imaging, 584
- Nano-scale, 76
- Nano-scale aperture, 76
- Nanoantennae/metamaterials, 112
- Nanobiomedical, 571
- Nanobot components, 112
- Nanocarriers, 571
- Nanochannels, 539
- Nanocrystal radius, 334
- Nanoelectronics, 246, 262
- Nanofiber, 295, 299, 301
- Nanogold, 109
- Nanomachines, 109
- Nanomachining, 175, 176
- Nanomaterials, 292
- Nanomaterials embedded in complex matrix materials, 124
- Nanomedicine, 572, 579, 585
- Nanometrology, 111
- Nanoparticulate, 581
- Nanopharmaceuticals, 581

- Nanopillar arrays, 366
 - Nanopillars, 351
 - Nanoporous, 474, 477, 488, 501, 546
 - Nanoshells, 583
 - Nanostructures, 49, 172, 352
 - Nanosystems, 108
 - Nanotechnology, 580
 - Nanotheranostics, 584
 - Nanotube, 380
 - Nanowire and nanotube crossbar arrays, 638
 - Nanowire growth, 317
 - Nanowire growth mode, 320
 - Nanowire structure, 320
 - Nanowires, 2, 299, 390
 - Narrow band gap, 310
 - Natural adhesives, 540
 - Nearest neighbor (NN) sp^3 tight binding theory, 16
 - Nearest neighbor interaction, 30
 - Neurodegenerative diseases, 581
 - Next generation QDs, 564
 - NH_4OH , 430
 - N-(2-Hydroxypropyl) methacrylamide (HPMA), 582
 - NIR, 522
 - Nitrification, 500
 - NMA, 144
 - 2NN and TB models, 28
 - 2NN sp^3s^* orbitals, 22
 - 2NN sp^3s^* orbitals sets, 36
 - 2NN sp^3s^* TB model, 23, 31
 - 2NN sp^3s^* TB parametrization, 26
 - 2NN sp^3s^* tight binding parametrization, 22
 - 2NN sp^3 TB Hamiltonian matrix elements, 17
 - NN $sp^3d^5s^*$ orbitals set, 27
 - NN sp^3d^5 TB Hamiltonian matrix, 30
 - NN sp^3d^5 , 36
 - NN sp^3d^5 TB model, 31
 - NN sp^3d^5 and 2NN sp^3s^* TB models, 37
 - NN sp^3d^5 tight binding model, 29
 - NN sp^3 tight binding parametrization, 29
 - NO_2 , O_2 , H_2O , NH_3 , CH_4 , CO_2 , H_2 , N_2 , Ar, 251
 - NO_2 or NH_3 , 250
 - Non-covalent interactions, 512
 - Non-diffracting, 177
 - Non-Markovian, 609
 - Non-Markovian master equation, 622
 - Non-Markovian QSD equation, 614, 623
 - Non-optimal, 641
 - Non-parabolic dispersion, InN conduction band, 316
 - Nonlinear dispersion, 225
 - Nonlinear mode, 225
 - Nonlinear plasmon, 224, 225
 - Nonlocal, 206, 208, 212, 229
 - Nonlocal plasmon, 229
 - Nonlocality, 213
 - Nonparabolic dispersion, InN conduction band, 315
 - Nonpolar m-plane, 159
 - Nonspecific binding, 570
 - Nonspecific uptake, 520
 - Novikov's theorem, 615, 616, 621, 628
 - NSET, 567
 - Nucleation, 173, 181–185, 187, 188, 190, 199, 201, 361
 - Nudged elastic band, 183
 - Number of 2D layers, 216
 - Number of literals, 640
 - Number of products, 640
 - Numerical analysis, 86, 88
 - N vacancy, 315, 316
 - Nyquist plots, 468
- O**
- Obtain any probability, 658
 - Off-diagonal matrix elements, 14
 - OH/H, 268
 - Ohmic type contacts, 377
 - On-site atomic energies, 16
 - One-dimensional nanostructures, 355
 - One-dimensional ZnO nanostructures, 401
 - One-to-one matching, 647
 - One qubit dissipative model, 616
 - Open circuit voltage, 365, 376, 381
 - O operator, 614, 619
 - O operator coefficient function, 620
 - O operator initial condition, 614
 - OPLS-All Atom, 136
 - Opsins, 139
 - Opsonins, 520
 - Optical, 291, 305
 - Optical absorption, 327, 336
 - Optical absorption tail, 343
 - Optical coloration efficiency, 305
 - Optical dielectric constants, 42
 - Optical imaging agents, 113
 - Optical phonons, 224
 - Optical ruler imaging system, 121
 - Optical Tweezers, 111
 - Optimal phase ratio, 454
 - Optimal size, 642
 - Optimum phase content, 454
 - Opto-electronic devices, 371
 - Organic dyes, 447
 - Organic solar cells, 368
 - Ornstein-Uhlenbeck, 622

- Ornstein-Uhlenbeck correlation function, 613
 Oscillator eigenfunctions, 53
 Ostwald ripening, 506
 Outlook, 456
 Overdamped, 217
 Overlap matrix, 14
 Oxide nanoparticles, 256
 Oxygen, 188–190, 192, 269
 Oxygen deficiency, 293
- P**
- P25, 441, 446
 Pair of layers, 229
 Passive nanostructures, 108
 Pauli repulsion, 193
 PBC method, 263
 PC, 470
 PCBM, 368
 p-d interactions, 27, 39
 p-doped, 250
 PEDOT:PSS, 389
 PEG, 511
 PEG polymer, 298
 PEGylated QDs, 562
 Peierls phase integral, 57
 Peptide conjugation, 517
 Peptide libraries, 537
 Peptideconjugated QDs, 568
 Perforated 2D layer, 79
 Perforated GaAs layer, 91, 92–94
 Perforated plasmonic layer, 74, 79, 85, 86, 96, 97, 102
 Perforated plasmonic wave, 87
 Permeable, 498
 Permutations, 647
 Perovskite structure, 292
 PET, 572
 Phase content, 436
 Phase transformation, 426, 438, 440
 Phase transition, 427
 Phenol, 449
 Phonon confinement, 346
 Phospholipid micelles, 512
 Photo anode, 384, 387
 Photocatalysis, 423
 Photocatalytic, 441
 Photocatalytic activity, 425, 427, 428
 Photocatalytic decomposition, 447
 Photocatalytic degradation, 448, 450
 Photocatalytic reduction activity, 445
 Photocurrent, 370, 384
 Photodetectors, 132
 Photo-generated carriers, 451
 Photo-generated electrons, 452
 Photogenerated free carriers, 376
 Photoisomerisation, 140
 Photoluminescence (PL), 168, 332, 333
 Photoluminescence spectroscopy, 318, 320, 322
 Photon Factory Advanced Ring, KEK, 160
 Photovoltaic, 248, 352
 Photovoltaic behavior, 370
 PL spectra, 334
 Planar boundaries, 210
 Plane wave-train, 83
 Plasma, 173
 2D, 214
 layer, 207
 slab, 208
 torch, 245
 π plasmon, 228
 Plasmon, 205, 206, 213, 220
 Plasmon dispersion, 64, 217, 225
 Plasmon modes, 218, 219
 Plasmonic layer, 74, 76, 80, 95, 103
 Plasmonic response, 74
 Plasmonics, 177
 Plasmonic screen, 73
 Plastic substrates, 255
 3D plots, 89, 91–95
 p-n heterojunction, 365
 p-n junction, 249, 365
 Poisson ratio, 7
 Polarizability, 60, 73, 74, 76, 79, 207, 214
 Polarization, 221
 Polarization of gapped graphene, 231
 Poly(L-arginine), 518
 Polymer, 296
 Polymerase Chain Reaction (PCR), 591–593
 Polyvinylpyrrolidone (PVP), 299, 301
 POPC, 143
 Popular topic, 456
 Population heterogeneity, 601
 Porous, 474
 Porous anodic alumina, 354
 Positional broadening, 103
 Post-annealing, 363
 Potassium coverage, 227
 Potential effect, 468
 Potentiodynamically, 465
 Power conversion efficiency, 248, 354
 20–30 ppb range, 256
 PProDOT-Me₂, 462
 Pretreatment, 485
 Printed/flexible electronics, 654
 Probabilities, 652
 Probability generation tree, 658
 Probe-sample interaction, 123

- Problems, 456
Production of CO and CH₄, 445
Propan-2-ol, 135
Proposed mechanism, 435
Prostate cancer, 519
Protein-based computer, 108
Protein activation, 139
Proteins, 138
Pseudomorphic InGaAs/GaAs heterostructures, 19, 42
p-type doping, 226, 323
PVD, 329
PW, 262
- Q**
QD—antibody conjugates, 521
QD toxicity, 560
QDs bioconjugates, 515
QDs with antibodies, 516
Quantitative reverse transcription polymerase chain reaction (qRT-PCR), 591
Quantized energy levels, 5
Quantum-confined Stark effects (QCSEs), 159
Quantum confinement effect, 333
Quantum dot, 2, 50, 390, 505
 three dimensional, 51
 wires wells, 50
Quantum effect, 506
Quantum mechanical amplitude, 69
Quantum mechanical perturbation theory, 151
Quantum state diffusion equation, 609
Quantum systems, 49
Quantum trajectory, 611
Quantum well, 5, 390
Quantum wire, 50, 56
Quantum yield, 507
Quasi-Fermi level, 384
Quasi-freestanding, 228
- R**
Radial, 353
Radial breathing mode, 242
Radial distribution functions, 156
Raman, 242, 318
Raman shift, 345
Raman spectra, 295
Raman spectroscopy, 249, 331, 340, 345
Random bit assigning, 654
Random bit shuffling, 655
Random phase approximation (RPA), 60, 208, 229
Random walk theory, 185–187, 189, 201
Rate constants, 440
RBAM, 655
RBSM, 655
Reaction time, 430
Real time, 591
Real-time PCR, 592
Realisation of the metre, 123
Recombination, 424
Reconfigurable crossbar architectures, 638
Rectification ratio, 365, 369
Red-shift, 227, 330
Red shifted, 346
Reduced density matrix, 615
Reduced matrix, 646
Reduction of CO₂ with water, 444
Reflected wave, 88
Regeneration, 487
Relationship, 435
Relativistic Dirac electrons, 206
Removal, 480
Repeatability, 116
Representation, 105
Resistance, 474
Resonant Raman scattering, 340
Responsivity, 379
Reversible coloration ability, 303
Review, 423, 425
Ring diagram, 208, 231
 2D, 214
RNA technology, 525
Rolled, 239
Room temperature, 253
RPA, 207, 214
2D RPA, 231
2D RPA density perturbation, 229
 $\Delta R/R_0$, 255
Rutile, 426, 436, 437
- S**
S₁₁, 242
S₂₂, 242
Saccharomyces cerevisiae, 601
Scalar field, 79, 85
Scalar Green's function, 74, 76
Scalar wave, 73, 74, 86, 103
Scanning tunneling microscope, 275
Schrödinger Green's function, 50
Schrodinger equation, 33
Screening, 223
Screening function, 59, 206
Seed layer, 372
Self-assemble, 547
Self-assembled nanoscale systems, 636
Self-assembly, 512, 536, 544
Self-cleaning, 542, 546
Self-doped graphene, 224

- Self-formation, 172, 173
- SEM, 113, 337, 347, 377, 474
- SEM and TEM images, 438
- II–VI semiconductors, 360
- III–V semiconductors, 360
- Semi-infinite conductor, 217–219
- Semi-infinite limit, 212
- Semi-infinite slab, 213
- Semi-infinite substrate, 216
- Semiconductor, 49, 150, 207
- Semi conductor nanocrystals, 556
- Semiconducting, 241
- Semiconductor devices, 2
- Semiconductor Industry, 109
- Semiconductor nanocrystals, 328
- Semiempirical NN sp^3d^5 TB model, 46
- Semiempirical Sp^3 tight binding modeling, 13
- Semiempirical Sp^3d^5 tight binding modeling, 30
- Semiempirical Sp^3s^* tight binding modeling, 21
- Semiempirical tight binding theory, 1
- Sensitivity, 255
- Series resistance, 376
- S^* -excited state, 31
- Sheet resistances, 382
- Shift the binary value, 658
- Shift the Fermi level, 253
- Short circuit current, 365
- Short circuit current density, 366, 376
- Short current density, 381
- Shunt resistance, 376
- SI microscopic representations, 107
- Si nanowires, 368
- Sidewall electron accumulation, 317
- SIESTA, 262
- Silanization, 509
- Silicene, 206–208, 229
- Silicon, 181, 182, 186, 187, 206
- Silicon carbide (SiC), 182, 190, 192, 193, 199, 201
- Silicon dioxide, 181–183, 190–193, 196, 197, 199, 201
- Silicon dots, 565
- Silicon oxide layer, 249
- Silicon, zinc-oxide, 361
- Simulation, 473, 649
- Single-cell characterization and analysis, 602
- Single-particle electronic orbitals, 32
- Single-particle excitation (SPE), 225
- Single-walled carbon nanotubes, 387
- Single cell genomics, 602
- Single walled carbon nanotube, 241
- Size- and shape-dependent energetics, 413
- Size distribution, 455
- Slab plasma, 206, 214
- Slater-Koster formalism, 13
- S-layer proteins, 547
- Small angle X-ray scattering diffractometer, 120
- SnO_2 , 256
- SO_4^{2-} concentration, 434
- Sol-gel method, 436, 508
- Solar cell, 351, 369, 539
- Solar energy, 352
- Solar simulator, 376
- Solar spectrum, 362, 365, 378
- Solid-solution-solid mechanism, 358
- Solution-based techniques, 356
- Solution concentration, 430
- Solution temperature, 363
- Solvent mixing and calcination method, 437
- Solvents, 461
- Solvothermal, 432
- Sonication, 383
- Sonochemical technique, 357
- Sort matrices, 645
- Source, 80
- $Sp^3d^5 s^*$ orbitals set, 29
- Sp^3d^5 orbital basis, 31
- Space elevator, 108
- Spatial compression, 103
- SPE interband, 225
- SPE intraband, 225
- Spectroscopic ellipsometers, 113
- Spectroscopy, 467
- Speed of generating quantum entanglement, 626
- Spin coating, 369
- Spin orbit coupling of p-states, 30
- Split-off valence band, 8
- SPring-8, 156
- Sputtering technique, 377
- Square-root-like dispersion, 221
- SQW, 155
- Stability, 426, 449
- Standard, 297
- Standard Berendsen, 144
- Standard heat capacity, 9
- Standard Reference Materials (SRM's), 121
- S^* state, 29
- π states, 228
- Static shielding, 60
- Statistical average, 612
- Statistical thermodynamic model, 9, 21
- Step height standard, 115
- Sterically hindered, 138
- Sticking coefficient, 182, 186, 187, 189, 201

- Stober approach, 510
Stochastic computing, 636, 652
Stochastic computing models, 659
Stochastic multiplication, 653
Strain, 343
Strain effects, 20
Strong coupling, 610
Structural, 291, 298, 305
Structural disorder, 343
Structural materials, 247
Structural properties, 36
Structure factor, 63
Structure properties, 425
Structured medium, 610
Stuck-closed, 643, 649
Stuck-open, 643, 649
Sturm-Liouville, 64
Sub-division of optical laser wavelengths, 123
Sub-nanometre step heights, 119
Sub-nm precision, 118
Sub-nm scales, 107
Subarray, 646
Substitutional atom, 188, 190
Substrate, 208
Substrate plasma, 206, 214
Subwavelength, 73, 74, 79, 86, 103
Sum-of-products (SOP), 640
Super insulation, 114
Supercapacitors, 463
Supercell, 263
Supercell approach, 401, 403, 404
Superhydrophobic, 542
Superlattice, 49, 59, 157
Supramolecular chemistry, 538
Surface, 205
Surface charge accumulation on InN
 Nanowires, 317–319
Surface electron accumulation, 309, 310, 312, 314–316
Surface modification, 508, 560
Surface morphology, 188–190, 197, 200, 202
Surface passivation, 367
Surface passivation treatment, 355
Surface plasmon, 213
Surface steps, 185, 188–190
Surface-to-layer separation, 218, 219
Surfactant, 298, 382
Switching crossbar nanoarray, 637
Symmetric/antisymmetric modes, 205
Synchrotron light sources, 149
Synchrotron radiation beam, 120
Synergistic effects, 442
Synthesis, 429, 506
Synthesis methods, 506
Synthesis results, 649
Synthetic, 485
- T**
Target function, 642
Targeting of drugs, 110
Tartaric acid, 432
Telomerisation, 525
TEM, 378
TEM images, 445
Terahertz spectroscopy, 131
Ternary/binary heterointerfaces, 43
Ternary semiconductor, 11
Tethered bilayer membranes, 545
Tetrabutylammonium hexafluorophosphate, 464
Tetrabutylammonium tetrafluoroborate, 464
Tetraethylammonium hexafluorophosphate, 464
Tetrahedral semiconductors, 16
Theranostic, 583
Thermal equilibrium state, 612
Thermal evaporation, 361
Thermal expansion, 6
Thermodynamic Green's function, 208
Thick slab, 210, 212
Thickness of the film, 342
Thin film, 175, 177, 292, 295, 327, 351, 376
Thiolated molecules, 525
Thiolated proteins, 518
Third-generation solar cells, 353
Three-qubit dissipative model, 623
THz emitter, 133
Tight binding, 242
Tight binding method, 12
Tight binding model with an sp^3d^5 orbital basis, 38
Time-local, 628
 TiO_2 , 256, 378, 424
 TiO_2 nanofiber, 385
 TiO_2 nanorods, 376
Tissue engineering, 113
Titania nanoparticles, 385
Titanium-dioxide, 361, 376
Titanium nitrides, 389
TM6 helix, 146
Top-down approach, 356
Topological insulators, 229
Total reflection, 84
Toxicity, 585
Toxicological studies, 559
Traceability, 105
Traceability to the lattice constant, 123
Traceable measurement techniques, 113

Transfer artefacts, 106
 Transmembrane, 139
 Transmission, 73, 74, 79, 103, 364
 Transmission maximum, 103
 Transmission mode, 154
 Transmitted field, 88
 Transmitted field distribution, 91–102
 transmitted field intensities, 89
 Transmitted scalar wave, 86
 Transmitted wave, 83
 Transparent conductive oxide, 369
 Trapping viruses, 110
 Treatment, 581
 Triphasic TiO₂, 448, 449
 Truouiller-Martins, 263
 Truth table, 642
 Transport measurements, 322
 Tungstic acid, 299
 Tungsten hexachloride, 301
 Tungsten oxide, 291, 295, 297
 Tungsten trioxide, 292
 Tunneling of electrons, 194
 Two dimensional (2D), 74
 Two dimensional electron gas (2DEG), 4
 Two dimensional plasmon, 103
 Two-qubit system, 619
 Two-step procedures, 353
 Type 1 superlattice, 50
 Types and characteristics of QD, 556

U

Ultrasonic waves, 358
 Ultra-violet, 369
 Uncertainty, 106, 116
 Undamped, 218, 219
 Undamped plasmons, 220
 Undamped plasmon spectrum, 205
 Undistorted bond length, 11
 Upper bound, 641
 Upper plasmon branch, 220
 Urbach, 345
 Urbach energy, 335, 342
 Urbach rule, 341
 Urbach tail, 341, 342
 Urea concentration, 435
 Use of quantum dots in biological systems, 557
 UV/Vis/NIR absorption spectroscopy, 242
 UV illumination, 296

V

Vacancy defects, 279
 Valence and conduction bands, 215
 Valence band edge, 321
 Valence band offset, 4, 41

van der Waals (vdW), 267
 van Hove singularities (vHs), 242, 248
 Vanadium doping, 448
 Vapor-liquid-solid (VLS), 359
 Vapor-phase based synthesis, 359
 Vapor-solid (VS), 359
 Variable mass, 64
 VASP, 262
 vdW forces, 187
 Virtual crystal approximation (VCA), 10
 Visible, 369

W

W–O–W mode, 301
 Wafer inspection, 110
 Wafer transfer artefacts, 115
 Wastewater, 497
 Water, 135
 Water solubility, 512
 Wave packet transmission, 69
 Wave propagation, 73, 74, 103
 WCl₆/PVP blend, 301
 Wet-chemistry, 382
 Wetland, 500
 WIEN2K simulations package, 11, 34, 46
 WKB approximation, 341
 WO₃, 256
 Wurtzite, 157

X

Γ , X and L symmetry points, 38
 X-ray absorption fine structure (XAFS), 149
 X-ray absorption near edge structure (XANES), 153
 X-ray diffraction (XRD), 339, 347, 426
 X-ray fringes, 119
 X-ray interferometry, 119
 X-ray reciprocal space mapping (XRSM), 160
 XANADU, 152
 XPS, 252
 XRD pattern, 294

Z

Zeolite, 477, 480, 484, 485, 493, 501
 Zero bandgap, 218, 219
 Zero radial size, 78
 Zero temperature, 231
 Zinc-blende crystal structure, 34
 Zinc nitrate hexahydrate, 372
 Zinc oxide varistors, 402
 Zinc oxide (ZnO), 371, 372
 ZnO nanowires, 401–405, 412, 413
 ZnO thin film, 374
 ZnSSe/ZnSe, 44