Topics in Heterocyclic Chemistry 38
Series Editors: B.U.W. Maes · J. Cossy · S. Polanc

Frank De Proft
Paul Geerlings *Editors*

Structure, Bonding and Reactivity of Heterocyclic Compounds

Springer

38
Topics in Heterocyclic Chemistry

Aims and Scope

The series Topics in Heterocyclic Chemistry presents critical reviews on present and future trends in the research of heterocyclic compounds. Overall the scope is to cover topics dealing with all areas within heterocyclic chemistry, both experimental and theoretical, of interest to the general heterocyclic chemistry community.

The series consists of topic related volumes edited by renowned editors with contributions of experts in the field.

More information about this series at
http://www.springer.com/series/7081

Frank De Proft • Paul Geerlings
Editors

Structure, Bonding and Reactivity of Heterocyclic Compounds

With contributions by
R. Broer • J. Cadet • S. Catak • N. De Kimpe • F. De Proft • M. D'hooghe • F. Feixas • P. Geerlings • H. Goossens • A. Grand • M.Z. Griffiths • R.W.A. Havenith • D. Hertsen • V.T.T. Huong • N. Jorge • L. Joubert • V. Labet • E. Matito • K. Mollet • C. Morell • M.T. Nguyen • J. Poater • P.L.A. Popelier • Z. Rashid • M. Solà • O.A. Syzgantseva • T.B. Tai • V. Tognetti • J.H. van Lenthe • V. Van Speybroeck • M. Waroquier

Springer

Editors
Frank De Proft
Paul Geerlings
Research group of General Chemistry
Vrije Universiteit Brussel (VUB)
Brussels
Belgium

ISSN 1861-9282 ISSN 1861-9290 (electronic)
ISBN 978-3-642-45148-5 ISBN 978-3-642-45149-2 (eBook)
DOI 10.1007/978-3-642-45149-2
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014953968

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This volume of "Topics in Heterocyclic Chemistry" focuses on the aspects of structure, bonding and reactivity of heterocyclic systems. It aims at presenting a series of different state-of-the-art methods and concepts available in quantum mechanics and theoretical chemistry to probe these different important aspects of heterocyclic compounds.

In quantum chemistry, quantum mechanics is applied to problems in chemistry. An important goal in this field, among others, is to accurately solve the time-independent Schrödinger equation for molecular systems, yielding the different energy states of the system and the accompanying wave functions, from which all properties can be obtained. Unfortunately, this equation can only be solved exactly for one-electron systems; for many-electron systems, approximations have to be made. In recent decades, powerful and accurate methods have been developed. Accompanied by the ever-increasing power of computers and the availability of dedicated specialized software, accurate calculations of different properties of compounds of ever-increasing size can now be performed. These properties include, among others, molecular structures and energetics, such as reaction energies and activation barriers, both of importance to describe thermodynamic and kinetic aspects of a chemical reaction. This often requires the explicit treatment of the effect of the solvent, which can play a crucial role in determining chemical reactivity. Often, use is made of density functional theory methods, which, instead of the wave function, adopt the electron density as the basic quantity to describe all atomic and molecular properties. Another important aspect of quantum chemistry is that it serves as a source for the evaluation and computation of concepts. Chemistry has traditionally used and continues to use a series of powerful concepts in order to explain chemical behaviour. These concepts are not only invoked to provide rationalization, but they are also applied in a predictive manner. Already from the early days of quantum mechanics, a lot of attention has been devoted to the computation of these chemical concepts, which are often not experimentally measurable or, in quantum mechanical terms, correspond to non-observables. In this volume, attention will also be devoted to the use of a selected number of these concepts to rationalize the bonding and reactivity of heterocyclic compounds.

In the first chapter in this volume, Van Speybroeck et al. show how computational chemistry and molecular modelling, in synergy with experiment, can play an important role in the elucidation of various aspects related to the reactivity of aziridinium ions, important intermediates in the synthesis of functionalized amines and other nitrogen-containing organic compounds. This chapter focuses on the computation of ring strain and the role of the solvent, N- and C-substituents and nucleophile on the kinetics (through activation-free energies) of regio- and stereocontrolled ring opening of the aziridinium ions. Density functional theory-based reactivity descriptors are introduced to probe the regio- and stereochemical preferences in the ring-opening reaction. These reactivity descriptors are discussed in more detail by Morell et al. in the subsequent chapter, paying attention, among others, to the intricacies of the less frequently used selectivity descriptors for excited states. In addition, these authors describe their application, in particular of the so-called dual descriptor, to a very important series of heterocyclic ring systems, DNA bases. It is investigated how both global and local DFT-based reactivity indices can provide insights into the occurrence and formation mechanisms of DNA lesions.

The role of the electron density as a central quantity in density functional theory is already emphasized in the above chapters. However, the analysis of the topology of this function has been proven to also give important insights into the structure, bonding and reactivity. This approach is the central topic in the chapter of Popelier et al., focusing in detail on the characterization of five-membered heterocyclic ring systems through the theory of quantum chemical topology (QCT). After a short tour d'horizon of QCT, the authors show that ring atom properties can be obtained from the ring critical point (RCP) properties, and in their quest for determining a ring's net charge, they introduce the concept of ring characteristic orthogonality, with different characteristics of a ring affecting the ring's properties orthogonally. The RCP and ring atom properties of the model are finally tested against the IGPD inhibitor molecules.

The majority of quantum chemical calculations nowadays are performed within the molecular orbital (MO) framework. In this approach, molecular orbitals are constructed as a linear combination of (atom-centred) atomic orbitals. These molecular orbitals on the other hand are delocalized over the whole molecule and, thus, provide no straightforward connection to the Lewis picture of bonding often used by chemists. Valence bond (VB) theory was introduced almost at the same time as the MO approach around 1930, but received much less attention in subsequent decades due to its computational intricacies. In recent years, it regained interest as described by Havenith et al. in the chapter "Valence Bond Theory in Heterocyclic Chemistry," providing an alternative method that more closely connects to the concept of a molecule being atoms connected by bonds consisting of electron pairs. The method is subsequently discussed within the context of heterocyclic rings, leading to the conclusion that VB is a strong tool in the interpretation of the electronic structure in terms of classical chemical concepts such as bond bending, hyperconjugation and aromaticity. The last concept, of ubiquitous importance when studying unsaturated rings, is further elaborated on

in the final two chapters of the volume in the case of heterocyclic ring systems. As stated and discussed in detail in the contribution of Sola et al., this concept, related to the cyclic delocalization of electrons, cannot be defined without ambiguity, and it is often indirectly "measured" on the basis of its repercussion on various properties of the molecule. These authors discuss this concept for both organic and inorganic monocycles with special emphasis on the measures based on electron delocalization properties of aromatic rings.

In the final chapter, Nguyen et al. focus on the aromaticity of polyheterocyclic systems containing sulphur and their derivatives, such as annulated oligothiophenes (sulflower) and their O, Se, PH, PF and NH analogues. In addition, this chapter also discusses the use of the electron localization function (ELF) to probe aromaticity and to gain insight into bonding in these particular systems. The concept of disk aromaticity is thereby proposed to be an effective measure of the aromaticity of polyheterocycles.

In summary, the volume presents a broad overview of concepts and the way they can be quantified by state-of-the-art techniques to study the structure, bonding and reactivity of heterocyclic compounds.

Brussels, Belgium

Frank De Proft
Paul Geerlings

Contents

Reactivity of Aziridinium Salts in Different Solvents Unraveled by a Combined Theoretical and Experimental Approach

Hannelore Goossens, Dietmar Hertsen, Karen Mollet, Saron Catak, Matthias D'hooghe, Frank De Proft, Paul Geerlings, Norbert De Kimpe, Michel Waroquier, and Veronique Van Speybroeck

Contents

H. Goossens • D. Hertsen • M. Waroquier • V. Van Speybroeck (✉)
Center for Molecular Modeling, Ghent University, Technologiepark 903, 9052 Zwijnaarde, Belgium
e-mail: veronique.vanspeybroeck@ugent.be

K. Mollet • M. D'hooghe • N. De Kimpe
Department of Sustainable Organic Chemistry and Technology, Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, 9000 Ghent, Belgium

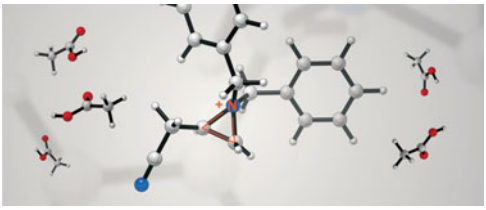
S. Catak
Center for Molecular Modeling, Ghent University, Technologiepark 903, 9052 Zwijnaarde, Belgium

Department of Chemistry, Bogaziçi University, Bebek, 34342 Istanbul, Turkey

F. De Proft • P. Geerlings
Research group of General Chemistry, Vrije Universiteit Brussel (VUB), Pleinlaan 2, 1050 Brussels, Belgium
e-mail: fdeproft@vub.ac.be; pgeerlin@vub.ac.be

F. De Proft, P. Geerlings (eds.), *Structure, Bonding and Reactivity of Heterocyclic Compounds*, Topics in Heterocyclic Chemistry 38, DOI 10.1007/978-3-642-45149-2_1

Abstract This chapter focuses on the importance of aziridinium ions as intermediates in organic chemistry. The principal aim is to gain insight into the factors to take into account for the selective synthesis of a variety of functionalized amines via aziridinium salts, such as the nature of the aziridinium ion (ring strain and *N*- and *C*-substituents of the aziridine ring), the nucleophile, and the solvent environment. Molecular modeling is used to investigate kinetics, electrostatics, and frontier molecular orbitals of reactions involving intermediate aziridinium ions, such as the nucleophilic ring opening of aziridines, the ring expansion of nitrogen heterocycles, and the ene reactions with triazolinedione.



Keywords Aziridinium ions • DFT • Ring strain • Regioselectivity • Solvation

Abbreviations

AI	Aziridinium imide intermediate
CM3	Charge model 3
CSE	Conventional strain energy
CSG	Gibbs free coordination solvation energy
DFT	Density functional theory
Hirshfeld-I	Iterative Hirshfeld
LG	Leaving group
NPA	Natural population analysis
OI	Open intermediate
PM3	Semi-empirical parameterized model number 3
QM/MM	Quantum mechanics/molecular mechanics
TAD	Triazolinedione
TS_{abs}	Transition state for proton abstraction
TS_{add}	Transition state for addition
TS_{iso}	Transition state for isomerization

1 Introduction

The aziridine moiety **A** (Fig. 1) represents one of the most valuable three-membered ring systems in organic chemistry, due to the uncommon combination of reactivity, synthetic flexibility, and atom economy [1–11]. Indeed, ring strain renders aziridines susceptible to ring-opening reactions that dominate their chemistry and makes them useful synthetic intermediates in the arsenal of the organic chemist. The regio-controlled ring opening of *C*-substituted aziridines constitutes a powerful approach toward the preparation of a large variety of functionalized nitrogen-containing target compounds [12–15]. The ring opening of activated aziridines, i.e., aziridines with an electron-withdrawing group at the nitrogen atom, has shown to be quite straightforward, mostly involving the nucleophilic attack at the less-hindered aziridine carbon atom [16–20]. On the other hand, nonactivated aziridines, which do require quaternization toward an aziridinium intermediate **B** for nucleophilic ring-opening reactions due to an electron-donating group at the nitrogen atom, often have different reactivity and applications. They provide interesting opportunities for the selective synthesis of a variety of functionalized amines. It should be noted that aziridinium ions not only can be obtained through *N*-functionalization of neutral aziridines [21] but also through intramolecular substitution of amines bearing a leaving group at the β-position [22].

This chapter mainly covers work that was performed by the authors. In the first part, different types of strain and different methods to determine ring strain are described and a rationalization of the ring strain of different aziridinium ions will be given based on theoretical calculations. Next, a short overview will be given of the formation of aziridinium ions **B** from both the rearrangement of β-amino alcohols and the ring opening of nonactivated aziridines, and the reactivity of aziridinium ions will be investigated theoretically. In addition, the role of bicyclic aziridinium ions **C** as intermediates in the ring expansion of aziridines to azetidines will be unraveled. In the last part, we highlight the importance of spiro aziridinium ylides **D** as intermediates in ene reactions involving the highly reactive electrophile triazolinedione (Acevedo et al. [23]).

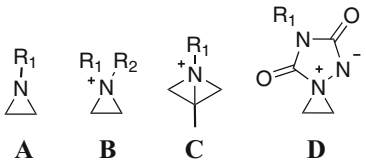


Fig. 1 Aziridines **A**, aziridinium ions **B** and **C**, and aziridinium ylides **D**

2 Ring Strain

Ring strain is an effect that destabilizes cyclic compounds and in this context may make aziridines and aziridinium ions, which act as versatile intermediates in organic chemistry [7], susceptible to ring-opening reactions. Ring strain is characterized by the increase in energy due to the ring structure [24–26] and plays an important role in the thermodynamics of aziridinium ions, as their existence makes the synthetically useful ring-opening reactions of these intermediates favorable. Despite the important role of ring strain in certain activated aziridinium intermediates, it is rarely assessed quantitatively in experimental or theoretical studies, and the concept has not often been introduced in a formal way. A better understanding of various ring structures will be performed from a detailed theoretical analysis. For that reason, the ring strain of several ammonium species **3–8** (Fig. 2) is investigated.

Ring strain itself is composed of at least four different types of strain: Baeyer or angle strain, stretching strain, Pitzer or conformational strain, and Dunitz-Schomaker or transannular or steric strain (Fig. 3) [27–32].

Baeyer strain arises from bond angle expansion or compression relative to acyclic, unstrained species. This type is most pronounced in small rings. In aziridinium ions, for example, the endocyclic bond angle centered on a tetrahedral center is reduced from 109.5° to 60°. Baeyer strain is the most obvious type of strain. However, its effects are not always straightforward, as illustrated by the relatively low strain energy of cyclopropane compared to cyclobutane (115.1 and 110.1 kJ/mol, respectively) [33]. Indeed, three-membered rings are often stabilized by σ delocalization, which is characterized by an increase of electron density in the center of the ring [34].

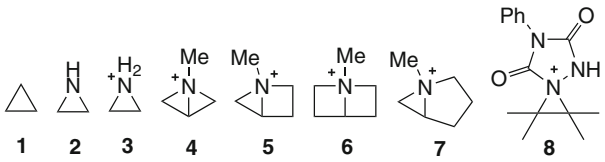


Fig. 2 Cyclopropane **1**, aziridine **2**, and substituted ammonium ions **3–8**

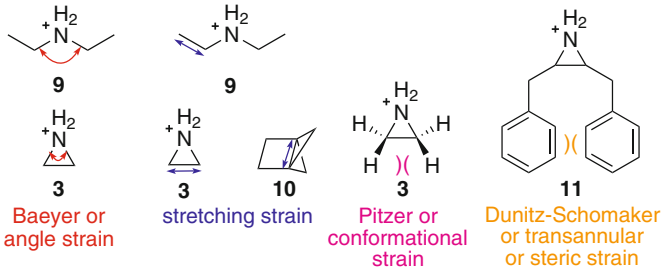
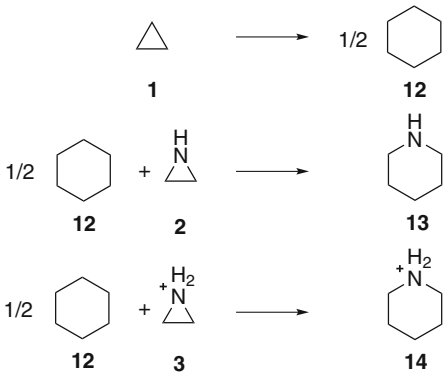


Fig. 3 Types of (ring) strain



Scheme 1 Unstrained isomers for ring strain determination of cyclopropane **1**, aziridine **2**, and aziridinium ion **3**

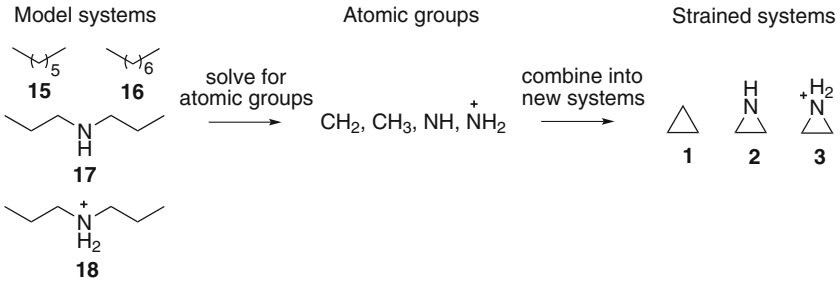
Stretching strain includes bond elongation and compression. Single carbon-carbon bonds are slightly shorter in three-membered rings than in acyclic alkanes (1.51 and 1.53 Å, respectively). However, this effect is more pronounced in tricyclo [2.1.1.$0^{1,4}$]hexane **10** [35]. In this tricyclic ring system, the bond between the two bridgehead carbon atoms is 1.75 Å.

Pitzer strain involves the 1,4-interactions across a bond and is determined by the dihedral angle. The stability of the chair conformation of cyclohexane, for example, is a result of a low Pitzer strain, since all 1,4-positioned hydrogen atoms are staggered. Finally, transannular strain is caused by repulsive steric interactions between atoms across a ring and is the sum of the van der Waals interactions between atoms that are separated by more than three formal bonds. The repulsion between axial substituents in cyclohexanes is a nice example.

Thus, strain energy is the sum of multiple effects, which can be either repulsive (e.g., angle strain in small rings) or stabilizing (e.g., σ delocalization in three-membered rings). This sum is often denoted as conventional strain energy (CSE) to highlight the subtle interplay between these effects [36].

CSE is defined as the difference between the heat of formation (or combustion) of a strained molecule and an equivalent system, as will be explained below [25]. The definition of an equivalent system is dependent on the determination method (Schemes 1 and 2), and the required thermodynamic data can be obtained either from experiments or from calculations. CSE is a gas-phase property and thus the effect of solvation is not taken into account.

The most straightforward way to determine ring strain is by comparing a strained molecule with an unstrained molecule (unstrained isomers, Scheme 1) [37, 38]. Strained cyclopropane **1**, for example, can be seen as one half of an unstrained cyclohexane **12** molecule. The equivalent systems must conserve certain chemical characteristics in a stoichiometric way, as will be explained later on. This approach has two drawbacks: to start with, it is not always possible to find an equivalent system and furthermore, six-membered rings are often the only good choice for an equivalent system. The latter problem is circumvented by assuming that the chair



Scheme 2 Additive schemes for ring strain determination of cyclopropane **1**, aziridine **2**, and aziridinium ion **3**

conformation of six-membered rings is strain-free, although this assumption can be questioned [38].

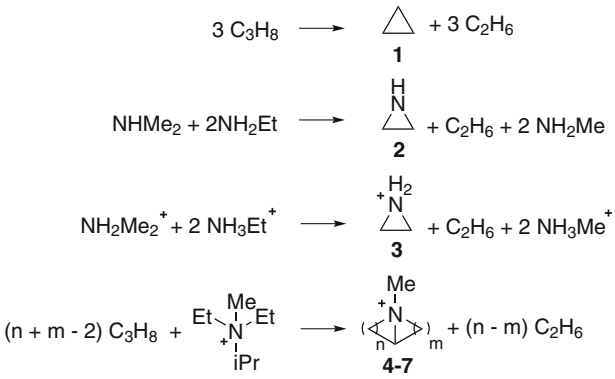
Additive schemes represent a second method for ring strain determination (Scheme 2) [39, 40]. In this approach, a strained molecule is divided into atomic groups. An aziridinium ion, for example, is a combination of two methylene groups (CH_2) and a nitrenium group ($^+NR_2$). The unstrained energy of these groups is estimated from unstrained, acyclic molecules in which the groups occur in a similar chemical environment and the sum of these energies is then subtracted from the energy of the strained species, yielding the ring strain.

The most commonly used methods are extensions of the first approach. In these methods, the strain energy is determined as the enthalpy change of a reaction in which all products and reagents, except the strained molecule, are acyclic. A number of chemical properties are conserved by the reaction, according to specific schemes. Wheeler et al. have constructed a generalized classification of these schemes, though only for hydrocarbons [25]. The differences between the schemes become less clear for heterocyclic compounds, although the classification is still manageable for simple systems [35, 41, 42].

The most common scheme is based on homodesmotic reactions (Scheme 3) [25, 43]. They conserve the number of atoms with the same atom number, hybridization state, and bonding partners (condition a) and the number of bonds with the same bonding partners and the same formal valence (condition b). A homodesmotic reaction for an aziridinium ion should conserve, inter alia, the number of carbon centers attached to a positive, tetravalent nitrogen center, two hydrogen atoms and a tetravalent carbon center (condition a) and the number of bonds between a positive, tetravalent nitrogen center and a tetravalent carbon center (condition b).

As mentioned above, ring strain can be determined either experimentally or computationally. Since experimental thermodynamic data for aziridinium ions are limited, the ring strain of several ammonium species **3–8** is investigated computationally. This approach can be systematically extended to estimate the ring strain of a large number of related compounds.

Unstrained isomers, an additive scheme, and a scheme based on homodesmotic reactions have been used to estimate the conventional ring strain of cyclopropane **1**,



Scheme 3 Homodesmotic reactions for strained compounds **1**–**7** ($n = m = 1$, $n = 1$ and $m = 2$, $n = m = 2$, and $n = 1$ and $m = 3$ for **4**, **5**, **6**, and **7**, respectively)

Table 1 Conventional strain energies (kJ/mol) of cyclopropane **1**, aziridine **2**, and ammonium ions **3**–**7** (M06-2X/6-311 + G(2df,2pd)//B3LYP/6-31 + G(d,p), 298 K, thermal enthalpy corrections were taken from B3LYP/6-31 + G(d,p) optimizations)

Three-membered rings					Bicyclic systems	
	Unstrained isomers	Additive schemes	Homodesmotic reactions	Experimental	Homodesmotic reactions	
1	102.4	104.1	105.7	115.1 [32]	**4**	279.7
2	105.4	107.3	110.2	111.7 [43]	**5**	208.7
3	152.8	163.0	160.1		**6**	178.2
					7	103.3

aziridine **2**, and aziridinium ion **3** (Table 1). In all cases, the conventional strain energy was obtained as a stoichiometric sum of the computationally calculated enthalpies of the molecules or groups involved.

Based on literature [25, 37–40, 43], it can be assumed that the scheme based on homodesmotic reactions gives the best results. This is not very surprising considering the rather simplistic nature of the additive scheme: an atomic group is defined solely by atomic number, charge, and valence of the center atom and the number of hydrogen atoms attached to it. For example, a methylene group in an aziridine is considered to be the same as a methylene group in an alkane. Furthermore, piperidine **13** and piperidinium **14** are used as six-membered equivalents for aziridine **2** and aziridinium **3**. Piperazine and piperazinium would be more convenient equivalent systems, though there is a significant repulsion between the two nitrogen atoms in these species.

However, there are no significant numerical differences between the three methods and it is not possible to compare the calculated numbers with the limited experimental results (Table 1).

Cyclopropane **1** and aziridine **2** have similar strain energies of circa 110 kJ/mol, which is the enthalpy reduction of a reaction that cleaves one of these rings. The kinetic barrier for this reaction will be too high for cyclopropane **1** because carbanions are extremely poor leaving groups. However, in the case of aziridine **2**, ring-opening reactions become more feasible since amide anions are better leaving groups.

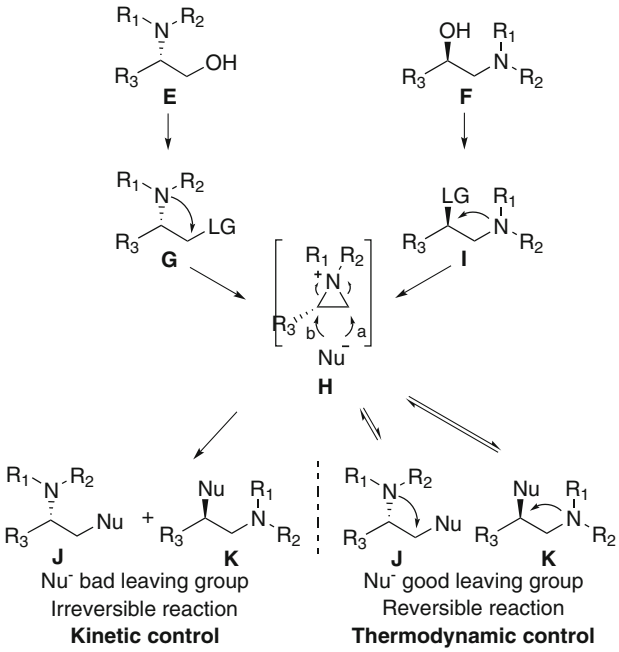
The most simple aziridinium ion **3** has a ring strain of 160.1 kJ/mol. The enthalpy reduction of ring-opening reactions is even larger in this case. Furthermore, the kinetic barrier is lower since neutral amines are excellent leaving groups. This explains the generally accepted statement that nonactivated aziridines must be activated prior to ring-opening reactions. The aziridine nitrogen atom can be activated by alkylation, acylation, or protonation to an aziridinium ion or by complexation with a Lewis acid.

Schemes based on homodesmotic reactions in combination with computational methods have been proven to be reliable to determine conventional strain energies [25]. This technique was applied to a series of bicyclic ions **4**–**7**, which have a single bond between the bridgehead atoms, defining two rings. Ions **4**–**6** have two aziridinium and/or azetidinium moieties, which results in a higher ring strain than the parent aziridinium ion. When the size of one of these rings increases, the ring strain decreases (Table 1). Remarkably, ion **7**, which has an aziridinium and a pyrrolidinium moiety, has a lower strain than ion **6**, which has two azetidinium moieties. According to these results, ion **7** even has a smaller ring strain than the parent aziridinium ion **3**.

As stated earlier, it is not always easy to define a good scheme for heterocyclic compounds. As an example, trying to define a scheme based on homodesmotic reactions for the spiro aziridinium imide ion **8**, one encounters the difficulty of conserving the size of the conjugated π orbital. A large acyclic molecule with hydrogen bond donors and acceptors should be included. The most stable conformation of such a molecule would involve hydrogen bonds that are not present in the ring structure, which leads to an artificially high strain energy. As a conclusion, a precise definition of the conventional ring strain and a general protocol to determine it for heterocyclic compounds are needed.

3 Reactivity of Aziridinium Ions

As mentioned in the introduction, aziridinium ions can be obtained through intramolecular substitution of amines bearing a leaving group at the β-position [20] or through *N*-functionalization of neutral aziridines [21]. An overview will be given of both pathways and the reactivity of aziridinium ions, regardless of their origin, will be investigated theoretically.



Scheme 4 Rearrangement of β-amino alcohols via aziridinium ions

3.1 Rearrangement of β-Amino Alcohols via Aziridinium Ions

The cyclization of β-amino alcohols is the most prominent representative of the formation of aziridinium ions through intramolecular substitution of amines bearing a leaving group at the β-position. Activation of the hydroxy group followed by addition of nucleophiles leads to the rearrangement of β-amino alcohols, mostly through an aziridinium intermediate (Scheme 4). The ratio of amines resulting from attack of nucleophiles at either the less substituted or the more substituted carbon atom of the aziridinium intermediate depends on the nature of the nucleophiles, the aziridinium substituents, and, to a lesser extent, the solvent and the temperature.

β-Amino alcohols **E** or **F** are converted into a derivate bearing a good leaving group LG (**G** or **I**, respectively) followed by displacement of the LG by a nucleophilic substitution reaction with formation of an aziridinium ion **H**. Subsequently, a nucleophile Nu, which can be the leaving group LG^- or an "external nucleophile", can attack **H** at either the less substituted (path a) or the more substituted (path b) carbon atom to produce amines **J** or **K**, respectively. If no external nucleophile is present in the reaction medium, β-amino alcohols **E** or **F** are rearranged if LG^- attacks the aziridinium **H** at the more substituted carbon atom with formation of amine **I** or at the less substituted carbon atom with formation of amine **G**, respectively. However, a mixture of **G** (**J** with Nu = LG) and **I** (**K** with Nu = LG) could be

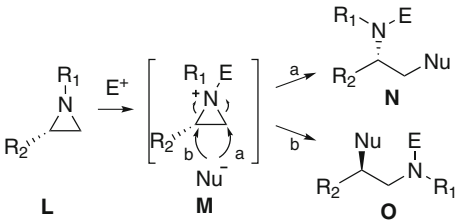
obtained as these compounds are in thermodynamic equilibrium with **H**. The proportion of amines **G** and **I** is the result of thermodynamic control and complete rearrangement of **E** or **F** will only be observed if **I** or **G**, respectively, is the more stable product. If an external nucleophile Nu is present in the reaction medium, the ratio of amines **J** and **K** depends mainly on Nu. If Nu is a good leaving group, the reaction is reversible and the proportion of **J** and **K** corresponds to a thermodynamic equilibrium (thermodynamic control). If Nu is a bad leaving group, the reaction is irreversible and the proportion of **J** and **K** corresponds to the regioselectivity of the nucleophilic attack on the aziridinium ion **H** (kinetic control).

The rearrangement of β-amino alcohols toward vicinally substituted amines via aziridinium intermediates has been discussed frequently in the literature [22]. It has been accepted that the regioselectivity of the nucleophilic attack across the aziridinium ions depends on the nature of the nucleophile and the nature of the substituents on the β-amino alcohol substrate. As already stated, the leaving group ability of the nucleophile has a major impact on the reaction outcome. If the nucleophile is a good leaving group, the thermodynamic product will prevail. On the other hand, if the nucleophile is a poor leaving group, kinetic control will dedicate the reaction outcome. Based on the literature data, a general overview is provided in Table 2 as a practical guide. From this overview, it can be deduced that when the R_3 substituent is an electron-withdrawing group (CO_2R), the more substituted carbon atom of the aziridinium intermediates **H** is preferentially attacked by the nucleophile, as this carbon atom is more electropositive, leading to the formation of amines **K** if the nucleophile is a bad leaving group (kinetic control). However, ring opening of 2-CF_3-aziridinium ions has recently been shown to proceed exclusively at C3, starting from the corresponding 1-alkyl-2-CF_3-aziridinium through *N*-protonation or alkylation [44]. In case the R_3 substituent is an aryl group, the more substituted carbon atom, i.e., the benzylic position, is in general prone to undergo nucleophilic attack, resulting in the formation of amines **K** regardless of the nucleophile used. The regioselectivity of the ring opening of

Table 2 Regioselectivity in the rearrangement of β-amino alcohols depending on the substrate and the nucleophile (R_1 = alkyl; R_2 = alkyl) [22]

	Nu									
R_3	Br^-	Cl^-	F^-	S[a]	O[b]	$CF_3CO_2^-$	SO_4^{2-}	DMF	N[c]	C[d]
Alkyl	**K**	**K**	**K + J**	**K**	**K + J**	**K**	**K**	**K + J**	**K + J**	**K + J**
CF_3	**J**	**J**	–[e]	–[e]	–[e]	–[e]	–[e]	–[e]	–[e]	–[e]
CO_2R	–[e]	**K**	**K**	**K**	–[e]	–[e]	–[e]	–[e]	**K + J**	–[e]
Aryl	**K**	**K**	**K**	**K**	**K + J**	**K**	**K**	–[e]	**K**	**K**

[a]Nucleophilic sulfur (SCN^-, *n*PrSH, *t*BuSH, PhSH)
[b]Nucleophilic oxygen (DMF, 4-$NO_2C_6H_4OH$, 2-$OMeC_6H_4OH$)
[c]Nucleophilic nitrogen ($(Me)_2NH$, $(Et)_2NH$, $(allyl)_2NH$, $(Bn)_2NH$, pyrrolidine, piperidine, piperazine, morpholine, imidazole, N_3^-, PhtNH, $TBSONH_2$, $MeNH_2$, *n*$BuNH_2$, *t*$BuNH_2$)
[d]Nucleophilic carbon (CN^-, 3-$OMeC_6H_4MgBr$, 4-ClC_6H_4MgBr, Me_2CuLi)
[e]No experimental data available



Scheme 5 Ring opening of nonactivated aziridines via aziridinium ions (E^+ stands for electrophile)

2-alkylaziridinium ions **H** is strongly influenced by the nucleophile employed for the ring opening.

3.2 Ring Opening of Nonactivated Aziridines via Aziridinium Ions

Nonactivated aziridines require quaternization toward an aziridinium intermediate prior to nucleophilic ring-opening reactions. *N*-alkylation, *N*-acylation, *N*-protonation, or *N*-complexation with Lewis acids gives rise to the formation of highly electrophilic aziridinium ions which can then easily be opened by different types of nucleophiles. Ring opening of aziridinium ions **M** can occur at the less substituted aziridine carbon atom (Scheme 5, path a) leading to amines **N**, in which the stereochemistry of the substituted aziridine carbon atom remained intact. Alternatively, ring opening can take place at the more substituted aziridine carbon atom (Scheme 5, path b) leading to amines **O**, with inversion of the stereochemistry of the substituted aziridine carbon by a S_N2 reaction.

The regioselectivity in the ring-opening reactions of nonactivated 2-substituted aziridines **L**, i.e., ring opening at the less (path a) and/or the more substituted carbon atom (path b) toward amines **N** and **O**, depends on a number of factors, including the nature of the nucleophile, the type of electrophile used for the activation of the aziridine moiety, and the nature of the substituents on the aziridine ring. Based on data published in the literature [21], a general overview is provided in Table 3. From this table, it can be concluded that when the R_2 substituent is a 1-alkenyl group, the nucleophilic attack generally occurs at the more substituted carbon atom of the aziridine moiety, independent of the type of electrophile used in these reactions. Also, activation of 2-arylaziridines is followed by nucleophilic attack at the more substituted carbon atom (benzylic position), furnishing the C2 ring-opening products **O** as the single or the major regioisomers. The regioselectivity of the ring opening of 2-acylaziridines appears to be electrophile dependent. However, in most cases, the products obtained result from the attack at the more substituted aziridinium carbon atom (i.e., the α-carbon atom with respect to the carbonyl moiety). The ring opening of 2-alkylaziridines is also influenced by the electrophile used for the activation, in most cases resulting in the ring opening of the aziridine moiety at the less-hindered aziridinium carbon atom. Bearing this in mind, it might

Table 3 Regioselectivity in the ring opening of nonactivated 2-substituted aziridines depending on the substrate and the electrophile (R_1 = alkyl; LA: *N*-complexation with Lewis acids; RCO^+: *N*-alkoxycarbonylation, *N*-carbamoylation, *N*-acylation, *N*-carboxylation; H^+: *N*-protonation; R^+: *N*-alkylation; TMS^+: *N*-silylation) [21]

	E^+				
R_2	LA	RCO^+	H^+	R^+	TMS^+
–CH = CH_2, –CH = CH–COOEt	**O**[b]	**O**	**O**[b]	**O**	**O**
Aryl	**O**	**O**	**O**	**O**[b]	**O**[b]
COR, COOR, $CONH_2$	**N**	**O**	**O**	**O**	**O**
Alkyl	**N**	**O** and/or **N**	**N**	**N**[a]	**N**

[a]Only halides attack the C2 position
[b]Proposed regioselectivity (no experimental data available)

be possible to predict a regioselective preference for other so far unexamined ring-opening reactions of nonactivated 2-substituted aziridines as well.

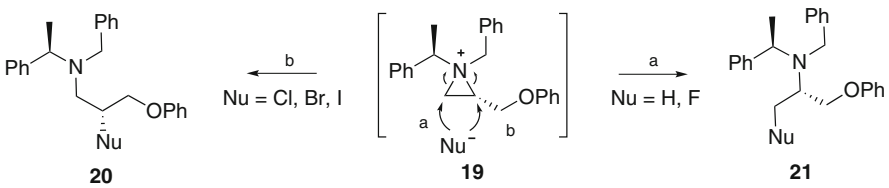
3.3 *Theoretical Investigation of the Reactivity of Aziridinium Ions*

Density functional theory (DFT) calculations have been performed to shed more light on the reactivity of aziridinium ions and to rationalize experimentally observed regio- and stereochemical preferences in their ring-opening reactions [45–50]. The nucleophile-dependent regioselectivity in the ring opening of 2-substituted *N*,*N*-dibenzylaziridinium ions is evaluated in detail. Comparison of reaction barriers and relative product stabilities helps to rationalize the nucleophile-dependent reactivity. Furthermore, the different regioselectivity in the ring opening of *N*,*N*-dibenzyl-2-(cyanomethyl)aziridinium ion and protonated *N*-benzyl-2-(cyanomethyl)aziridine by bromide will be investigated by a thorough analysis of the electronic structure of the aziridinium ions.

Nucleophile-Dependent Regioselectivity in the Ring Opening of 2-Substituted *N*,*N*-Dibenzylaziridinium Ions

The ring opening of 1-benzyl-1-(α(*R*)-methylbenzyl)-2(*S*)-(phenoxymethyl) aziridinium ion **19** with halides and hydrides is shown in Scheme 6. While the ring opening with chloride, bromide, and iodide exclusively takes place at the substituted aziridine carbon atom with formation of secondary β-halo amines **20** (path b), the reaction with fluoride gives rise to mainly primary amines **21** and ring opening with hydride occurs at the less substituted aziridine carbon atom with exclusive formation of amines **21** (path a) [48].

Since aziridinium ion **19** bears three aromatic functionalities, different conformations can occur as a result of stabilizing intramolecular π-π stacking interactions.



Scheme 6 Ring opening of 1-benzyl-1-(α(*R*)-methylbenzyl)-2(*S*)-(phenoxymethyl)aziridinium ion **19**

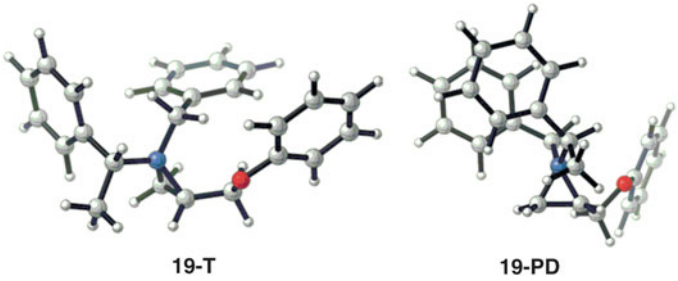


Fig. 4 T-stacking (*T*) and parallel-displaced (*PD*) interactions in aziridinium ion **19** [47]

It was found that T-stacking interactions are more favorable than parallel-displaced interactions (Fig. 4). Therefore, a conformer with a T-shaped configuration was used for further calculations.

The two competing paths for the ring opening of aziridinium ion **19** with halides (bromide and fluoride) and hydride donors (borohydride and aluminum hydride) were modeled, and consequent comparison of reaction barriers and relative product stabilities helps to understand the factors controlling regioselectivity for different nucleophiles.

Since nucleophilic substitution reactions are known to be influenced by the reaction conditions, the reactions under study were modeled with a proper solvent environment. Simulation of organic solvents can be performed in a continuum model [51–54], where the solvent is modeled as a continuous medium characterized by a dielectric constant. However, if explicit solvent interactions are present, discrete solvent molecules should be placed around the chemically active species to form a so-called supermolecule structure [45, 46, 55–60]. Alternatively, the supermolecule can be placed in a dielectric continuum leading to a mixed implicit/explicit model [61–63], but these models should be used with caution since they can give unreliable results with increasing amounts of solvent molecules and are highly influenced by their orientation [64]. Furthermore, embedding the supermolecule in a dielectric continuum does not always have an appreciable effect [50].

Ring-opening reactions of aziridinium ion **19** with bromide and fluoride have been modeled with and without the use of three explicit acetonitrile molecules to

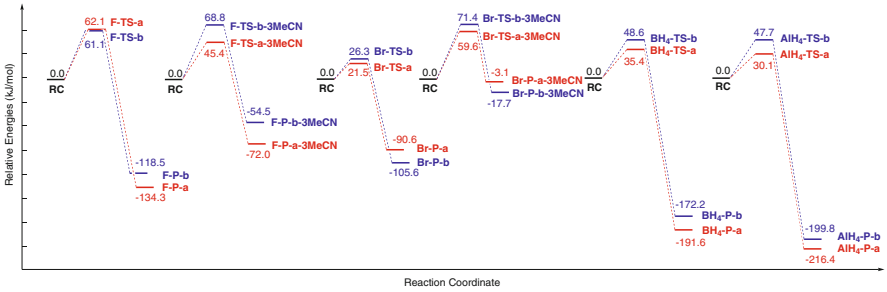
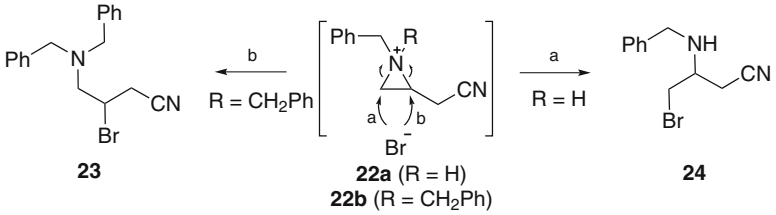


Fig. 5 Gibbs free energy profiles for the nucleophilic ring opening of aziridinium ion **19** with various nucleophiles (SCS-MP2/6-31++G(d,p)//B3LYP/6-31++G(d,p), free energies in kJ/mol at 298 K, thermal free energy corrections were taken from B3LYP/6-31++G(d,p) optimizations, 3MeOH stands for three explicit acetonitrile molecules used to solvate the fluoride or bromide ion) [47]

illustrate the benefit and necessity of solvating halide ions. The Gibbs free energy profiles for the nucleophilic ring-opening reactions of aziridinium ion **19** are shown in Fig. 5. Solvation significantly changes the nature of the potential energy surfaces. For bromide, barriers for paths a (ring opening at the less substituted aziridine carbon atom, Scheme 6) and b (ring opening at the more substituted aziridine carbon atom, Scheme 6) are quite comparable in the non-solvated case and tend to favor path a in the solvated case. However, relative product stabilities favor path b, and therefore, it is likely that thermodynamic equilibration dictates the outcome of the ring opening with bromide, if the barrier for the back reaction is feasible. In the non-solvated case, the back reaction barriers are too high for the reverse reaction to occur. However, in the solvated case, the back reaction barriers are much lower, allowing thermodynamic equilibration to yield the more stable product, which is the experimentally observed product. This is in line with the fact that bromide is a good leaving group. In the non-solvated fluoride case, barriers for paths a and b are quasi equal. However, there is a remarkable difference between the two pathways in the solvated fluoride case, manifestly favoring path a to be the kinetically preferred route and confirming the experimentally observed regioselectivity. Furthermore, relative product stabilities indicate that the kinetic product resulting from pathway a is also the thermodynamically most stable one. In the case of hydride donors borohydride and aluminum hydride, path a is again the kinetically preferred route and the corresponding product is the thermodynamically most stable and the experimentally observed product.

Thus, for hydride donors, attack at the less-hindered ring carbon of aziridinium ion **19** is the kinetical route and incidentally leads to the thermodynamic product. However, for the ring opening with halides, attack at the less-hindered ring carbon is always the kinetical route, but the hindered route can lead to the thermodynamic product. The eventual outcome depends on the leaving group ability of the halide. If the halide is a good leaving group, back reactions are sufficiently low to allow thermodynamic equilibration, but if the halide is a poor leaving group, the back



Scheme 7 Ring opening of 1-benzyl-2-(cyanomethyl)aziridinium ions **22** by bromide

reaction is unlikely and the kinetic route will dictate the reaction outcome. Depending on the equilibration rate, the kinetic product can be observed during the reaction or not.

Different Regioselectivity in the Ring Opening of *N*,*N*-Dibenzyl-2-(cyanomethyl)aziridinium Ion and Protonated *N*-Benzyl-2-(cyanomethyl)aziridine by Bromide

The ring opening of 1-benzyl-2-(cyanomethyl)aziridinium ions **22a** (R = H) and **22b** (R = CH_2Ph) by bromide is shown in Scheme 7. Ring opening of aziridinium ion **22a** affords 3-amino-4-bromobutanenitrile **24** via ring opening at the less substituted aziridine carbon atom (path a), whereas ring opening of aziridinium ion **22b** gives rise to 4-amino-3-bromobutanenitrile **23** via ring opening at the more substituted aziridine carbon atom (path b) [49].

In order to compare the two reaction mechanisms and rationalize the different reactivity, the effect of *N*-substituents on the aziridinium electronic structure was investigated. The correlation between aziridinium *N*-substituent identity and regioselective preference in the ring-opening reactions is explored by a thorough analysis of the electronic structures of aziridinium ions **22a** and **22b** from an electrostatic and a frontier molecular orbital point of view.

Population analysis can be used to gain insight into the electronic structure of molecular systems by calculating net atomic charges using a population scheme [65–70]. NPA (Natural Population Analysis) [71] and Hirshfeld-I (Iterative Hirshfeld) [72] charges for aziridine carbon atoms C2 (more substituted) and C3 (less substituted) of aziridinium ions **22a** and **22b** are shown in Table 4. Since C2 bears a more positive charge than C3 for both aziridinium ions, it is tempting to conclude that C2 is always the preferred atom to undergo a nucleophilic attack by the bromide ion. Hence, it is clear that the regioselectivity has no electrostatic driving force.

In a second attempt to rationalize the observed difference in nucleophilic attack, local Fukui functions $f(\mathbf{r})$ [73, 74], which can be used to describe orbital-controlled reactions, are analyzed. The three-dimensional Fukui function for nucleophilic attack is defined as

Table 4 NPA and Hirshfeld-I charges (MPW1B95/6-31++G(d,p)//B3LYP/6-31++G(d,p)) for aziridine carbon atoms C2 (more substituted) and C3 (less substituted) of aziridinium ions **22a** and **22b** [48]

	NPA		Hirshfeld-I	
	C2	C3	C2	C3
Aziridinium ion **22a**	−0.047	−0.235	0.107	−0.197
Aziridinium ion **22b**	−0.042	−0.229	0.070	−0.260

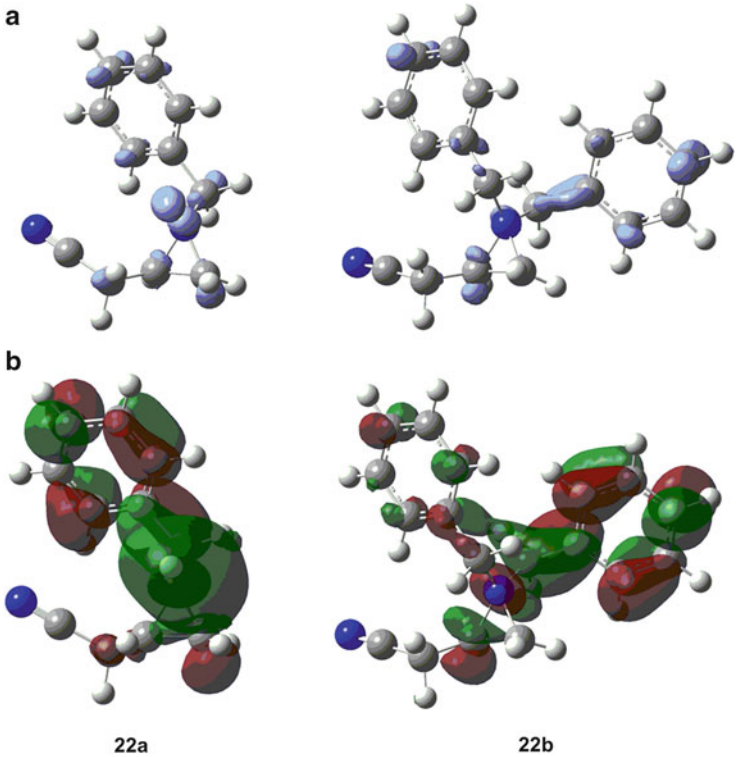
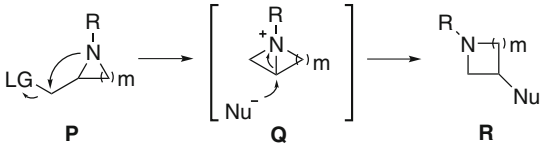


Fig. 6 Nucleophilic Fukui functions (isovalue 0.003 au) (**a**) and LUMOs (isovalue 0.03 au) for aziridinium ions **22a** and **22b** (**b**) [48]

$$f^{+}(\mathbf{r}) \approx \rho_{N+1}(\mathbf{r}) - \rho_{N}(\mathbf{r}) \approx \rho_{LUMO}(\mathbf{r})$$

Fukui functions and LUMOs for aziridinium ions **22a** and **22b** are shown in Fig. 6. The main focus is on the difference in reactivity between aziridine carbon atoms C2 and C3 in each aziridinium ion. The largest (positive) value of the Fukui function indicates the most reactive site, while the site where the LUMO is localized is a good electrophilic site, susceptible to nucleophilic attack. Both properties show that

Scheme 8 Ring expansion of aziridines, azetidines, and pyrrolidines **P** ($m = 1$, 2, and 3, respectively)



aziridinium ions **22a** and **22b** have a clearly opposite preference, i.e., nucleophilic attack at the substituted versus unsubstituted aziridine carbon atoms, respectively.

Although the electrostatic picture fails to explain the opposite regiospecific nature of the reaction, frontier molecular orbital analysis of LUMOs and nucleophilic Fukui functions show a clear preference of attack for the substituted aziridine carbon in the benzyl bromide case and for the unsubstituted aziridine carbon in the HBr case, successfully rationalizing the experimentally observed regioselectivity.

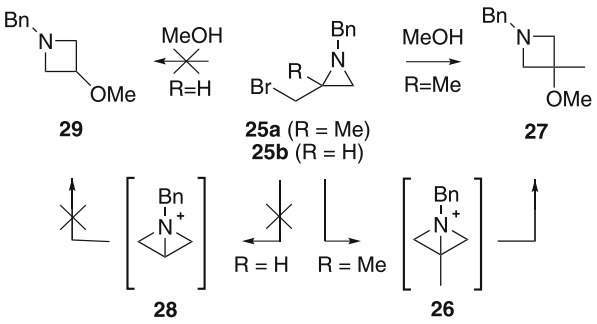
4 Bicyclic Aziridinium Intermediates in Ring-Expansion Reactions

In this section, the role of bicyclic aziridinium ions as intermediates in the ring expansion of aziridines, azetidines, and pyrrolidines to azetidines, pyrrolidines, and piperidines, respectively, will be unraveled. Nitrogen heterocycles different from aziridines, such as azetidines, pyrrolidines, and piperidines, can be found in numerous natural products and possess a wide range of biological activities [75–81]. Many efforts have been performed to find new synthetic routes toward such heterocycles and to control the stereochemistry [79, 82–94]. One method to prepare n-membered azacycles **R** is the ring expansion of (n-1)-membered azacycles **P** via 1-azabicyclo[m.1.0]alkane intermediates **Q** (Scheme 8).

The ring expansion of pyrrolidines **P** ($m = 3$) to piperidines **R** ($m = 3$) has been known for decades, and this reaction has been used as a key step in a number of syntheses [95–102]. However, the similar rearrangements of aziridines and azetidines **P** ($m = 1$ and 2, respectively) leading to azetidines and pyrrolidines **R** ($m = 1$ and 2, respectively) have only been reported in a few exceptional cases [93, 103–109]. The role of bicyclic aziridinium intermediates **Q** ($m = 1$) in the ring expansion of aziridines **P** ($m = 1$) to azetidines **R** ($m = 1$) will be investigated in detail.

4.1 Investigation of the Ring Expansion of Aziridines to Azetidines via Bicyclic Aziridinium Intermediates

DFT studies have been performed on the ring-expansion reactions of 2-(bromomethyl)aziridines to azetidines. The influence of different nucleophiles,



Scheme 9 Ring expansion of 1-benzyl-2-(bromomethyl)aziridines **10** in methanol

different solvents, and an additional methyl group at the 2-position is investigated to rationalize why the ring expansion takes place or does not take place depending on the conditions. Solvation and close-packing of the nucleophiles, the nucleofuge, and the bicyclic aziridinium intermediates have a profound influence on the reaction outcome.

Experimentally, the transformation of 2-bromomethyl-2-methylaziridines to functionalized aziridines and azetidines, and the influence of the choice of the solvent, has been investigated in the presence of different oxygen (OAc, OPh), carbon (CN), and sulfur (SCN) nucleophiles [110]. Dimethylformamide (DMF) was found to be a suitable solvent for the synthesis of the corresponding aziridines, while acetonitrile (MeCN) rather favors the formation of azetidines via strained bicyclic aziridinium intermediates. The reaction outcome provides an efficient strategy toward the selective synthesis of a large variety of functionalized aziridines and azetidines.

The presence of an extra substituent at the 2-position is crucial for a ring-expansion process, since 2-(bromomethyl)aziridines without an extra substituent at the 2-position are not susceptible to a ring-expansion process. All these experimental findings were tested by theoretical DFT calculations.

4.2 2-(Bromomethyl)aziridines With and Without an Extra Substituent at the 2-Position

Intramolecular cyclization and further transformation to azetidines is observed for 2-bromomethyl-2-methylaziridine **25a** and not for 2-(bromomethyl)aziridine **25b**, which lacks an additional methyl group at the 2-position (Scheme 9) [109]. Therefore, cyclization pathways for both aziridines are investigated by means of DFT calculations.

Transition-state geometries and free energies of activation for the formation of the bicyclic aziridinium intermediates **26** and **28** from aziridines **25** are shown in Fig. 7. Three explicit methanol molecules are used to solvate the bromide ion, which is expelled during the S_N2 reaction. The difference in free energy of

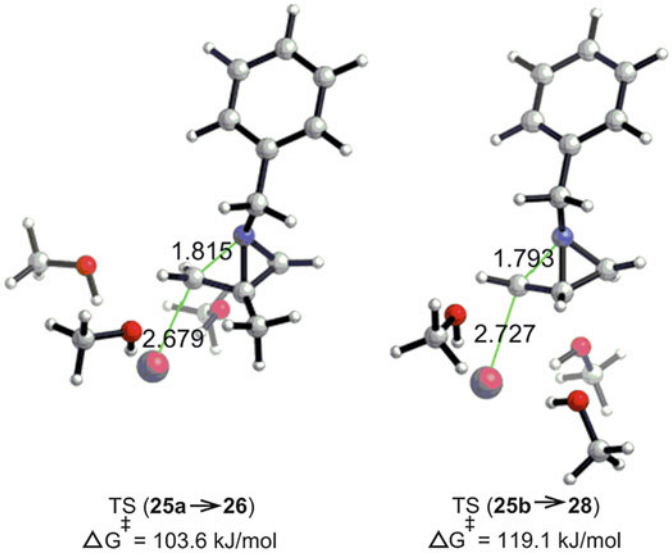


Fig. 7 Transition-state geometries and free energies of activation for the formation of the bicyclic aziridinium intermediates **26** and **28** from aziridines **25** (MPW1B95/6-31++G(d,p)//B3LYP/6-31++G(d,p), free energies in kJ/mol at 298 K, thermal free energy corrections were taken from B3LYP/6-31++G(d,p) optimizations, critical distances in Å, three explicit methanol molecules are used to solvate the bromide ion) [108]

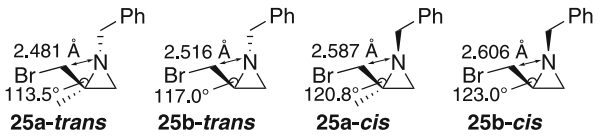


Fig. 8 Invertomers of **25a** and **25b**

activation can explain why 2-bromomethyl-2-methylaziridine **25a** undergoes cyclization and further transformation, in contrast with 2-(bromomethyl)aziridine **25b**, which lacks an additional substituent at the 2-position. This difference can be rationalized considering the Thorpe-Ingold effect due to *gem*-disubstitution at the 2-position of the aziridine, resulting in a more favorable geometry for nucleophilic attack. Replacement of the hydrogen atom at the 2-position of aziridines **25** by a methyl group decreases the distance between the nucleophilic nitrogen atom and the brominated carbon atom, which in turn gives rise to a higher reactivity (Fig. 8).

Since the bicyclic aziridinium ion **26** could be in equilibrium with its nonbridged carbenium ion counterpart **30**, the free energy profile for this equilibrium is shown in Fig. 9. Relative stabilities show that the bicyclic aziridinium ion is far more stable. This, together with the fact that the free energies of activation for the formation of the aziridinium ion **26** are reasonable, makes it very likely that the rearrangement of aziridine **25a** toward azetidine **27** takes place via the bicyclic species.

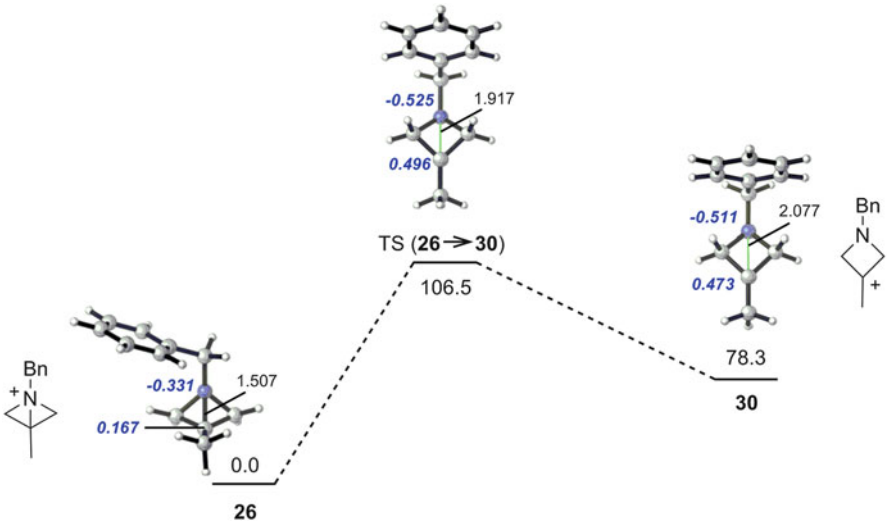
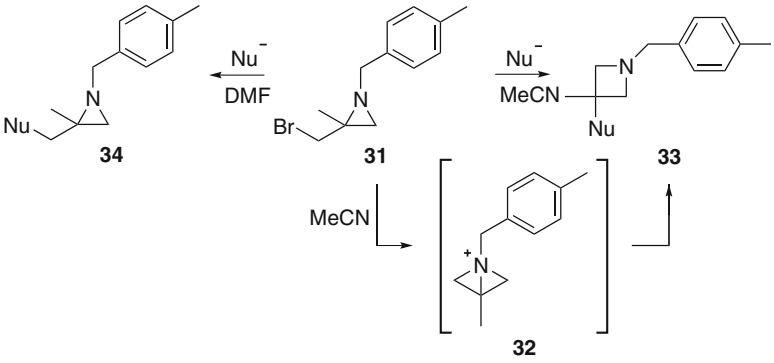


Fig. 9 Gibbs free energy profile for the equilibrium of bicyclic aziridinium ion **26** with its nonbridged carbenium ion counterpart **30** (CPCM ($\varepsilon = 32.6$) MPW1B95/6-31++G(d,p)//B3LYP/6-31++G(d,p), free energies in kJ/mol at 298 K, thermal free energy corrections were taken from B3LYP/6-31++G(d,p) optimizations, critical distances in Å, atomic NPA charges in italic) [108]



Scheme 10 Reactivity of 2-bromomethyl-2-methylaziridines **31** toward different nucleophiles in dimethylformamide and acetonitrile

4.3 Solvent-Dependent Reactivity of 2-Bromomethyl-2-methylaziridines Toward Different Nucleophiles

The choice of the solvent has a profound influence on the reactivity of 2-bromomethyl-2-methylaziridine **31** toward different nucleophiles, enabling the selective formation of either functionalized aziridines **34** in dimethylformamide or azetidines **33** in acetonitrile (Scheme 10) [110]. The difference in observed

selectivity in DMF and MeCN was unexpected, since both solvents are polar and aprotic and have very similar dielectric constants. However, they have different densities, molar volumes, shapes, and dipoles. The latter properties can influence the solvation of ions. Different close-packing of ions by DMF and MeCN can be crucial for their reactivities and was investigated thoroughly.

A stronger coordination and thus better stabilization of the nucleophiles by MeCN lies on the origin of the difference in the regioselectivity observed with DMF and MeCN as solvent. It leads to a higher probability for the formation of the bicyclic intermediate **32** and Br^- in MeCN. A nucleophilic attack at the more hindered carbon atom of the strained bicyclic intermediate **32** leads to the formation of the 3-substituted azetidines **33**. Following the literature, it is not excluded that in this MeCN-mediated reaction, a nucleophilic attack at the less-hindered carbon atom of the bicyclic aziridinium ion **32** can take place forming the aziridines **34**. But this path will probably not prevail. On the contrary, a direct nucleophilic displacement of bromide by the nucleophile is more probable as indeed observed in DMF mediated reactions (Scheme 10).

In the theoretical calculation, we keep only the carbon and sulfur nucleophiles CN^- and SCN^- into consideration and investigate their close-packing. The convergence behavior of the coordination solvation energy is inspected by means of systematically increasing the number of explicit solvent molecules around the ions.

Coordination of the nucleophile CN^- by DMF and MeCN is shown in Fig. 10. Coordination solvation energies converge after coordination with three DMF molecules and four MeCN molecules. Coordination with three explicit solvent molecules shows that CN^- is more strongly coordinated by MeCN than by DMF (CSG (Gibbs free coordination solvation energy) $= -56.3$ and -41.0 kJ/mol, respectively). Furthermore, MeCN molecules pack more closely than DMF molecules, as seen in the relative distances between the hydrogen atoms of the solvent and the CN^- nitrogen atom (2.14–2.17 and 2.40–2.67 Å, respectively) and the CN^- carbon atom (2.36–2.51 and 2.57–2.89 Å, respectively). The stronger coordination and better stabilization of CN^- by MeCN points to a lower reactivity, allowing the formation of azetine **33** via the bicyclic intermediate **32**. Since CN^- is less stabilized in DMF, it will be more reactive, and nucleophilic substitution in aziridine **31** will lead to aziridine **34**.

Coordination of the nucleophile SCN^- by DMF and MeCN is shown in Fig. 11. Coordination solvation energies converge after coordination with two DMF molecules and three MeCN molecules. Coordination with two explicit solvent molecules shows that, like CN^-, SCN^- is stronger coordinated by MeCN than by DMF (CSG $= -30.4$ and -13.5 kJ/mol, respectively). However, packing distances of MeCN and DMF molecules are not remarkably different. On the other hand, MeCN molecules pack more closely to CN^- than to SCN^- (distances between the hydrogen atoms of MeCN and the CN^- and SCN^- nitrogen atoms are 2.14–2.17 and 2.22–2.72 Å, respectively, and distances between the hydrogen atoms of MeCN and the CN^- carbon atom and the SCN^- sulfur atom are 2.36–2.51 and 2.93–3.07 Å, respectively). This, together with the stronger coordination of CN^- by MeCN than

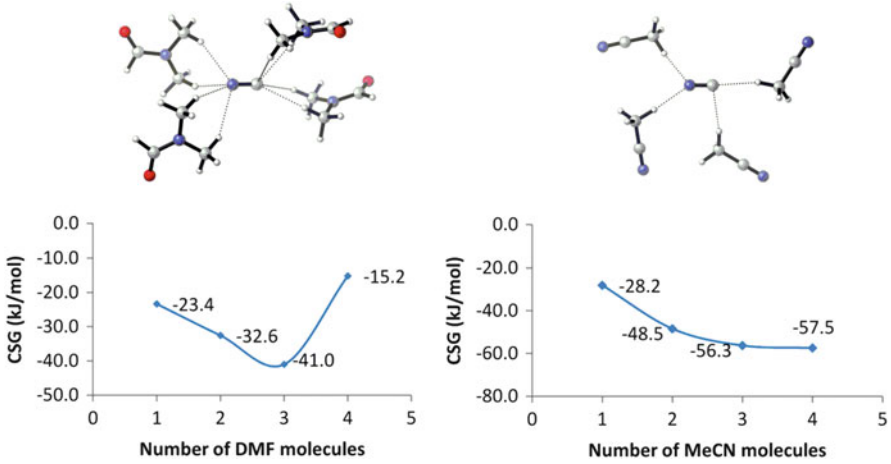


Fig. 10 Coordination of CN^- by an increasing number of DMF and MeCN molecules (MPW1K/6-31++G(d,p)//B3LYP/6-31 + G(d,p), free energies in kJ/mol at 298 K, thermal free energy corrections were taken from B3LYP/6-31 + G(d,p) optimizations) [109]

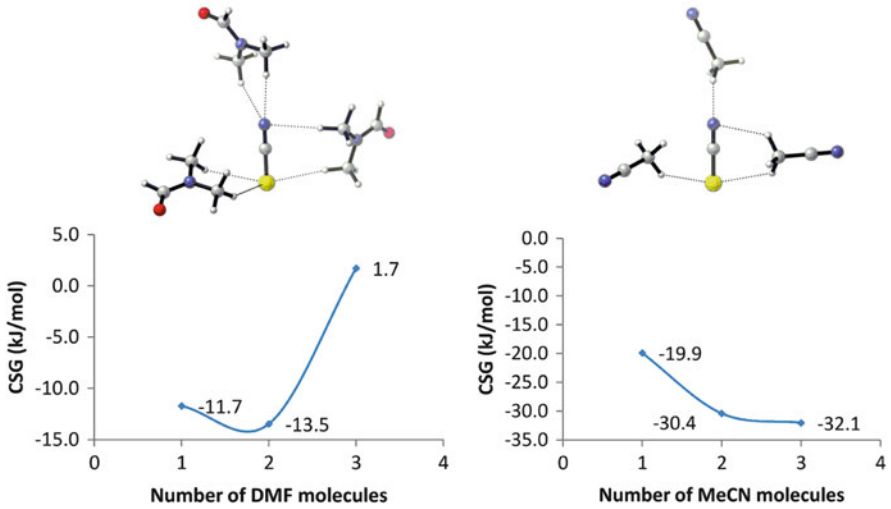
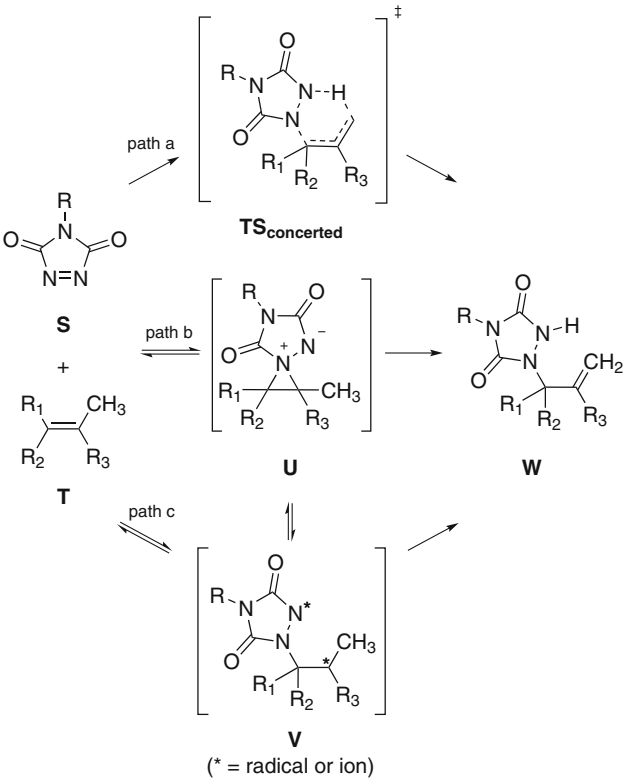


Fig. 11 Coordination of SCN^- by an increasing number of DMF and MeCN molecules (MPW1K/6-31++G(d,p)//B3LYP/6-31 + G(d,p), free energies in kJ/mol at 298 K, thermal free energy corrections were taken from B3LYP/6-31 + G(d,p) optimizations) [109]

SCN^- (CSG = −56.3 and −32.1 kJ/mol, respectively, for coordination with three explicit MeCN molecules), can explain the experimentally observed formation of both azetidine **33** and aziridine **34** in case of the nucleophile SCN^-.

Thus, the close-packing of CN^- by MeCN makes it less reactive, allowing the formation of azetidine **34** via the formation of the bicyclic intermediate **32**.



Scheme 11 Proposed mechanisms for the triazolinedione ene reaction

Furthermore, the weaker coordination of SCN^- by MeCN, compared to CN^-, can explain the formation of both azetidine **33** and aziridine **34** in case of the nucleophile SCN^- in MeCN.

5 Aziridinium Intermediates in Triazolinedione Ene Reactions

The 1,2,4-triazoline-3,5-dione ene reaction is another example where aziridinium ion-like species play an important role in the reaction mechanism. Triazolinediones (TAD) are electrophiles with high reactivity, particularly in ene reactions with alkenes [111–115]. The different mechanisms proposed for the ene reaction are shown in Scheme 11. Initially, triazolinedione ene reactions were believed to proceed via a concerted mechanism (path a). However, it has been demonstrated experimentally that the reaction of triazolinediones **S** and alkenes **T** follows a

stepwise route [116–119], via spiro aziridinium imide intermediates **U** or open intermediates **V**.

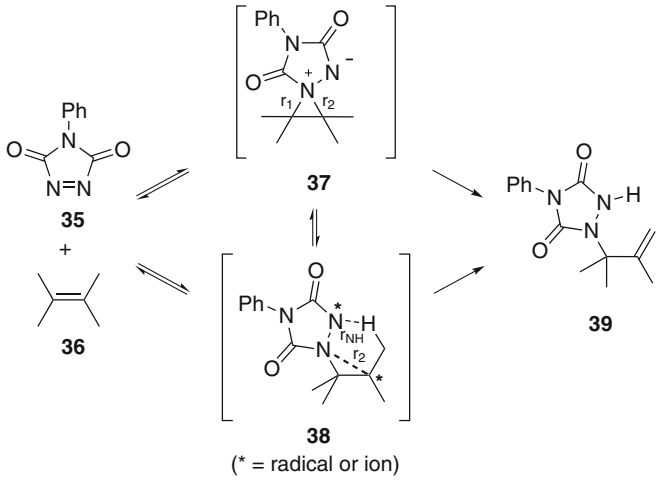
5.1 Experimental Section

The above-described triazolinedione ene reaction has attracted considerable mechanistic attention [111–119]. Most of the experimental studies support a stepwise mechanism with formation of intermediate aziridinium imides **U** in the rate-determining step [111–119], and the presence of the latter intermediates has been observed spectroscopically in a number of reports [120–122]. In addition, the regioselectivity of this synthetically useful transformation has been studied as well, pointing to the importance of the double bond substitution in that respect [111, 114, 123, 124]. For example, the ene reaction with unsymmetrically dialkyl-substituted *cis*-alkenes showed a general preference for hydrogen abstraction from the largest group of the double bond. These effects were proven to be more general, as in tri- and tetrasubstituted alkenes, a remarkable geminal selectivity was observed, leading to hydrogen abstraction from the group geminal to the largest substituent. The ene reaction with α,β-unsaturated ketones, esters, and lactones was described to be highly regioselective as α-alkylated, β-monoalkylated substrates selectively furnish new α,β-unsaturated carbonyl compounds resulting from hydrogen abstraction from the α-alkyl group. However, α-alkylated, β-dialkylated substrates give rise to hydrogen abstraction from one of the two β-alkyl groups, leading to the formation of the corresponding substituted β,γ-unsaturated compounds [125].

5.2 Theoretical Investigation

The role and stability of an aziridinium imide intermediate **37** and an open intermediate **38** in the ene reaction of 4-phenyl-1,2,4-triazoline-3,5-dione **35** and tetramethylethylene **36** (Scheme 12) have been investigated by means of QM/MM calculations (quantum mechanics/molecular mechanics) in protic and aprotic solvents by Acevedo et al. [23].

Free energy maps for the ene reactions in water and DMSO are shown in Fig. 12. Starting from semiempirical (PM3) gas-phase structures, the two reacting C–N bonds, r_1 and r_2 (Scheme 12), between triazolinedione **35** and alkene **36** are perturbed in order to find the aziridinium imide intermediate (AI) **37**, the open intermediate (OI) **38**, and the transition states for the initial addition of alkene **36** to triazolinedione **35** ($\mathbf{TS_{add}}$) and for the isomerization between the intermediates ($\mathbf{TS_{iso}}$). Furthermore, r_2 and r_{NH} (Scheme 12) are perturbed to locate the transition state for the allylic proton abstraction ($\mathbf{TS_{abs}}$), the intermediates **37** and **38**, and the ene product **39**. The geometries of the intermediates are identical in both free energy maps and are used to link the energy values between the maps.



Scheme 12 Ene reaction of 4-phenyl-1,2,4-triazoline-3,5-dione **35** and tetramethylethylene **36**

The resulting free energy profiles and relative free energies for the ene reactions in water, methanol, DMSO, and acetonitrile are shown in Fig. 13 and Table 5. The simulations predict a stepwise mechanism with three transition states and two intermediates. The initial addition of alkene **36** to triazolinedione **35** proceeds exclusively to an open intermediate **38**, and the transition-state geometries, TS_{add}, have very similar reacting bond lengths in all solvents ($r_1 = 1.78$–1.84 Å and $r_2 = 2.45$–2.50 Å). This step is found to be the rate-determining step and is faster in aprotic solvents than in protic solvents (58.2, 62.4, 72.9, and 93.4 kJ/mol in DMSO, MeCN, water, and MeOH, respectively). Direct formation of an aziridinium imide intermediate **37** is unlikely, since it would require an extra 40–60 kJ/mol (10–15 kcal/mol, Fig. 12). Instead, the aziridinium imide intermediate **37** is formed by closure of the open intermediate **38**. The activation barrier for the formation of **37** from **38** is significantly lower in aprotic solvents than in protic solvents (3.3, 5.0, 13.4, and 23.1 kJ/mol in DMSO, MeCN, MeOH, and water, respectively).

The transition-state geometries for the allylic proton abstraction, TS_{abs}, are significantly more solvent dependent than TS_{add}. An earlier transition state is found for the polar solvents ($r_2 = 2.33$–2.43 Å and $r_{NH} = 2.08$–2.11 Å) than for the apolar solvents ($r_2 = 1.99$–2.00 Å and $r_{NH} = 1.75$–1.77 Å). This difference results from different intermediate structures prior to TS_{abs} caused by the stabilization of charge separations over two and four atoms in the aziridinium imide intermediate **37** in aprotic solvents and the open intermediate **38** in protic solvents, respectively. Indeed, **38** is more stable in protic solvents than in aprotic solvents because of the stabilization by a hydrogen-bonding environment. Once the intermediates are formed, the reaction is expected to be faster in protic solvents than in aprotic solvents (34.2, 22.6, 77.0, and 74.9 kJ/mol in water, MeOH, DMSO, and MeCN, respectively).

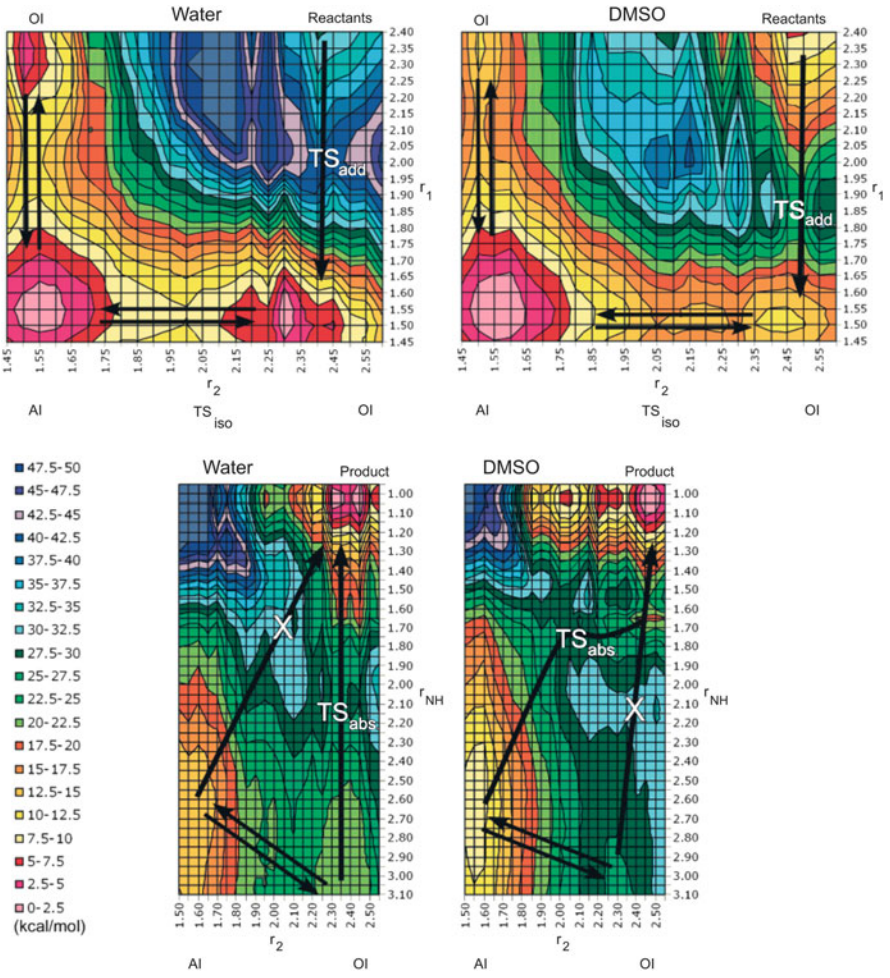


Fig. 12 Free energy maps for the ene reaction of 4-phenyl-1,2,4-triazoline-3,5-dione **35** and tetramethylethylene **36** in water (QM/MM (PM3/TIP4P)) and DMSO (QM/MM (PM3/OPLS)) (AI stands for aziridinium imide intermediate **37**, OI for open intermediate **38**, $\mathbf{TS_{add}}$ for the initial addition, $\mathbf{TS_{iso}}$ for the isomerization between the intermediates, and $\mathbf{TS_{abs}}$ for the allylic proton abstraction; distances in Å; maximum free energy values truncated to 50 kcal/mol for clarity; X denotes an energetically unfavorable reaction path) (Reprinted with permission from Acevedo et al. [23]; copyright 2008 American Chemical Society)

A remaining question is the nature of the open intermediate **38**. CM3 charges [126] of the intermediates were calculated. The computed charge separation between N1 and C4 of intermediate **38** is 1.26, 1.18, 1.22, and 1.21 in water, MeOH, DMSO, and MeCN, respectively, which is indicative of a dipolar, and not diradical, intermediate and is supported by experiments [127].

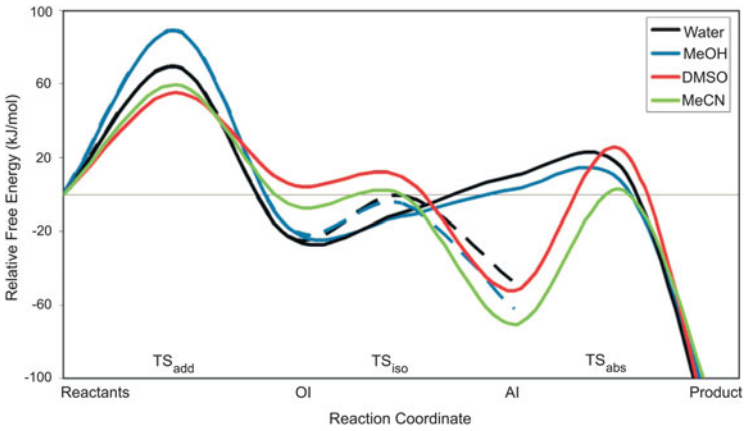


Fig. 13 Free energy profiles for the ene reaction of 4-phenyl-1,2,4-triazoline-3,5-dione **35** and tetramethylethylene **36** in water (QM/MM (PM3/TIP4P)), methanol, DMSO, and acetonitrile (QM/MM (PM3/OPLS)) (Al stands for aziridinium imide intermediate **37** and OI for open intermediate **38**) (Reprinted with permission from Acevedo et al. [23], copyright 2008 American Chemical Society)

Table 5 Relative free energies for the ene reaction of 4-phenyl-1,2,4-triazoline-3,5-dione **35** and tetramethylethylene **36** in water (QM/MM (PM3/TIP4P)), MeOH, DMSO, and MeCN (QM/MM (PM3/OPLS)) (free energies in kJ/mol at 298 K) [23]

	Water[a]	MeOH[a]	DMSO[b]	MeCN[b]
Reactants	0.0	0.0	0.0	0.0
TS_{add}	72.9	93.4	58.2	62.4
OI (**38**)	−23.9	−18.8	6.3	−5.4
TS_{iso}	−0.8	−5.4	9.6	−0.4
AI (**37**)	−49.4	−64.9	−54.4	−74.1
TS_{abs}	11.3	3.8	22.6	0.8
Product	−199.7	−172.9	−203.9	−165.8

[a]OI leads to product
[b]AI leads to product

In conclusion, the ene reaction of 4-phenyl-1,2,4-triazoline-3,5-dione **35** and tetramethylethylene **36** was found to proceed via a stepwise mechanism with the initial addition of alkene **36** to triazolinedione **35** with formation of an open dipolar intermediate **38** as rate-determining step. In protic solvents, the ene product **39** is formed from this open intermediate, but in aprotic solvents, the open intermediate is short-lived and the ene product is formed from an aziridinium imide intermediate **37**.

6 Conclusions

Aziridinium ions have been found to be important intermediates in the synthesis of functionalized amines. The strain of several aziridinium ions was studied and calculated theoretically, confirming the generally accepted statement that nonactivated aziridines have to be activated prior to ring-opening reactions and pointing to the conclusion that a precise protocol to quantify the ring strain of heterocyclic compounds is required. Furthermore, the factors that control the synthesis of functionalized amines via aziridinium salts were investigated. A combined experimental and theoretical approach makes it possible to better understand and predict the outcome of reactions for the synthesis of amines involving aziridinium intermediates. The *N*- and *C*-substituents of the aziridine ring, the nucleophile, and the solvent environment were found to be important in the formation and the regio- and stereocontrolled ring opening of aziridinium ions. In addition, bicyclic aziridinium ions and spiro aziridinium ions can be important intermediates in certain chemical reactions.

Furthermore, the choice of the solvation method for molecular modeling was proven to be a critical point in the calculations. More specifically, explicit solvent molecules are necessary to simulate reactions correctly if explicit solvent interactions are present. Alternatively, molecular dynamics simulations can be performed. Moreover, next to kinetics, reactivity indices and the close-packing of molecules can be useful to better understand the outcome of chemical reactions.

Acknowledgments This work was supported by the Research Foundation-Flanders (FWO-Vlaanderen), the Research Board of Ghent University (BOF-GOA), and the IAP-BELSPO program in the frame of IAP 7/05. Computational resources and services used in this work were provided by Ghent University.

References

1. Tanner D (1994) Chiral aziridines–their synthesis and use in stereoselective transformations. Angew Chem Int Ed Engl 33:599–619
2. Osborn HMI, Sweeney J (1997) The asymmetric synthesis of aziridines. Tetrahedron Asymmetry 8:1693–1715
3. Lindström UM, Somfai P (1998) Aminolysis of vinyl epoxides as an efficient entry to *N-H* vinylaziridines. Synthesis 1:109–117
4. McCoull WM, Davis FA (2000) Recent synthetic applications of chiral aziridines. Synthesis 10:1347–1365
5. Zwanenburg B, ten Holte P (2001) The synthetic potential of three-membered ring aza-heterocycles. Top Curr Chem 216:93–124
6. Sweeney JB (2002) Aziridines: epoxides' ugly cousins? Chem Soc Rev 31:247–258
7. Hu XE (2004) Nucleophilic ring opening of aziridines. Tetrahedron 60:2701–2743
8. Watson IDG, Yu LL, Yudin AK (2006) Advances in nitrogen transfer reactions involving aziridines. Acc Chem Res 39:194–206
9. Padwa A, Murphree SS (2006) Epoxides and aziridines–a mini review. ARKIVOC 3:6–33

10. Singh GS, D'hooghe M, De Kimpe N (2007) Synthesis and reactivity of *C*-heteroatom-substituted aziridines. Chem Rev 107:2080–2135
11. Tsang DS, Yang S, Alphonse FA et al (2008) Stereoselective isomerisation of *N*-allyl aziridines into geometrically stable *Z* enamines by using rhodium hydride catalysis. Chem-Eur J 14:886–894
12. De Kimpe N, Verhé R (1983) α-Halogenated imines. In: Patai S, Rappoport Z (eds) Halides, pseudo-halides and azides, vol 1. Wiley, Chichester
13. Helmut S (1999) Nucleophilic ring opening of aziridines. J Prakt Chem/Chem Ztg 341:319–331
14. Dahanukar VH, Zavialov LA (2002) Aziridines and aziridinium ions in the practical synthesis of pharmaceutical intermediates–a perspective. Curr Opin Drug Discov Dev 5:918–927
15. Schneider C (2009) Catalytic, enantioselective ring opening of aziridines. Angew Chem Int Ed 48:2082–2084
16. Roehn U, Becaud J, Mu L et al (2009) Nucleophilic ring opening of activated aziridines: a one-step method for labeling biomolecules with fluorine-18. J Fluor Chem 130:902–912
17. Sureshkumar D, Ganesh V, Vidyarini RS et al (2009) Direct synthesis of functionalized unsymmetrical β-sulfonamido disulfides by tetrathiomolybdate mediated aziridine ring-opening reactions. J Org Chem 74:7958–7961
18. Ottesen LK, Jaroszewski JW, Franzyk H (2010) Ring opening of a resin-bound chiral aziridine with phenol nucleophiles. J Org Chem 75:4983–4991
19. Bornholdt J, Felding J, Kristensen JL (2010) Synthesis of enantiopure 3-substituted morpholines. J Org Chem 75:7454–7457
20. Bornholdt J, Felding J, Clausen RP et al (2010) Ring opening of pymisyl-protected aziridines with organocuprates. Chem-Eur J 16:12474–12480
21. Stanković S, D'hooghe M, Catak S et al (2012) Regioselectivity in the ring opening of non-activated aziridines. Chem Soc Rev 41:643–665, and references cited herein
22. Métro T-X, Duthion B, Pardo DG et al (2010) Rearrangement of β-amino alcohols via aziridiniums: a review. Chem Soc Rev 39:89–102, and references cited herein
23. Acevedo O, Squillacote ME (2008) A new solvent-dependent mechanism for a triazolinedione ene reaction. J Org Chem 73:912–922
24. Wiberg KB (2004) Strain, structure, stability and reactivity. Found Chem 6:65–80
25. Wheeler SE, Houk KN, Schleyer PR et al (2009) A hierarchy of homodesmotic reactions for thermochemistry. J Am Chem Soc 131:2547–2560
26. Eliel EL, Wilen SH, Mander LN (1994) Stereochemistry of organic compounds. Wiley, New York
27. Baeyer A (1885) Ueber polyacetylenverbindungen. Ber Dtsch Chem Ges 18:2269–2281
28. Pitzer KS (1937) Thermodynamic functions for molecules having restricted internal rotations. J Chem Phys 5:469–472
29. Pitzer KS (1945) Strain energies of cyclic hydrocarbons. Science 101:672
30. Dunitz J, Schomaker V (1952) The molecular structure of cyclobutane. J Chem Phys 20:1703–1707
31. Westheimer FA (1956) Calculation of the magnitude of steric effects. In: Newman MS (ed) Steric effects in organic chemistry. Wiley, New York
32. Lewis LL, Turner LL, Salter EA et al (2002) Computation of the conventional strain energy in oxaziridine. J Mol Struct-Theochem 592:161–171
33. Wiberg K, Bader R, Lau C (1987) Theoretical analysis of hydrocarbon properties. 2. Additivity of group properties and the origin of strain energy. J Am Chem Soc 109:1001–1012
34. Dewar M (1984) Chemical implications of sigma-conjugation. J Am Chem Soc 106:669–682
35. Orchin M, Macomber RS, Pinhas AR et al (2005) The vocabulary and concepts of organic chemistry, 2nd edn. Wiley, Hoboken
36. Cox J (1963) A bond energy scheme–II: strain and conjugation energies in cyclic compounds. Tetrahedron 19:1175–1184

37. Skancke A, Van Vechten D, Liebman JF et al (1996) Strain energy of three-membered rings: a new ultradiagonal definition as applied to silicon- and carbon-containing species. J Mol Struct 376:461–468
38. Bach RD, Dmitrenko O (2006) The effect of carbonyl substitution on the strain energy of small ring compounds and their six-member ring reference compounds. J Am Chem Soc 128:4598–4611
39. Gimarc BM, Zhao M (1997) Strain and resonance energies in main-group homoatomic rings and clusters. Coord Chem Rev 158:385–412
40. Dudev T, Lim C (1998) Ring strain energies from ab initio calculations. J Am Chem Soc 120:4450–4458
41. Magers DH, Davis SR (1999) Ring strain in the oxazetidines. J Mol Struct-Theochem 487:205–210
42. Verevkin SP, Emel'yanenko VN, Pimerzin AA et al (2011) Thermodynamic analysis of strain in the five-membered oxygen and nitrogen heterocyclic compounds. J Phys Chem A 115:1992–2004
43. Smith SA, Hand KE, Love ML et al (2013) Conventional strain energies of azetidine and phosphetane: can density functional theory yield reliable results? J Comput Chem 34:558–565
44. Kenis S, D'hooghe M, Verniest G et al (2011) Straightforward synthesis of 1-alkyl-2-(trifluoromethyl)aziridines starting from 1,1,1-trifluoroacetone. Org Biomol Chem 9:7217–7223
45. D'hooghe M, Van Speybroeck V, Waroquier M et al (2006) Regio- and stereospecific ring opening of 1,1-dialkyl-2-(aryloxymethyl)aziridinium salts by bromide. Chem Commun 1554–1556
46. D'hooghe M, Van Speybroeck V, Van Nieuwenhove A et al (2007) Novel synthesis of 3,4-diaminobutanenitriles and 4-amino-2-butenenitriles from 2-(cyanomethyl)aziridines through intermediate aziridinium salts: an experimental and theoretical approach. J Org Chem 72:4733–4740
47. Yun SY, Catak S, Lee WK et al (2009) Nucleophile-dependent regioselective ring opening of 2-substituted *N*,*N*-dibenzylaziridinium ions: bromide versus hydride. Chem Commun 2508–2510
48. Catak S, D'hooghe M, De Kimpe N et al (2010) Intramolecular pi-pi stacking interactions in 2-substituted *N*,*N*-dibenzylaziridinium ions and their regioselectivity in nucleophilic ring-opening reactions. J Org Chem 75:885–896
49. Catak S, D'hooghe M, Verstraelen T et al (2010) Opposite regiospecific ring opening of 2-(cyanomethyl)aziridines by hydrogen bromide and benzyl bromide: experimental study and theoretical rationalization. J Org Chem 75:4530–4541
50. D'hooghe M, Catak S, Stanković S et al (2010) Systematic study of halide-induced ring opening of 2-substituted aziridinium salts and theoretical rationalization of the reaction pathways. Eur J Org Chem 2010:4920–4931
51. Cramer CJ, Truhlar DG (1996) Continuum solvation models. In: Tapia O, Bertrán J (eds) Solvent effects and chemical reactivity. Kluwer, Dordrecht
52. Barone V, Cossi M (1998) Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. J Phys Chem A 102:1995–2001
53. Cossi M, Rega N, Scalmani G et al (2003) Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. J Comput Chem 24:669–681
54. Takano Y, Houk KN (2005) Benchmarking the conductor-like polarizable continuum model (CPCM) for aqueous solvation free energies of neutral and ionic organic molecules. J Chem Theory Comput 1:70–77
55. Van Speybroeck V, Moonen K, Hemelsoet K et al (2006) Unexpected four-membered over six-membered ring formation during the synthesis of azaheterocyclic phosphonates: experimental and theoretical evaluation. J Am Chem Soc 128:8468–8478

56. Catak S, Monard G, Aviyente V et al (2006) Reaction mechanism of deamidation of asparaginyl residues in peptides: effect of solvent molecules. J Phys Chem A 110:8354–8365
57. Catak S, Monard G, Aviyente V et al (2008) Computational study on nonenzymatic peptide bond cleavage at asparagine and aspartic acid. J Phys Chem A 112:8752–8761
58. Catak S, Monard G, Aviyente V et al (2009) Deamidation of asparagine residues: direct hydrolysis versus succinimide-mediated deamidation mechanisms. J Phys Chem A 113:1111–1120
59. Hermosilla L, Catak S, Van Speybroeck V et al (2010) Kinetic and mechanistic study on p-quinodimethane formation in the sulfinyl precursor route for the polymerization of poly (p-phenylenevinylene) (PPV). Macromolecules 43:7424–7433
60. Dedeoglu B, Catak S, Houk KN et al (2010) A theoretical study of the mechanism of the desymmetrization of cyclic meso-anhydrides by chiral amino alcohols. ChemCatChem 2:1122–1129
61. Pliego JR, Riveros JM (2001) The cluster-continuum model for the calculation of the solvation free energy of ionic species. J Phys Chem A 105:7241–7247
62. Kelly CP, Cramer CJ, Truhlar DG (2006) Adding explicit solvent molecules to continuum solvent calculations for the calculation of aqueous acid dissociation constants. J Phys Chem A 110:2493–2499
63. da Silva EF, Svendsen HF, Merz KM (2009) Explicitly representing the solvation shell in continuum solvent calculations. J Phys Chem A 113:6404–6409
64. Kamerlin SCL, Haranczyk M, Warshel A (2009) Are mixed explicit/implicit solvation models reliable for studying phosphate hydrolysis? A comparative study of continuum, explicit and mixed solvation models. ChemPhysChem 10:1125–1134
65. Mora JR, Tosta M, Domínguez RM et al (2007) Joint theoretical and experimental study of the gas-phase elimination kinetics of tert-butyl ester of carbamic, *N*,*N*-dimethylcarbamic, *N*-hydroxycarbamic acids and 1-(tert-butoxycarbonyl)-imidazole. J Phys Org Chem 20:1021–1031
66. Kaur D, Kohli R (2008) Intra and intermolecular hydrogen bonding in formohydroxamic acid. Int J Quantum Chem 108:119–134
67. Hedegard ED, Bendix J, Sauer SPA (2009) Partial charges as reactivity descriptors for nitrido complexes. J Mol Struct-Theochem 913:1–7
68. Parks JM, Guo H, Momany C et al (2009) Mechanism of Hg-C protonolysis in the organomercurial lyase MerB. J Am Chem Soc 131:13278–13285
69. Schwöbel J, Ebert R-U, Kühne R et al (2009) Modeling the H-bond donor strength of –OH, –NH, and –CH sites by local molecular parameters. J Comput Chem 30:1454–1464
70. Cheshmedzhieva D, Ilieva S, Hadjieva B et al (2009) Reactivity of acetanilides in the alkaline hydrolysis reaction: theory vs. experiment. Mol Phys 107:1187–1192
71. Reed AE, Weinstock RB, Weinhold F (1985) Natural population analysis. J Chem Phys 83:735–746
72. Bultinck P, Ayers PW, Fias S et al (2007) Uniqueness and basis set dependence of iterative Hirshfeld charges. Chem Phys Lett 444:205–208
73. Parr RG, Yang W (1984) Density functional approach to the frontier-electron theory of chemical reactivity. J Am Chem Soc 106:4049–4050
74. Ayers PW, Levy M (2000) Perspective on "density functional approach to the frontier-electron theory of chemical reactivity". Theor Chem Acc 103:353–360
75. Cromwell NH, Phillips B (1979) The azetidines. Recent synthetic developments. Chem Rev 79:331–358
76. Moore JA, Ayers RS (1983) Azetidines. In: Hassner A (ed) Chemistry of heterocyclic compounds: small ring heterocycles, part 2. Wiley, New York
77. Davies DE, Storr RC (1984) Azetidines, azetines and azetes. In: Lwowski W (ed) Comprehensive heterocyclic chemistry, vol 7. Pergamon, Oxford
78. De Kimpe N (1996) Azetidines, azetines and azetes: monocyclic. In: Padwa A (ed) Comprehensive heterocyclic chemistry II, vol 1. Elsevier, Oxford

79. Bellina F, Rossi R (2006) Synthesis and biological activity of pyrrole, pyrroline and pyrrolidine derivatives with two aryl groups on adjacent positions. Tetrahedron 62:7213–7256
80. Fraser HL, Floyd MB, Hopper DW (2007) Six-membered ring systems: pyridines and benzo derivates. In: Gribble GW, Joule JA (eds) Progress in heterocyclic chemistry, vol 18. Elsevier, Oxford
81. Källström S, Leino R (2008) Synthesis of pharmaceutically active compounds containing a disubstituted piperidine framework. Bioorg Med Chem 16:601–635
82. Felpin FX, Lebreton J (2003) Recent advances in the total synthesis of piperidine and pyrrolidine natural alkaloids with ring-closing metathesis as a key step. Eur J Org Chem 2003:3693–3712
83. Pandey G, Banerjee P, Gadre SR (2006) Construction of enantiopure pyrrolidine ring system via asymmetric [3 + 2]-cycloaddition of azomethine ylides. Chem Rev 106:4484–4517
84. Minatti A, Muniz K (2007) Intramolecular aminopalladation of alkenes as a key step to pyrrolidines and related heterocycles. Chem Soc Rev 36:1142–1152
85. Wolfe JP (2007) Palladium-catalyzed carboetherification and carboamination reactions of gamma-hydroxy- and gamma-aminoalkenes for the synthesis of tetrahydrofurans and pyrrolidines. Eur J Org Chem 2007:571–582
86. Schomaker JM, Bhattacharjee S, Yan J et al (2007) Diastereomerically and enantiomerically pure 2,3-disubstituted pyrrolidines from 2,3-aziridin-1-ols using a sulfoxonium ylide: a one-carbon homologative relay ring expansion. J Am Chem Soc 129:1996–2003
87. Sherman ES, Fuller PH, Kasi D et al (2007) Pyrrolidine and piperidine formation via copper (II) carboxylate-promoted intramolecular carboamination of unactivated olefins: diastereoselectivity and mechanism. J Org Chem 72:3896–3905
88. Sulmon P, De Kimpe N, Schamp N (1985) A novel synthesis of 2-cyano-3,3-dimethylazetidines. Chem Commun 715–716
89. Sulmon P, De Kimpe N, Schamp N et al (1988) Synthesis of azetidines from beta-chloro imines. Tetrahedron 44:3653–3670
90. De Kimpe N, De Smaele D (1995) A convenient synthesis of 3-alkoxyazetidines. Tetrahedron 51:5465–5478
91. Dejaegher Y, Kuzmenok N, Zvonok A et al (2002) The chemistry of azetidin-3-ones, oxetan-3-ones and thietan-3-ones. Chem Rev 102:29–60
92. Van Brabandt W, Dejaegher Y, Van Landeghem R et al (2006) Reduction of 4-(haloalkyl) azetidin-2-ones with $LiAlH_4$ as a powerful method for the synthesis of stereodefined aziridines and azetidines. Org Lett 8:1101–1104
93. Van Brabandt W, Van Landeghem R, De Kimpe N (2006) Ring transformation of 2-(haloalkyl)azetidines into 3,4-disubstituted pyrrolidines and piperidines. Org Lett 8:1105–1108
94. Van Brabandt W, Mangelinckx S, D'hooghe M et al (2009) Synthesis and reactivity of 3-haloazetidines and 3-sulfonyloxyazetidines. Curr Org Chem 13:829–853
95. Fuson RC, Zirkle CL (1948) Ring enlargement by rearrangement of the 1,2-aminochloroalkyl group; rearrangement of 1-ethyl-2-chloromethylpyrrolidine to 1-ethyl-3-chloropiperidine. J Am Chem Soc 70:2760–2762
96. Wilkem J, Kossenjans M, Saak W et al (1997) Synthesis of new chiral bicyclic 3-hydroxypiperidines – highly diastereoselective ring expansion of the azabicyclo[3.3.0] octane system to chiral piperidine derivatives. Liebigs Ann/Recueil 1997:573–579
97. Cossy J, Dumas C, Pardo DG (1997) Synthesis of (-)-pseudoconhydrine through ring enlargement of a L-proline derivative. Synlett 1997:905–906
98. Cossy J, Dumas C, Pardo DG (1997) A short and efficient synthesis of zamifenacin a muscarinic M-3 receptor antagonist. Bioorg Med Chem Lett 7:1343–1344
99. Calvez O, Chiaroni A, Langlois N (1998) Enantioselective synthesis of 2,3-disubstituted piperidines from (S)-methylpyroglutamate. Tetrahedron Lett 39:9447–9450

100. Michel P, Rassat A (2000) An easy access to 2,6-dihydroxy-9-azabicyclo[3.3.1]nonane, a versatile synthon. J Org Chem 65:2572–2573
101. Cossy J, Mirguet O, Pardo DG (2001) Ring expansion: synthesis of the velbanamine piperidine core. Synlett 2001:1575–1577
102. Mena M, Bonjoch J, Pardo DG et al (2006) Ring expansion of functionalized octahydroindoles to enantiopure cis-decahydroquinolines. J Org Chem 71:5930–5935
103. Gaertner VR (1970) Reactions of nucleophiles with 1-tertiary-butyl-3-chloroazetidine and 1-tertiary-butyl-2-chloromethylaziridine. J Org Chem 35:3952–3959
104. Outurquin F, Pannecoucke X, Berthe B et al (2002) Stereocontrolled synthesis of 1,2-dialkyl-4-halopyrrolidines through PhSeX-induced cyclization of secondary homoallylamines. Eur J Org Chem 2002:1007–1014
105. Couty F, Durrat F, Prim D (2003) Highly stereoselective ring expansion of enantiopure alpha-hydroxyalkyl azetidines. Tetrahedron Lett 44:5209–5212
106. Durrat F, Sanchez MV, Couty F et al (2008) Ring expansion of 2-(alpha-hydroxyalkyl) azetidines: a synthetic route to functionalized pyrrolidines. Eur J Org Chem 2008:3286–3297
107. Mangelinckx S, Žukauskaitė A, Buinauskaitė V et al (2008) Synthesis of alkyl 2-(bromomethyl)aziridine-2-carboxylates and alkyl 3-bromoazetidine-3-carboxylates as amino acid building blocks. Tetrahedron Lett 49:6896–6900
108. Žukauskaitė A, Mangelinckx S, Buinauskaitė V et al (2011) Synthesis of new functionalized aziridine-2-and azetidine-3-carboxylic acid derivatives of potential interest for biological and foldameric applications. Amino Acids 41:541–558
109. Stanković S, Catak S, D'hooghe M et al (2011) Synthesis of 3-methoxyazetidines via an aziridine to azetidine rearrangement and theoretical rationalization of the reaction mechanism. J Org Chem 76:2157–2167
110. Stanković S, Goossens H, Catak S et al (2012) Solvent-controlled selective transformation of 2-bromomethyl-2-methylaziridines to functionalized aziridines and azetidines. J Org Chem 77:3181–3190
111. Cheng CC, Seymour CA, Petti MA (1984) Reaction of electrophiles with unsaturated systems – triazolinedione-olefin reactions. J Org Chem 49:2910–2916
112. Orfanopoulos M, Foote CS, Smonou I (1987) Stereochemical dependence of isotope effects in the ene reaction of *N*-phenyl-1,2,4-triazoline-3,5-dione with isomers of butene-d_3. Tetrahedron Lett 28:15–18
113. Elemes Y, Stratakis M, Orfanopoulos M (1989) Reactions of triazolinediones with *cis*-alkenes. A highly regioselective ene reaction. Tetrahedron Lett 30:6903–6906
114. Orfanopoulos M, Elemes Y, Stratakis M (1990) Reactions of triazolinedione with alkenes. A remarkable geminal selectivity. Tetrahedron Lett 31:5775–5778
115. Chen JS, Houk KN, Foote CS (1997) The nature of the transition structures of triazolinedione ene reactions. J Am Chem Soc 119:9852–9855
116. Seymour CA, Greene FD (1980) Mechanism of triazolinedione-olefin reactions – ene and cycloaddition. J Am Chem Soc 102:6384–6385
117. Smonou I, Orfanopoulos M, Foote CS (1988) Reaction of 4-phenyl-1,2,4-triazoline-3,5-dione with substituted indenes. Tetrahedron Lett 29:2769–2772
118. Orfanopoulos M, Smonou I, Foote CS (1990) Intermediates in the ene reactions of singlet oxygen and *N*-phenyl-1,2,4-triazoline-3,5-dione with olefins. J Am Chem Soc 112:3607–3614
119. Smonou I, Khan S, Foote CS et al (1995) Reactions of phenyltriazolinedione with alkenes – stereochemistry of methanol adducts to aziridinium imide intermediates. J Am Chem Soc 117:7081–7087
120. Nelsen SF, Kapp DL (1985) Direct observation of the aziridinium imide intermediates in the reaction of biadamantylidene with triazolinediones. J Am Chem Soc 107:5548–5549
121. Squillacote M, Mooney M, De Felippis J (1990) An aziridinium imide intermediate in the ene reaction of trans-cycloheptene and *N*-methyl-1,2,4-triazoline-3,5-dione. J Am Chem Soc 112:5364–5365

122. Poon THW, Park SH, Elemes Y et al (1995) Reaction of *N*-substituted 1,2,4-triazoline-3,5-diones and trans-cyclooctene. Direct observation of an aziridinium imide. J Am Chem Soc 117:10468–10473
123. Orfanopoulos M, Stratakis M, Elemes Y et al (1991) Do rotational barriers dictate the regioselectivity in the ene reactions of singlet oxygen and triazolinedione with alkenes? J Am Chem Soc 113:3180–3181
124. Elemes Y, Stratakis M, Orfanopoulos M (1997) An interplay in the regioselectivity induced by non bonding interactions, in the ene reactions of singlet oxygen and triazolinediones with tetrasubstituted alkenes. Tetrahedron Lett 38:6437–6440
125. Hoye TR, Bottorff KJ, Caruso AJ et al (1980) Regio- and stereoselectivity in the ene reaction of *N*-phenyl-1,2,4-triazoline-3,5-dione with α, β-unsaturated carbonyl substrates. J Org Chem 45:4287–4292
126. Winget P, Thompson JD, Xidos JD et al (2002) Charge model 3: a class IV charge model based on hybrid density functional theory with variable exchange. J Phys Chem A 106:10707–10717
127. Roubelakis MM, Vougioukalakis GC, Angelis YS et al (2006) Solvent-dependent changes in the ene reaction of RTAD with alkenes: the cyclopropyl group as a mechanistic probe. Org Lett 8:39–42

Characterization of the Chemical Reactivity and Selectivity of DNA Bases Through the Use of DFT-Based Descriptors

Vanessa Labet, Christophe Morell, Vincent Tognetti, Olga A. Syzgantseva, Laurent Joubert, Nelly Jorge, André Grand, and Jean Cadet

Contents

V. Labet
Sorbonne Universités, UPMC Univ Paris 6, MONARIS, UMR 8233, 75005 Paris, France

C. Morell (✉)
Université de Lyon 1, ISA, UMR 5280, 5 rue de la Doua, 69100 Villeurbanne, France
e-mail: Christophe.morell@univ-lyon1.fr

V. Tognetti • O.A. Syzgantseva • L. Joubert
Normandie Université, COBRA, UMR6014 & FR 3038, Université de Rouen, INSA Rouen, CNRS, 1 rue Tesnière, 76821 Mont-Saint-Aignan Cedex, France

N. Jorge
Área Fisicoquímica, Facultad de Ciencias Exactas y Naturales y Agrimensura, UNNE, Campus Universitario, Av. Libertad 5400, (3400), Corrientes, Argentina

A. Grand • J. Cadet
CEA Grenoble -Institut Nanosciences et Cryogénie/SCIB/LAN (UMR-E n 3 CEA-UJF), CEA-Grenoble, 17, rue des Martyrs, 38054 Grenoble Cedex 9, France

F. De Proft, P. Geerlings (eds.), *Structure, Bonding and Reactivity of Heterocyclic Compounds*, Topics in Heterocyclic Chemistry 38, DOI 10.1007/978-3-642-45149-2_2

Abstract In this chapter, the use of conceptual DFT descriptors for understanding the occurrence and likely mechanisms of formation of DNA lesions is reviewed. After a synthetic presentation of the principal DFT-based descriptors, the global reactivity and selectivity of DNA bases are investigated from global and local descriptors. Then, the formation of several DNA lesions is studied including cytosine compound deamination, intra-strand DNA cross-links, and pyrimidine dimer photoproducts. It appears from the use of the global and local DFT-based descriptors that most of the experimental facts can be theoretically rationalized.

Keywords Descriptors of chemical reactivity • Conceptual DFT • DNA damage • Cytosine deamination • Cyclobutane pyrimidine dimers • Pyrimidine (6–4) pyrimidone photoproducts • Tandem base lesions • Purine 5′,8-cyclonucleosides • Dual descriptors

1 Introduction

Deoxyribonucleic acid (DNA) in its cellular environment is prone to a wide range of degradation pathways involving hydrolysis and reactions with various chemicals and physical agents [1, 2]. One of the two main hydrolytic reactions involving DNA leads to the cleavage of the *N*-glycosidic bond of 2′-deoxyribonucleosides, preferentially at guanine (Gua) and to a lesser extent at adenine (Ade) sites with subsequent formation of highly mutagenic abasic sites [3]. Purine bases are also subject to deamination as well as cytosine (Cyt) and 5-methylcytosine (5-MeCyt) [1, 2], the fifth nucleobase [4] that has recently received major attention due to its strong implication in epigenetic pathways [5] through enzymatic oxidation of the 5-methyl group of the pyrimidine ring [6, 7].

Early evidence has shown that the susceptibility of the nucleobases to hydrolytic deamination decreases for model compounds in the following order: Cyt > 5-MeCyt > purine bases [8–10]. It is worth noting that uracil (Ura) and thymine (Thy), the two main bases thus hydrolytically generated, are efficiently removed from DNA by uracil and thymine DNA glycosylases respectively, two mismatch repair enzymes [11, 12]. Other important endogenous reactions to cellular DNA are mediated by reactive oxygen species (ROS), actually mostly by hydroxyl radical ($^{\bullet}$OH) that efficiently reacts indistinctively with neighboring molecules at the site where it is generated [13].

The initial event of the pathway leading to cellular DNA oxidation is monoelectronic reduction of dioxygen (O_2) during mitochondrial respiration [14], inflammation [15], or phagocytosis [16] giving rise to an unreactive superoxide anion radical ($O_2^{\bullet-}$) [13]. Dismutation of $O_2^{\bullet-}$ leads to low reactive H_2O_2 that is able to migrate through the cells before eventually being reduced by transition metals such as Fe^{2+} ions and converted into highly reactive $^{\bullet}$OH. A large body of

information is now available on the reaction of $^{\bullet}OH$ with nucleobases [17, 18] and 2-deoxyribose [17, 19, 20] in terms of reactivity, mechanisms involved, and oxidation products. Thus more than 100 oxidatively generated modifications including single and tandem base lesions together with oxidized sugar residues have been characterized in model compounds [17–20]. It may be added that almost 20 modified nucleosides have been accurately measured in cellular DNA, mostly by using high-performance liquid chromatography (HPLC) coupled with either tandem mass spectrometry (MS/MS) or MS^3 experiments [21].

Carbonate anion radical ($CO_3^{\bullet-}$), another powerful oxidant, may be generated during inflammation through, in the initial step, recombination of $O_2^{\bullet-}$ and nitrite oxide ($^{\bullet}NO$), the product of NO synthase [22]. Subsequent reaction of peroxynitrite thus formed with CO_2 gives rise to unstable nitrosoperoxycarbonate that decomposes with simultaneous release of $CO_3^{\bullet-}$ and nitrating $^{\bullet}NO_2$ [23, 24]. $CO_3^{\bullet-}$ is able to selectively one-electron-oxidize guanine [25], producing a related radical cation that efficiently reacts with nucleophiles generating a variety of damage including single lesions, intra- and inter-strand DNA cross-links, and DNA-protein adducts [19, 26–28].

Numerous exogenous agents are also able to oxidize DNA either directly or after metabolic activation. It is well documented that ionizing radiation generates $^{\bullet}OH$ through radiolysis of water molecules and ionize DNA components via indirect and direct effects, respectively [17]. Other relevant oxidizing agents include endogenous and exogenous photosensitizers that may act as one-electron oxidants (type I photosensitization mechanism) and/or generators of singlet oxygen (type II photosensitization mechanism) [29, 30], another important ROS that selectively reacts with guanine [27]. The main important genotoxic effects of solar radiation on human skin are explained in terms of direct excitation of nucleobases by the UV components including mostly the more energetic UVB photons and to a lesser extent UVA radiation [31, 32]. An excited pyrimidine base is able to react with vicinal thymine or cytosine via two competitive photoreactions whose efficiency is strongly primary sequence dependent. The formation of *cis-syn* cyclobutane pyrimidine dimers (CPDs) involves a [2 + 2] photocycloaddition reaction, whereas competitive Paternò-Büchi photocycloaddition gives rise to pyrimidine (6–4) pyrimidone photoproducts (6–4PPs) through either a dioxetane or an azetidine intermediate [32]. Photoisomerization of the UVB-induced 6-4PPs into related Dewar valence isomers has been shown to imply mostly UVA radiation in cellular DNA [33].

The base modifications generated to cellular DNA by deamination, oxidants, or photochemical reactions are usually removed from double-stranded DNA either by base excision repair for single lesions [34] or nucleotide excision repair for bulky DNA damage such as CPDs and 6-4PPs [35–37]. However, if DNA replication occurs before repair of the modifications, there is a risk of mutation induction in the newly synthesized DNA strand when the damage is miscoding [38]. This may lead to deleterious effects including long-term occurrence of carcinogenesis as the result of incorrect processing of cytosine-containing CPDs that are likely to undergo deamination [37]. This explains why numerous and extensive investigations have

been devoted to the characterization, measurement, and biological assessment (repair, mutagenic potential) of DNA lesions.

More recently, major efforts have been made to complement the above information by using theoretical approaches for gaining further insights into the chemical reactivity of the bases. In the present chapter, emphasis is placed on the critical review of information gained from the use of density functional theory-based descriptors that shed new light on the endogenous and exogenous formation of several single and tandem lesions. Deamination of unsaturated and 5,6-saturated cytosine compounds that involves tautomerization reaction is first considered. In addition three main types of DNA modifications are reviewed. Two involve radical reactions initiated by $^{\bullet}$OH that lead to the formation of tandem base lesions between two adjacent thymine-purine bases [19, 39] and cross-links between the 2-deoxyribose from the intramolecular cyclization of C5′ radical to either adenine or guanine at C8 [40, 41]. Lastly the relevance of DFT descriptors for investigating the reactivity of excited pyrimidine bases has been challenged through consideration of the relative formation of CPDs and 6-4PPs, the main UVB-induced bipyrimidine photoproducts [32].

2 Theoretical Framework

It is a chemical fact that some molecules are more stable and others more reactive. It is also often observed that some atomic positions within molecules are more prone to react than their neighbors. Different chemical theories have the ambition to rationalize these experimental facts. The most fruitful and elegant framework so far is certainly the density functional theory of chemical reactivity, also called conceptual DFT [42, 43]. Conceptual DFT is a subfield of DFT in which one tries to extract from the electron density relevant concepts and principles that help us to understand and predict the chemical behavior of a molecule. The aim of this section is to briefly present the chemical descriptors used in the chapter and the way they are usually computed.

First an overall picture of the theoretical framework is depicted and discussed. Section 2.2 gives an introduction to global indexes such as chemical potential and chemical hardness. In Sect. 2.3 the local counterparts of the global descriptors are presented. The last section is dedicated to the extension of the local reactivity descriptor, called dual descriptor, to photoexcited biomolecules.

2.1 Conceptual DFT Framework

Density functional theory is original in the sense that within the DFT framework, a molecule is considered as an electron density stabilized and shaped by the electric field created by nuclei. In other words, in DFT a molecule is seen as a whole and not

as a collection of atoms. This way of considering a molecule simplifies the modeling of chemical processes. During a reaction, the molecular system undergoes either a loss or a gain of electrons or both and a relocation of the nuclei. These modifications of the molecular structure can be monitored through the successive responses of the energy with respect to the variation of the number of electrons and to the modification of the external potential [44].

$$d\mathcal{E} = \left(\frac{\partial \mathcal{E}}{\partial N}\right)_v dN + \int \left[\frac{\delta \mathcal{E}}{\delta v(r)}\right]_N dv(r)dr + \frac{1}{2}\left(\frac{\partial^2 \mathcal{E}}{\partial N^2}\right)_v dN^2 + dN \int \left[\frac{\delta^2 \mathcal{E}}{\partial N \delta v(r)}\right] dv(r)dr + \frac{1}{2}\iint \left[\frac{\delta^2 \mathcal{E}}{\delta v(r)\delta v(r')}\right]_N dv(r)dv(r')drdr' + \frac{1}{6}\left(\frac{\partial^3 \mathcal{E}}{\partial N^3}\right)_v dN^3 + \frac{1}{3}dN^2 \int \left[\frac{\delta^3 \mathcal{E}}{\partial^2 N \delta v(r)}\right] dv(r)dr + \ldots\ldots\ldots \quad (1)$$

Interestingly, each and every derivative of the energy within this framework has a chemical meaning generally related to a property already in use by theoreticians and experimentalists. The identification of the first derivative of the energy with respect to the number of electrons to the opposite of the electronegativity exemplifies this statement. It is worth noticing that three sets of derivatives exist. Derivatives with respect to the number of electrons give rise to global descriptors. Their values are the same wherever position they are calculated. The global descriptors measure the chemical reactivity of the system as a whole. The responses of the energy with respect to the external potential are local descriptors and constitute the second set of indexes. Their values are dependent on the position where they are evaluated. These indexes probe the chemical selectivity of the system. Generally, positions exhibiting the highest values of a local descriptor are more prone to react than positions exhibiting the lowest values. Finally, second derivatives with respect to the external potential give rise to nonlocal responses [45]. The values of this kind of descriptors depend on two sets of spatial variables, and they are likely related to the polarization of the system. Table 1 summarizes the names and definitions of the main global and local descriptors. The chemical relevance of the last two entries of Table 1 has not yet been investigated. As a consequence, they have not been named and will not been considered further in this chapter.

The two next sections detail the mathematical definition and chemical meaning of these indexes.

Table 1 Summary of global and local descriptors as energy derivatives

Derivative order	Mathematical definition	Name
0	$\mathcal{E}[N,v] = \mathcal{E}_{elec}[\rho]$	Energy
1	$\left(\frac{\partial \mathcal{E}}{\partial N}\right)_v = \mu$	Chemical potential [46]
	$\left(\frac{\delta \mathcal{E}}{\delta v(r)}\right)_N = \rho(r)$	Electron density
2	$\left(\frac{\partial^2 \mathcal{E}}{\partial N^2}\right)_v = \eta$	Chemical hardness [47]
	$\left(\frac{\delta^2 \mathcal{E}}{\partial N \delta v(r)}\right) = f(r)$	Fukui function [48]
	$\left(\frac{\delta^2 \mathcal{E}}{\delta v(r) \delta v(r')}\right)_N = \chi_1(r,r')$	Linear response [49]
3	$\left(\frac{\partial^3 \mathcal{E}}{\partial N^3}\right)_v = \left(\frac{\partial \eta}{\partial N}\right)_v = \gamma$	Hyper-hardness [50, 51]
	$\left(\frac{\delta^3 \mathcal{E}}{\partial^2 N \delta v(r)}\right) = \left(\frac{\partial f(r)}{\partial N}\right)_v = \left(\frac{\delta \eta}{\delta v(r)}\right)_N = f^{(2)}(r)$	Dual descriptor [50]
	$\left(\frac{\delta^3 \mathcal{E}}{\partial N \delta v(r) \delta v(r')}\right) = \left(\frac{\partial \chi(r,r')}{\partial N}\right)_v = \left(\frac{\delta f(r)}{\delta v(r')}\right)_N = \xi(r,r')$	
	$\left(\frac{\delta^3 \mathcal{E}}{\delta v(r'') \delta v(r') \delta v(r)}\right)_N = \left(\frac{\delta \chi(r,r')}{\delta v(r'')}\right)_N = \chi_2(r,r',r'')$	

2.2 Global Descriptions

As already mentioned in the previous section, global DFT-based descriptors connect the electronic features to the reactivity of the system. The very basis of conceptual DFT has been the identification of the first derivative of the energy with respect to the number of electrons (called chemical potential for obvious reasons) to the opposite of the electronegativity as it has been proposed by Iczkowski and Margrave [46, 52]:

$$\left(\frac{\partial \mathcal{E}}{\partial \mathrm{N}}\right)_v = \mu \tag{2}$$

Using the finite difference approximation, the chemical potential can be computed as

$$\mu \approx -\frac{1}{2}(\mathrm{I} + \mathrm{A}) \tag{3}$$

In Eq. 3, I and A respectively stand for the vertical ionization potential and the electron affinity. This working equation shows how the chemical potential is connected to Mulliken [53] electronegativity. By adding a second layer of approximation, based upon the Koopmans' theorem, one gets

$$\mu \approx \frac{1}{2}(\varepsilon_{\mathrm{LUMO}}[\mathrm{N}, \mathrm{v}] + \varepsilon_{\mathrm{HOMO}}[\mathrm{N}, \mathrm{v}]) \tag{4}$$

where $\varepsilon_{\mathrm{LUMO}}$ and $\varepsilon_{\mathrm{HOMO}}$ are the energies of the lowest unoccupied molecular orbital (LUMO) and the highest occupied molecular orbital (HOMO). Just like the electronegativity, the chemical potential evaluates the escaping tendency of the electrons.

Another great achievement of conceptual DFT was the proposal by Parr and Pearson [47] to measure the chemical hardness (concept introduced by the very same Pearson in the 1960s) using the second N-derivative:

$$\left(\frac{\partial^2 \mathcal{E}}{\partial \mathrm{N}^2}\right)_{\mathrm{v}} = \eta \tag{5}$$

Softness [54] has been naturally defined as the hardness inverse ($\mathrm{S} = 1/\eta$). Chemical hardness is a versatile concept upon which two quite important principles are built. Originally, chemical hardness has been designed to rationalize the outcome of Lewis acid-base reactions. As stated by Pearson [55], hard acids prefer to bind hard bases, and soft bases preferentially react with soft acids. This statement is the basis of the hard and soft acids and bases principle. As 90 % of the chemical reactions belong into this heading, the HSAB principle proved itself quite handy. Later considerations by Pearson led to believe that hardness is also a strong indicator of chemical stability. The maximum hardness principle [56, 57] that states that molecules tend to be as hard as possible is the consequence of the evolution of the concept.

Later on, another index built upon the chemical potential and the chemical hardness has been proposed to describe the electrophilicity of a system. The chemical potential, as first N-derivative, must be negative for an electronically stable system. Indeed, a positive chemical potential would lead to an autoionization of the system that would lower its energy. Since the chemical potential is always negative, all systems are considered as electrophile and their energies decrease as they acquire electrons. In a seminal publication, Parr and coworkers [58] have shown that the maximum number of electrons a system can acquire is given by

$$\Delta \mathrm{N} = -\frac{\mu}{\eta} \tag{6}$$

The lowering in energy is

$$\Delta \mathcal{E} = -\frac{\mu^2}{2\eta} \tag{7}$$

The electrophilicity index has been defined as

$$\omega = \frac{\mu^2}{2\eta} \tag{8}$$

The higher the value of this descriptor, the more electrophilic the system is.

2.3 Local and Nonlocal Descriptors

Besides global indexes, several local indexes have been defined as derivatives of the energy with respect to the external potential. Those descriptors whose values are dependent on the position they are evaluated ascribe the selectivity of atomic sites. In most cases, three different kinds of energies can control the selectivity of a chemical process, namely, electrostatic, frontier, and polarization interactions [59, 60]. Each local index happens to describe the corresponding local response.

The first v-derivative is the electron density, which enters the integral describing electrostatic interactions:

$$\left(\frac{\delta \mathcal{E}}{\delta \mathrm{v}(r)}\right)_{\mathrm{N}} = \rho(r) \tag{9}$$

Two kinds of second derivatives of the energy including the crossed N,v-derivative of the energy and the second v-derivative can be defined. The crossed derivative has been named Fukui function [48] after Kenichi Fukui, the founder of frontier molecular orbital theory [61, 62]. This function can be roughly approximated by the density of the frontier orbitals HOMO and LUMO.

$$\left(\frac{\delta^2 \mathcal{E}}{\delta v(r) \partial N}\right) = \left(\frac{\partial \rho(r)}{\partial N}\right)_v = f(r) \tag{10}$$

This quantity can be seen as the distribution function of one electron gained or lost over the system. To really get these two kinds of information, the right and left derivatives have been actually used. They are called electrophilic/nucleophilic Fukui functions [63] (f^+ and f^-):

$$\left(\frac{\partial \rho(r)}{\partial N}\right)_v^{+/-} = f^{+/-}(r) \tag{11}$$

From this function the actual location of the charge transfer can be monitored. Fukui function has been widely used in the last two decades to predict and rationalize the regioselectivity or regiospecificity of chemical processes [64–67].

The second v-derivative of the energy is the linear response function [68, 69]:

$$\left(\frac{\delta^2\mathcal{E}}{\delta v(r)\delta v(r')}\right) = \left(\frac{\delta\rho(r)}{\delta v(r')}\right)_v = \chi(r,r') \tag{12}$$

This function describes how the electron density is modified at point r′ when the external potential is changed at position r. Descriptors that depend on more than one set of coordinates are called nonlocal. This function is symmetrical with respect to the exchange of coordinates. The computation of the latter quantity is quite cumbersome. Besides, the chemical understanding of a function that depends on two sets of coordinates is quite difficult. It is therefore easily understandable that this function has been scarcely used so far. However, its usefulness in the rationalization of inductive and mesomeric effects [70] came out lately.

The last local descriptor presented here is the dual descriptor [71, 72], also known as second-order Fukui function. This index has been introduced in 2005 to predict whether an atomic site within a molecule is more electrophilic than nucleophilic or the other way around. The definition of the dual descriptor, symbolized by $\Delta f(r)$ or $f^{(2)}(r)$, as the third derivative of the energy is not very meaningful:

$$\left(\frac{\delta^3\mathcal{E}}{\partial^2 N\delta v(r)}\right) = f^{(2)}(r) = \Delta f(r) \tag{13}$$

An alternative and more useful definition is

$$\left(\frac{\partial^2\rho(r)}{\partial N^2}\right)_v = \Delta f(r) \tag{14}$$

Using the finite difference approximation and the Koopmans' theorem, the dual descriptor can be computed through

$$\left(\frac{\partial^2\rho(r)}{\partial N^2}\right)_v \approx f^+(r) - f^-(r) \approx \rho_{LUMO}(r) - \rho_{HOMO}(r) \tag{15}$$

From this working definition, it is plain that the sign of this descriptor describes the philicity of a molecular site. Positions with positive values of dual descriptor are more electrophilic than nucleophilic. The density of the LUMO at this location is higher than that of the HOMO. On the contrary, positions with negative values of dual descriptor are more nucleophilic than electrophilic. In this case, at those locations, the density of the HOMO is higher than that of the LUMO. This descriptor has proved quite useful to rationalize both the regioselectivity and the regiospecificity of chemical processes.

It is quite important to point out that local reactivity indexes described so far are not suited for a comparison between species. To perform these kinds of studies, one needs to combine selectivity and reactivity information. Thus several indexes have been designed to help rationalize the reactivity difference of the same atomic site

within a molecular family. Historically, the first descriptor proposed was the local softness [73]:

$$s(r) = \left(\frac{\partial \rho(r)}{\partial \mu}\right)_v = \left(\frac{\partial \rho(r)}{\partial N}\right)_v \left(\frac{\partial N}{\partial \mu}\right)_v = \frac{f(r)}{\eta} \tag{16}$$

Obviously, the local softness contains the same information about selectivity as the Fukui function. Local softness also contains additional information about the global reactivity of the molecule. Combined, these pieces of information allow making a comparison between molecules.

More recently, using a similar line of arguments, the grand canonical dual descriptor (GCDD), symbolized by $\Delta s(r)$ or $s^{(2)}(r)$, has been developed [74]:

$$s^{(2)}(r) = \Delta s(r) = \left(\frac{\partial s(r)}{\partial \mu}\right)_v = \frac{\partial}{\partial \mu}\left(\frac{f(r)}{\eta}\right)_v = \frac{\Delta f(r)}{\eta^2} - \frac{f(r)}{\eta^3}\left\{\frac{\partial \eta}{\partial N}\right\} \tag{17}$$

As the derivative of the chemical hardness with respect to the electron number is supposedly negligible, the GCDD is generally assessed through

$$s^{(2)}(r) = \Delta s(r) \approx \frac{\Delta f(r)}{\eta^2} \tag{18}$$

2.4 Selectivity Descriptors for Excited States

Descriptors presented in the previous section characterize either the reactivity or the selectivity for chemical processes that occur in the ground state. A recent extension of conceptual DFT has been proposed to tackle the selectivity of excited species [75, 76]. Basically, when an electronic system is excited by an electromagnetic field, its electron density is distorted. The main assumption is that the predominant process is the relaxation of the density (geometrical relaxation is overlooked). Mathematically, the whole extension is based upon a singularity of density functional theory. Indeed for whatever excited state, the electronic chemical potential is no longer a global quantity but on the contrary a local quantity. This property is due to the fact that the electron density of an excited species is not stationary. In other words, the excited electron density is not an optimum of the electron density functional. Translated into mathematics, this property reads

$$\left(\frac{\delta \mathcal{E}[\rho_i]}{\delta \rho(r)}\right)_v = \mu_i(r) \tag{19}$$

in which ρ_i is the density of the ith excited state. When the energy is computed using the true ground state electron density, Eq. 19 becomes

$$\left(\frac{\delta \mathcal{E}[\rho_0]}{\delta \rho(r)}\right)_v = \left(\frac{\partial \mathcal{E}[\rho_0]}{\delta N}\right)_v = \mu \equiv Cst \tag{20}$$

The chemical potential has then the same value anywhere within the molecular volume, noted *Cst* in Eq. 20. The minimization of the integral form of Eq. 19 provides an excellent descriptor of the chemical behavior of atomic site within the molecule:

$$\delta \mathcal{E} = \int\limits_{\rho_i \to \rho_0} \mu_i(r) \delta \rho(r) dr \tag{21}$$

Indeed, for this equation to be minimum, it is obvious that positions with high values of the local chemical potential are depleted during the relaxation process, while positions with low values of the local chemical potential gain electron density. To reformulate this statement, it can be established that during the relaxation process from the excited state to the ground state, an intramolecular electron flux appears driven by the opposite of the gradient of the local chemical potential. The main assumption upon which the whole development is based is that the physical relaxation can be induced by a chemical partner. Following this line, an electrophile would attack positions which are depleted during the density relaxation. On the other hand, a nucleophile would attack locations that gain electron density during the relaxation. If one is capable of getting an analytical expression of the local chemical potential, one can monitor the electron flux and get a chemical picture of the selectivity of excited species. From the basics of DFT, it is well known that

$$\mu_i(r) = v(r) + \frac{\delta F[\rho_i]}{\delta \rho(r)} \tag{22}$$

in which $F[\rho]$ is the Hohenberg-Kohn universal functional.

The next step is to assume that the excited density can be written as a distortion of the ground state density. Then, using a Taylor expansion, one gets after several cumbersome approximations the following expression:

$$\mu_i(r) \approx \mu_0 + \int \frac{\delta_0^i \rho(r')}{|r - r'|} dr' \tag{23}$$

Specifically, for the first excited state, the quantity $\delta_0^i \rho(r)$ has very recently been defined as the first-order state-specific dual descriptor [77]. It can be approximated by the difference between the densities of the LUMO and the HOMO and therefore is very similar to the usual dual descriptor defined in Sect. 2.3. Therefore, the second term of the right-hand side of Eq. 23 can be seen as an electrostatic potential created by the dual descriptor density. This potential constitutes the local response and has been called dual potential ($V^{\Delta f}$). It is quite straightforward to show that the

Table 2 Chemical behavior for the ground state and the first excited state

Ground state $\Delta f(r)$		Excited state $\delta_0^1\rho(r')$		
Sign	Character	Sign	Character	Interactions
+	Electrophilic	+	Nucleophilic	Favorable
–	Nucleophilic	–	Electrophilic	Favorable
+/–	el/nu	–/+	el/nu	Unfavorable

sign of this dual potential has a chemical meaning opposite to that of the dual descriptor for the ground state. As a consequence, the best interaction between two species, one of which being excited, is given in Table 2.

From these sets of descriptors, most of the chemical reactivity and selectivity can be explained. The next section is dedicated to the use of these indexes for the rationalization of the formation of several DNA lesions.

3 Application: Reactivity of Nucleic Acid Bases

3.1 Introduction

DNA, though responsible for carrying the genetic information, is characterized by some chemical instability; this may be particularly important when cellular DNA is exposed to external aggressions but also occurs under physiological conditions [1]. DNA is damaged continuously and, hopefully, in most cases efficiently repaired. Nonetheless, when repair is deficient, or if DNA replication takes place (mitosis) before repair is achieved, mutation may occur. Therefore genome integrity is lost and errors are susceptible to occur during transcription and translation processes. These can be responsible for cellular aging and cancers. Among the different DNA lesions that are known to be generated, many concern the nucleobases whose structures are depicted in Fig. 1. Despite many structural similarities – all of them being aromatic nitrogenated heterocycles – nucleobases show different susceptibilities to chemical or physical damaging agents. In order to find efficient ways to fight against the generation of DNA damage, it is fundamental on one hand to understand the mechanisms of formation of the different lesions that may be produced endogenously or exogenously in cellular DNA and, on the other hand, to be able to rationalize the differences of reactivity of the nucleobases toward a given genotoxic agent. The second aspect has been the subject of theoretical studies where the reactivity indexes presented in the previous section have been successfully applied to a number of cases [78–83]. Some of them will be described below with the aim of showing the wide range of reactivity issues that *conceptual DFT* can address.

In a first part, global and local reactivity indexes computed for the five main purine and pyrimidine nucleobases are gathered in the same place in order to have preliminary information on the differences between the nucleobases. The second

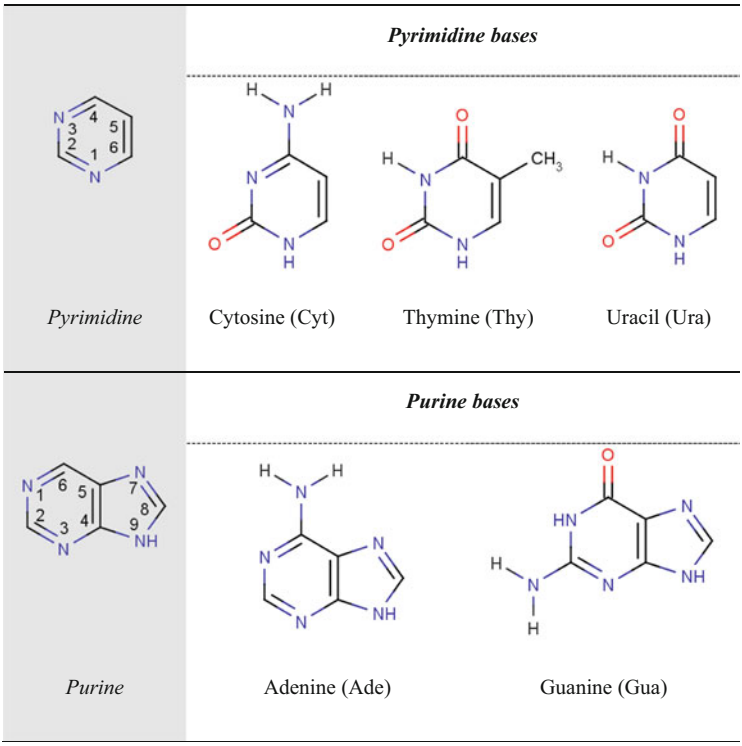


Fig. 1 Structure of DNA bases found in all living beings. Note that uracil is only present in RNA, contrary to thymine which is found in DNA only

part concentrates on information showing that reactivity descriptors can provide relevant structural features adopted by the nucleobases (base-pair formation, tautomers). Finally, the third and last part focuses, strictly speaking, on reactivity differences between the bases in the particular context of the formation of a given lesion.

3.2 *Reactivity Descriptors Computed for Isolated Nucleobases*

All the molecular properties have been computed through DFT methodology by using the B3LYP functional with the triple-zeta split valence basis set 6-311G(d,p). All the computations have been performed using the Gaussian package software [84]. The molecules have been fully optimized. Then the DFT-based descriptors have been calculated using the two usual sets of approximations, first the finite difference approximation, then the Koopmans' theorem. The latter approximation is questionable, especially for local indexes, but to our best knowledge, it has

always proved satisfactory for this kind of systems. Practically, the post calculations and condensations have been performed using the procedure proposed by Chamorro and Perez through their package "Global and Local Descriptor version 2.0" [85]. In this first part, the reactivity and selectivity descriptors of isolated bases are presented.

Global Indexes

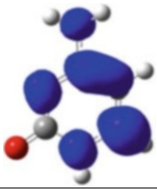
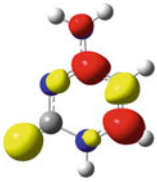
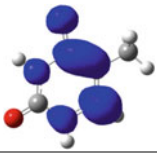
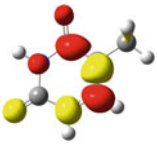
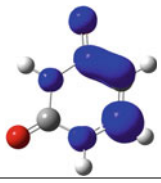
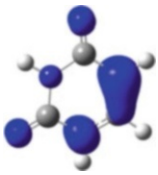
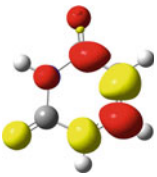
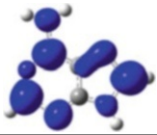
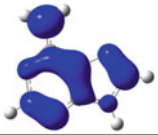
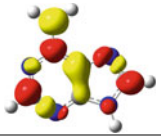
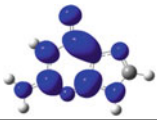
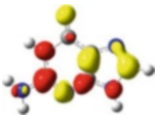
The global indexes calculated as stated above are gathered in Table 3. It is plain to note that the chemical potential (μ) of purine bases appears higher (less negative) than that of pyrimidine bases, making the former components better electron donors than the latter constituents. More precisely, Gua appears as the base which is the most easily oxidized, while Ura is the most resistant to oxidants. This is consistent with the fact that Gua is the preferential target in the formation of oxidatively generated damage at least when exposed to one-electron oxidants and singlet oxygen [18, 27]. In particular it can explain that usually Gua is the main target of ionization giving rise predominantly to 8-oxo-7,8-dihydroguanine (8-oxoGua) through the transient formation of the guanine radical cation either directly or after charge transfer from other one-electron oxidized nucleobases [18, 27, 86].

Besides, it is remarkable how close the hardness (η) of the five nucleobases are to each other (and thus their softness (S)), suggesting a similar resistance to charge transfer, though it can be noticed from comparison between Thy and Ura that methylation on C5 decreases the hardness by about 0.3 eV. Since the electrophilicity power is computed through ($\mu^2/2\eta$), the differences in electrophilicity power (ω) between the bases are mainly due to differences in their chemical potential. It appears that the purine bases are being better nucleophiles than the pyrimidine bases. Descriptors are also able to characterize the selectivity of molecules. This is the subject the next section focuses on.

Table 3 Global indexes computed from frontier orbitals at the B3LYP/6-311G(d,p) level

		Canonical ensemble		Grand canonical ensemble		
		First order	Second order	First order	Second order	Other
Bases		μ (eV)	η (eV)	N	S (eV^{-1})	ω (eV)
Pyrimidines	Cyt	−3.70	5.38	58	0.186	1.28
	Thy	−4.16	5.39	66	0.185	1.60
	Ura	−4.26	5.70	58	0.175	1.59
Purines	Ade	−3.40	5.44	70	0.184	1.06
	Gua	−3.33	4.58	78	0.179	1.20

Table 4 Local reactivity indexes computed from frontier orbitals at B3LYP/6-311G(d,p) level

		Second order		Third order
		$f^+(r)$	$f^-(r)$	$\Delta f(r)$
				Red: more E than Nu
Base		Electrophilic sites	Nucleophilic sites	Yellow: more Nu than E
Pyrimidine	Cyt			
	Thy			
	Ura			
Purine	Ade			
	Gua			

Local Indexes of the Canonical Ensemble

Local reactivity descriptors for pyrimidine and purine bases are reported in Tables 4 and 5. Specifically, in Table 4 spatial maps of electrophilic and nucleophilic Fukui functions along with dual descriptor are displayed, while in Table 5 the corresponding condensed values are given. From a general point of view, it can be noted that spatial maps of Fukui functions are not very conclusive, while the tabulated data are more useful. Indeed the most reactive site, with respect to either an electrophile or a nucleophile, is respectively obtained by looking at the highest value of the electrophilic or nucleophilic Fukui function. On the contrary, since the sign of the dual descriptor carries information about the chemical philicity (electrophilicity/nucleophilicity), the maps are very often self-sufficient for rationalizing the chemical selectivity.

Table 5 Local reactivity indexes of pyrimidine and purine bases condensed through Chamorro and Perez scheme

Pyrimidine bases	Cytosine	**Atom number**	$\mathbf{f^+_k}$	$\mathbf{f^-_k}$	$\mathbf{\Delta f_k}$
		N1	0.112	0.157	−0.045
		C2	**0.007**	**0.015**	**−0.007**
		N3	0.127	0.176	−0.049
		C4	0.190	0.007	0.183
		C5	0.097	**0.262**	**−0.164**
		C6	**0.393**	0.067	**0.326**
		N4 exo	0.001	0.002	−0.001
	Thymine	**Atom number**	$\mathbf{f^+_k}$	$\mathbf{f^-_k}$	$\mathbf{\Delta f_k}$
		N1	0.056	0.231	−0.175
		C2	**0.014**	**0.010**	**0.003**
		N3	0.061	0.002	0.060
		C4	0.185	0.008	0.177
		C5	0.158	**0.331**	**−0.173**
		C6	**0.381**	0.149	**0.233**
		O4 exo	0.122	0.085	0.037
	Uracil	**Atom number**	$\mathbf{f^+_k}$	$\mathbf{f^-_k}$	$\mathbf{\Delta f_k}$
		N1	0.059	0.253	−0.194
		C2	**0.019**	**0.009**	**0.010**
		N3	0.059	0.000	0.059
		C4	0.160	0.006	0.154
		C5	0.171	**0.375**	**−0.204**
		C6	**0.399**	0.123	**0.276**
		O4 exo	0.113	0.118	−0.006
Purine bases	Adenine	**Atom number**	$\mathbf{f^+_k}$	$\mathbf{f^-_k}$	$\mathbf{\Delta f_k}$
		N1	0.126	0.149	−0.023
		C2	0.185	0.069	0.117
		N3	0.005	0.045	−0.040
		C4	0.211	0.067	0.144
		C5	0.036	0.159	−0.124
		C6	0.007	0.056	−0.050
		N7	0.073	0.060	0.013
		C8	0.246	0.121	0.126
		N9	0.038	0.015	0.024
	Guanine	**Atom number**	$\mathbf{f^+_k}$	$\mathbf{f^-_k}$	$\mathbf{\Delta f_k}$
		N1	0.100	0.011	0.089
		C2	0.277	0.059	0.218
		N3	0.008	0.048	−0.040
		C4	0.192	0.112	0.080
		C5	0.110	0.220	−0.110
		C6	0.044	0.017	0.027
		N7	0.008	0.048	−0.040
		C8	0.028	0.175	−0.147
		N9	0.070	0.004	0.066

Pyrimidine Bases

The substantial contribution of C5 and C6 carbons of pyrimidine bases to the $f^{+}(r)$ and $f^{-}(r)$ descriptors indicates that the C5–C6 double bond of the pyrimidine ring is susceptible to both nucleophilic and electrophilic attacks. Nonetheless, the sign of the dual descriptor $\Delta f(r)$ indicates than carbon C5 is more nucleophilic than electrophilic ($\Delta f(r) < 0$), while this is the reverse for carbon C6 ($\Delta f(r) > 0$). Thus, an electrophile will add more efficiently to carbon C5 than to carbon C6, contrary to a nucleophilic reagent which will add more rapidly to carbon C6. This is consistent with the fact that the electrophilic $^{\bullet}$OH, which is responsible for the majority of oxidatively generated damage to DNA, adds preferentially to carbon C5 of pyrimidine bases compared to carbon C6 [87]. It may be pointed out that $^{\bullet}$OH addition to the 5,6-pyrimidine double bond is under kinetic control since several computational studies have shown that addition to carbon C6 leads to a more stable adduct than addition to carbon C5 [88, 89].

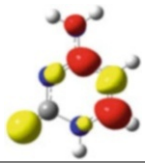
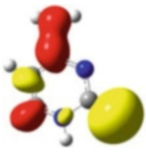
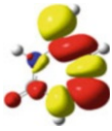
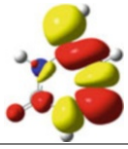
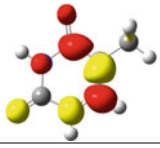
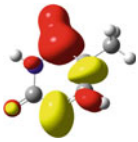
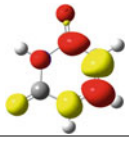
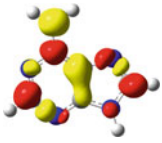
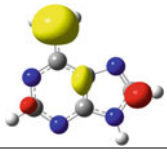
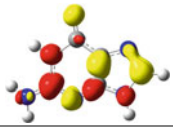
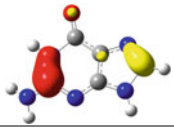
It is well documented that Cyt, Ade, and Gua may spontaneously deaminate through hydrolysis. Carbons holding the corresponding amino groups (C4 of Cyt, C6 of Ade, and C2 of Gua) all exhibit an electrophilic character. It will be shown in the following, in the particular case of Cyt (vide infra), that the reaction mechanism associated with spontaneous deamination reactions involves nucleophilic addition of a water molecule to those carbons (see Sect. 3.4).

Another well-documented fact is the chemical stability of the urea moiety of pyrimidine bases. Thus, position C2 is almost insensitive to both electrophiles and nucleophiles. All the selectivity indexes corroborate this statement for each pyrimidine base. As can be seen in Tables 4 and 5, neither the electrophilic nor the nucleophilic Fukui function exhibits significant values on C2. It is even clearer using the dual descriptor for which the value on C2 is roughly seven times lower than that of the other carbons.

Purine Bases

The most striking feature of purine bases is that two carbons of the 5-membered imidazole ring of the purine bases exhibit a different sign of the dual descriptor $\Delta f(r)$. C4 of Ade is nucleophilic only (almost zero contribution to $f^{+}(r)$), while that of Gua shows both nucleophilic and electrophilic characters (nonzero contribution to $f^{+}(r)$ and $f^{-}(r)$), the latter prevailing on the former. This is exactly the reverse for carbon C8 which corresponds to a pure nucleophilic site in Gua and to a nucleophilic and electrophilic site in Ade though more electrophilic than nucleophilic. In the following this difference in reactivity of the C8 carbon of purine bases is proposed to participate in the difference of reactivity of Gua and Ade toward the formation of tandem lesions in oxygen-free aqueous solution when DNA is exposed to $^{\bullet}$OH generated by radiolysis (see Sect. 3.4).

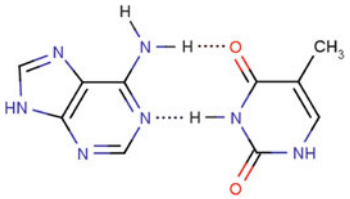
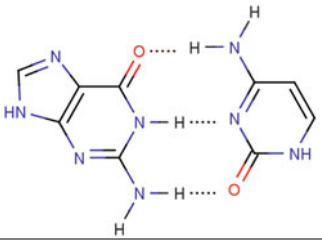
Table 6 Comparison of the dual descriptor and the dual potential computed from frontier orbitals at B3LYP/6-311G(d,p) level

Bases		Ground state $\Delta f(r)$ Red: more E than Nu Yellow: more Nu than E	First excited state $V^{\Delta f}(r)$ Red: more Nu than E Yellow: more E than Nu
Pyrimidines	Cyt (canonical)		
	Cyt (imino tautomer)		
	Thy		
	Ura		
Purines	Ade		
	Gua		

First Excited State

Finally, the latter section deals with the preliminary information that can be extracted from indexes describing excited states. As already explained in Sect. 2.4, it can be inferred from the sign of the dual descriptor and that of the dual potential (see Table 6) that the nucleophilicity/electrophilicity excess of the reactive sites of the nucleobases is reversed when moving from the ground state to the first excited state of a base. As a consequence, the chemical behavior toward a partner is reversed too: sites which most easily give electrons in the ground state

Table 7 Watson-Crick base pairs

	
AT base pair	GC base pair

most favorably accept electrons in the first excited state and vice versa. We will see that this plays an important role in the regioselectivity observed in the formation of pyrimidine dimers when DNA is exposed to UVB radiation as discussed in Sect. 3.4.

3.3 Structural Features of DNA

The results presented so far deal only with isolated DNA bases in their ground and first excited states. In this context, they are only relevant for the test tube study of chemical reactions. Within biological environment, the chemical reactivity and selectivity of DNA bases may change. For instance, what is the impact of pyrimidine-purine hydrogen bond coupling? Besides, the influence of DNA backbone and chemical environment may modify the relative contribution of the tautomers. In this section, some interesting results regarding the role of such factors upon the chemical behavior of DNA bases are presented.

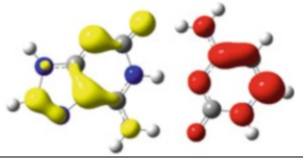
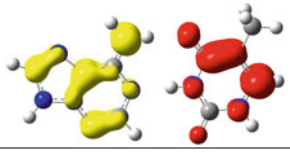
Formation of Base Pairs

The influence of the hydrogen bonding on the reactivity of DNA and RNA bases has been recently investigated by Toro-Labbé et al. [90] (Table 7). Their computation has been reproduced using B3LYP/6-311G(d,p) level of theory. The results are given in Tables 8 and 9. It appears that both the global and local reactivity are affected by the base pairing. Both chemical potential and hardness are of general lower value. The lowering of chemical potential signifies that, once paired, the bases are more easily excited and ionized. They are also more prone to react as their global hardness decreases. The local selectivity is also widely modified since it appears from the dual descriptor maps that electrophilic and nucleophilic sites are gathered on only one base. As it can be seen in Table 9, purine bases muster all the nucleophilic sites, while pyrimidine bases gather the electrophilic sites.

Table 8 Watson-Crick base pairs global descriptors computed from frontier orbitals at B3LYP/6-311G(d,p) level

		Canonical ensemble		Grand canonical ensemble		
		First order	Second order	First order	Second order	Other
DNA base pairs		μ (eV)	η (eV)	N	S (eV^{-1})	ω (eV)
Ade	Thy	−3.57	4.87	136	0.20	1.30
Gua	Cyt	−3.28	3.74	136	0.27	1.45

Table 9 Watson-Crick base pairs dual descriptor maps computed from frontier orbitals

DNA base pairs	$\Delta f(r)$ Red: more E than Nu Yellow: more Nu than E
Guanine-cytosine	
Adenine-thymine	

It is also interesting to compare the two base pairs. The couple Ade-Thy has a lower chemical potential than that of Gua-Cyt, which means that the latter is more easily ionized for generating a radical cation. As already discussed, nucleophilic zones are mainly located on the purine moieties. Therefore the generated radical is also very likely to be located on the purine bases. The couple Gua-Cyt also has a lower hardness and therefore is more prone to undergo chemical reactions than the couple Ade-Thy. From this investigation, two main conclusions can be drawn. First, the effect of pairing can be rather important upon selectivity; the meaning of the dual descriptor between two species in interaction might be questionable though. Second, the couple Gua-Cyt is more reactive than the couple Ade-Thy.

Tautomerism

It is generally assumed that DNA bases do not possess tautomers at neutral pH and that they are always present in DNA in their canonical forms (see Table 10). However, numerous tautomers can be built by modifying the position of a hydrogen atom (prototropy). An example is given for cytosine for which four tautomers are easily identified. Chemical potential, hardness, and electrophilicity values for each of the tautomers are given in Table 11. As can be seen in Table 11, the hydrogen shift mildly modifies the chemical reactivity. The chemical potential varies of about

Table 10 Different forms of cytosine along with dual descriptor maps computed from frontier orbitals at B3LYP/6-311G(d,p) level

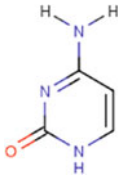
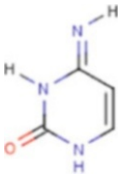
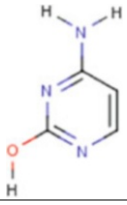
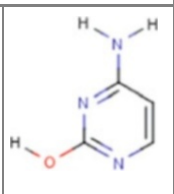
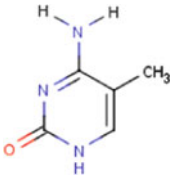
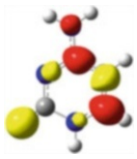
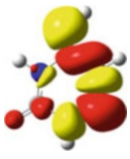
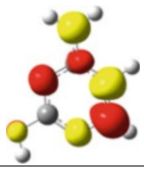
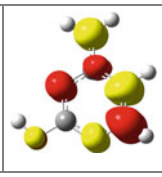
Cytosine				5-Methylcytosine
				
(1)	*(2)*	*(3)*	*(4)*	
Canonic	*Tautomers*			*Methylated (epigenetic mark)*
				*Δf(r)*

Table 11 Chemical potential, hardness, and electrophilicity values for the main tautomers of cytosine computed from frontier orbitals at B3LYP/6-311G(d,p) level

Numbering		μ (eV)	η (eV)	ω (eV)	Rel. E_{tot} (kJ/mol)
(1)	Cyt – amino/oxo	−3.70	5.38	1.28	0.0
(2)	Cyt – imino/oxo	−3.80	5.41	1.34	+4.5
(3)	Cyt – amino/enol – 1	−3.56	5.81	1.09	+3.3
(4)	Cyt – amino/enol – 2	−3.57	5.80	1.10	+6.6

0.25 eV. The variations of chemical hardness are larger than that of chemical potential and go up to 0.40 eV.

From the local point of view, it is worth noticing that the exocyclic N4 atom of the amino group is predicted as a nucleophile for all the tautomers, while for the canonical form it is proposed to be electrophile. This modification of the local reactivity has proved to be useful for understanding the occurrence of pyrimidine dimer lesions (see Sect. 3.4).

3.4 Reactivity Differences Between the Nucleobases

As mentioned earlier, experimentalists have identified many examples where the nucleobases show differences in the susceptibility to a given genotoxic agent/environment, for example, because they measured different rate constants associated with the formation of a transient damage, or because they observed different regioselectivities associated with a given set of experimental conditions. In those

cases where the lesion is formed under kinetic control, conceptual DFT reactivity descriptors can be helpful in providing rationalizing elements. As the nucleobases have different numbers of electrons (see Table 3), responses of the grand canonical potential are expected to be the relevant reactivity descriptors to analyze for this purpose. The spontaneous deamination of cytosine and its derivatives will be used to illustrate how such an approach can allow quantifying differences of reactivity.

Nonetheless, it will be shown beforehand that the consideration of the dual character of some reactivity descriptors of the canonical example can give useful information as long as one is interested by qualitative differences in reactivity. This will be illustrated through two examples: one lesion formed on the electronic ground state potential energy surface and another lesion involving a base in an electronic excited state.

Qualitative Differences of Reactivity

In the Ground State: Exploiting the Dual Descriptor

When DNA is exposed to ionizing radiation, surrounding water molecules are subject to radiolysis with the subsequent generation, among several reactive oxygen species (ROS), of highly reactive $^{\bullet}$OH [91]. The latter ROS is able to induce several types of oxidatively generated damage to DNA [18, 27]. Depending on the conditions, typically aerobic versus anaerobic, some lesions are specifically produced or their formation favored. In anaerobic conditions, a covalent link can be more efficiently formed between carbon C8 of a purine base and radicals issued from either its adjoining sugar moiety [40, 41] or the methyl group of an adjacent Thy [19, 39]. Though both purine bases can be involved in the formation of these lesions, Gua is experimentally observed to be more susceptible than Ade [92–94].

The general reaction mechanism proposed for the formation of these lesions involving Gua is shown in Fig. 2. The first step involves $^{\bullet}$OH-mediated hydrogen atom abstraction from either the sugar moiety of Gua or the methyl group of a Thy, leading to the formation of an electrophilic $RR'CH_2^{\bullet}$ or $RR'R''CH^{\bullet}$ radical. Subsequently, the carbon-centered radical thus formed adds to carbon C8 of the purine base by electrophilic radical addition. $H^{\bullet}$ abstraction by a radical from the environment or through the implication of O_2 in aerated solutions [40] finally leads to the formation of the tandem lesion.

This mechanism has been studied computationally by DFT in the cases of Gua and Ade [95–97].

For both types of lesions (cdG/cdA and G^T/A^T), the electrophilic radical addition to C8 appears to be responsible for the difference of reactivity observed between the two purine bases. Computed activation energies associated with this radical step are reported in Table 12.

First, from the point of view of the relative global reactivity of the purine bases, since the chemical potential of Gua is higher than that of Ade (see Table 3), an electrophilic addition to Gua is expected to be less destabilizing (lower activation

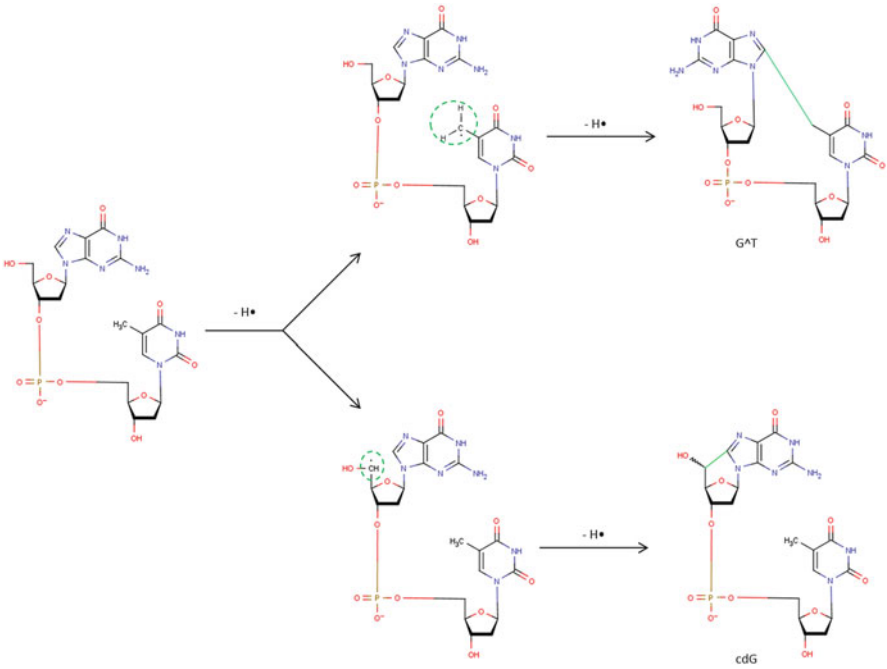


Fig. 2 Reaction scheme for the formation of cross-linked adducts between a purine base and either an adjacent thymine (G^T) or its sugar moiety (cdG)

Table 12 Activation energies (in kJ/mol) computed by DFT for the radical addition step leading to the formation of the tandem lesions sketched in Fig. 2

Lesion	$\Delta^{\neq}E$ (ZPE included)	Computational methodology	Reference
G^T	72.3	B3LYP/6-31G(d,p)	Xerri et al. [96], Labet et al. [98]
A^T	173.5		
cdG	41.0	B3LYP/6-31 + G(d,p)	Zhang and Eriksson [95]
cdA	55.2		

barrier) than to Ade. In addition, from a more local point of view, as we mentioned earlier, the two purine bases are characterized by different signs of the dual descriptor $\Delta f\left(\vec{r}\right)$ in the region of their C8 carbon, the site of addition during the formation of the tandem lesions. Thus, though both nucleophilic, the C8 carbons of the two purine bases are qualitatively different: carbon C8 of Gua is more nucleophilic than electrophilic ($\Delta f(r) < 0$), while carbon C8 of Ade is more electrophilic than nucleophilic ($\Delta f(r) > 0$) (see Table 4).

Now the dual descriptor $\Delta f(r)$ has been shown to be closely related to the maximum hardness principle [72]. When an electrophile approaches a nucleophile in a region characterized by $\Delta f(r) < 0$, then the hardness of the nucleophile is expected to increase, which is favorable. On the contrary, if the electrophile attacks the nucleophile on a reaction site characterized by $\Delta f(r) > 0$, the hardness of the nucleophile decreases, which is unfavorable.

As a conclusion, from both a global and local reactivity points of view, Gua is expected to be more susceptible than Ade to an electrophilic attack on its carbon C8. This explains the different formation yields experimentally observed between the purine bases for the tandem lesions reported in Fig. 2.

In an Excited State: Exploiting the Dual Potential

When DNA is exposed to UVB radiation (290–320 nm), the nucleobases can get electronically excited, thus allowing photochemical reactions to take place. Covalent links between two adjacent nucleobases can then be formed [32]. This is particularly frequent at bipyrimidine sites, i.e., when two pyrimidine bases are adjacent on the same DNA strand. As mentioned previously, the resulting photocycloadducts are called pyrimidine dimers. They are of two kinds: *i*) *cis-syn* cyclobutane pyrimidine dimers (CPDs) resulting from the [2 + 2] photocycloaddition between the C5 and C6 double bonds of two adjacent pyrimidine bases and (*ii*) pyrimidine (6–4) pyrimidone photoproducts (6-4PPs) resulting from a Paternò-Büchi [2 + 2] photocycloaddition between the C5 and C6 double bond of the 5′-end pyrimidine base and the exocyclic C4-O/C4-N bond of the 3′-end thymine/cytosine imino tautomer, followed by opening and rearrangement of the oxetane/azetidine ring. The two kinds of pyrimidine dimers involving two adjacent thymines are shown in Fig. 3.

A Regioselectivity Problem

The eight possible bipyrimidine lesions susceptible to be formed (two structural types of dimers – four different bipyrimidine sites) are produced in very different proportions as shown in Fig. 4 where the yields of formation of each photoproduct within isolated DNA exposed to UVB radiation are reported [99].

From the consideration of the formation yields, it can be noticed that at any given bipyrimidine site, CPDs are produced more efficiently than 6-4PPs. The effect is more pronounced when the 3′-end pyrimidine base is Thy. To rationalize this experimental finding, which is merely a regioselectivity issue, reactivity descriptors can be analyzed. This is indeed possible, despite the fact this concerns photochemical properties, because a local descriptor that is applicable to the first excited state (the dual potential presented in Sect. 2.4) is now available in addition to the "classical" dual descriptor available for the ground state.

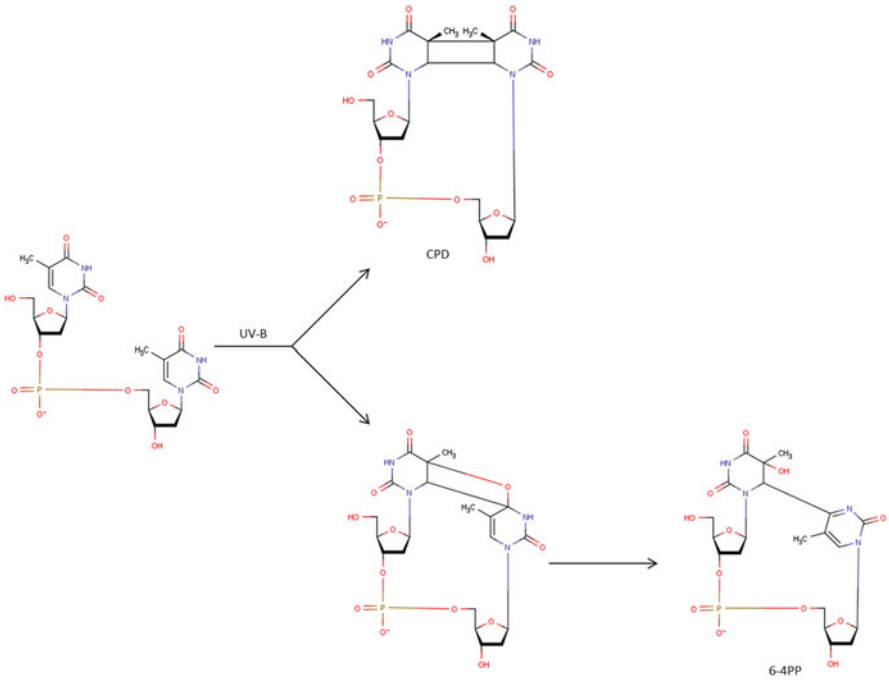


Fig. 3 Reaction scheme for the UV-induced formation of thymine dimers: cyclobutane pyrimidine dimer (CPD) and pyrimidine (6-4) pyrimidone photoproduct (6-4PP)

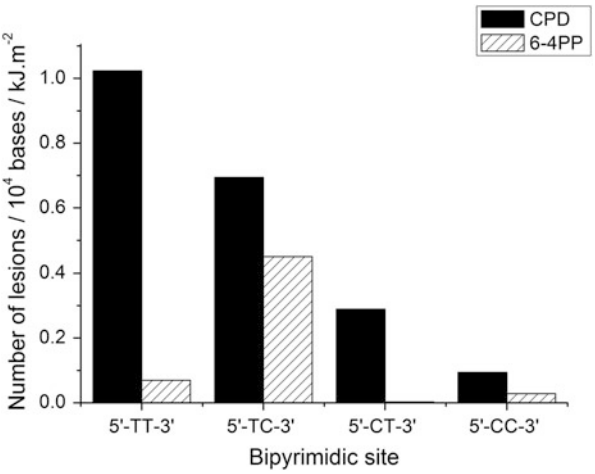


Fig. 4 Yield of formation of the main photoinduced bipyrimidine products within isolated DNA exposed to UVB radiation (dose range, 0–3.4 kJ·m^{-2}) [99]

Table 13 Chemical behavior related to the sign of the dual descriptor for the ground state and the dual potential for the first excited state

Electronic state	Sign of the relevant descriptor	Chemical behavior
Ground state	$\Delta f\left(\vec{r}\right) > 0$	Electrophile
	$\Delta f\left(\vec{r}\right) < 0$	Nucleophile
First excited state	$V^{\Delta f}\left(\vec{r}\right) > 0$	Nucleophile
	$V^{\Delta f}\left(\vec{r}\right) < 0$	Electrophile

During the [2 + 2] photocycloadditions involved in both the formation of CPDs and 6-4PPs, one of the nucleobases is still in the ground state, while the other lies in the first excited state corresponding to the first $\pi \rightarrow \pi^*$ transition. Reaction between the two bases is all the more favorable because electrophilic sites of the excited base interact with nucleophilic sites of its partner and vice versa. As mentioned earlier and reminded in Table 13, the chemical behavior of reactive sites of the excited base is characterized by the sign of the dual potential $V^{\Delta f}\left(\vec{r}\right)$; as for that of the reactive sites of the base in its ground state, it can be inferred from the sign of the dual descriptor $\Delta f(r)$. As a consequence interactions between a nucleophilic site of one base and an electrophilic site of the other base correspond to interactions between sites where the dual descriptor of one base and the dual potential of the other base are of the same sign.

About the Number of Stabilizing Interactions

Interacting sites are related to the pyrimidine dimer of interest. Formation of a CPD involves interaction between carbons C5 of both bases plus interaction between their carbons C6. The formation of a 6-4PP also involves two interactions: a first one between carbon C6 of the 5′-end base and carbon C4 of the 3′-end one and a second interaction between carbon C5 of the 5′-end base and O or N of the 3′-end base. In Fig. 5 the interactions that develop between the dual descriptor of the 5′-end base and the dual potential of the 3′-end base are shown for each of the eight potential photoadducts. Note that the fact to have considered the 3′-end base to be excited rather than the 5′-end one is purely arbitrary. Indeed, since the electrophilicity/nucleophilicity excess of the reactive sites is exactly reversed between the ground state and the first excited state (see Table 13), results would have been the same if we had considered the another alternative.

Since in a given electronic state the chemical behavior of C5 and C6 carbons is the same for Thy and Cyt and that this chemical behavior is reversed between the ground state and the first excited state, formation of CPDs systematically requires two stabilizing interactions, whatever the sequence of bases involved.

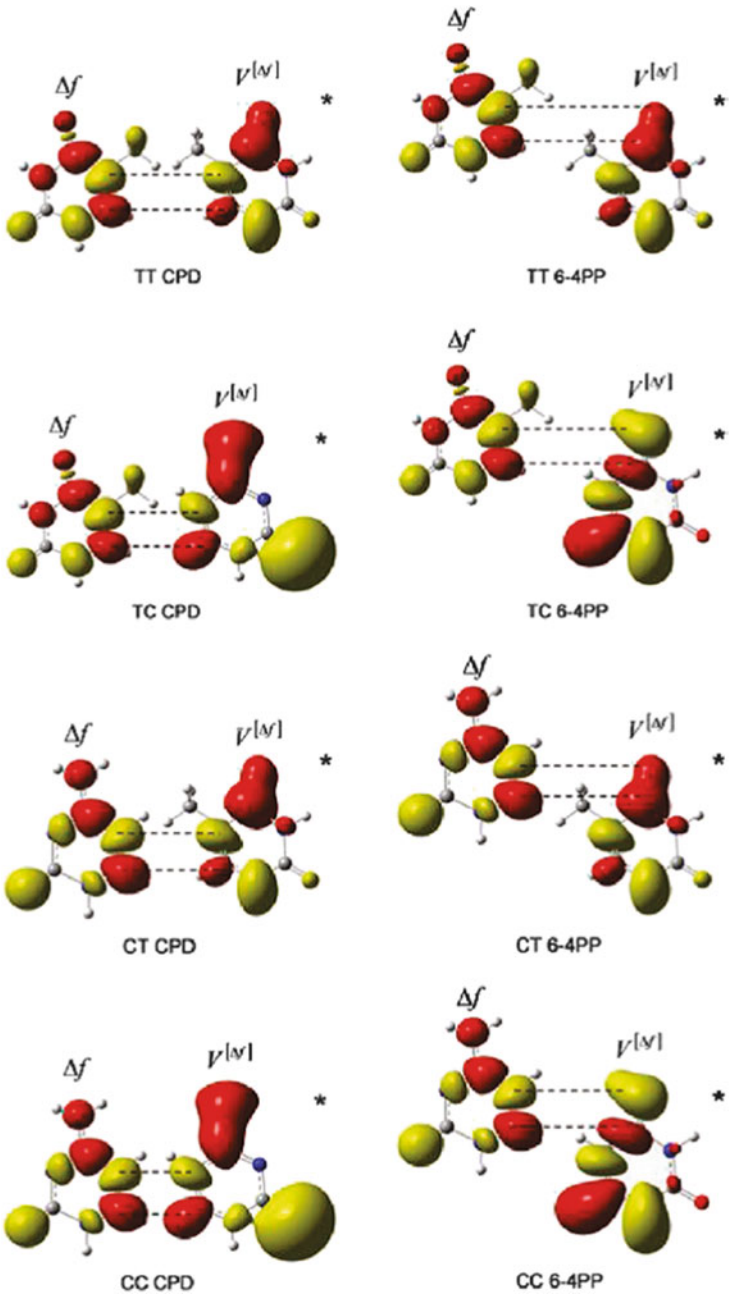


Fig. 5 Interactions at the four kinds of bipyrimidine sites between the dual descriptor of the 5′-end base in its ground state and the dual potential of the 3′-end base in its first excited state to form either a CPD or a 6-4PP lesion (Adapted from Ref. [76])

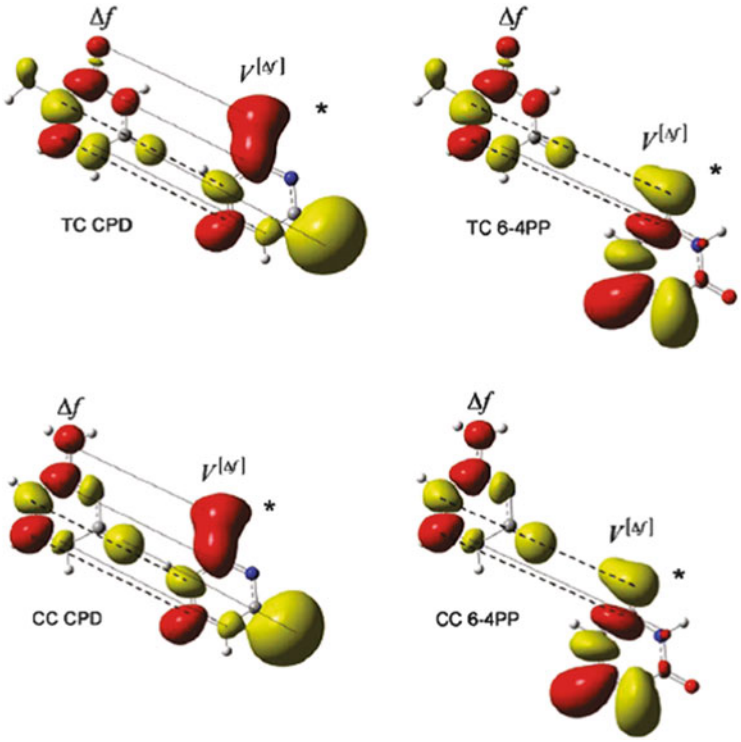


Fig. 6 Interactions at 5′-TC-3′ and 5′-CC-3′ bipyrimidine sites between the dual descriptor of the 5′-end pyrimidine base in its ground state and the dual potential of the 3′-end pyrimidine base in its first excited state to form either a CPD or a 6-4PP lesion. *Dashed lines* correspond to interactions leading to formation of the σ-bonds during the [2 + 2] photocycloaddition reaction. *Dotted lines* correspond to secondary interactions (Adapted from Ref. [76])

The formation of 6-4PPs, as shown in Fig. 5, involves two stabilizing interactions when a Cyt (in its imino form) constitutes the 3′-end base, but only one when this is a Thy. This is due to the fact that the oxygen linked to carbon C4 in Thy and the nitrogen linked to carbon C4 in the imino form of Cyt have reversed chemical behaviors. Therefore, the activation energy associated with the formation of a 6-4PP lesion is expected to be higher when a Thy base is at the 3′-end rather a Cyt. This explains why the formation of 6-4PP at 5′-TT-3′ and 5′-CT-3′ bipyrimidine sites is so inefficient compared to the corresponding CPDs.

About Secondary Interactions

The relative orientation of adjacent pyrimidine bases is governed by DNA double-stranded structure. Reacting bases approach each other in parallel planes (*supra-supra* approach). In addition to the interactions mentioned above which are

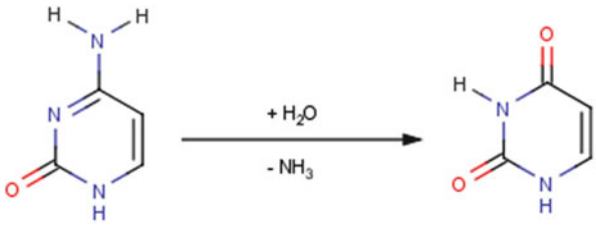


Fig. 7 Spontaneous cytosine deamination

responsible for the formation of two covalent links between adjacent pyrimidine bases, secondary interactions can develop. These are shown in Fig. 6 for the four dimers formed at bipyrimidine sites with 3′-end Cyt. The approach of pyrimidine bases to form 6-4PPs does not involve secondary interactions contrary to the approach giving rise to CPDs. In the latter case, four stabilizing secondary interactions develop (dotted lines in Fig. 6), facilitating the formation of CPDs compared to that of 6-4PPs, even in cases where the two primary interactions are stabilizing.

As a conclusion, the more efficient production of CPDs compared to 6-4PPs has two origins: (i) the number of stabilizing interactions responsible for the formation of covalent bonds between the two adjacent bases, which is 2 for all CPDs and 2 for 6-4PPs involving a 3′-end Cyt but only 1 for 6-4PPs involving a 3′-end Thy and (ii) the development of stabilizing secondary interactions during the formation of CPDs which do not exist in the case of 6-4PPs.

Quantitative Differences of Reactivity

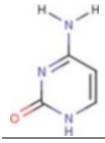
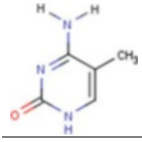
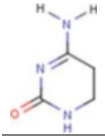
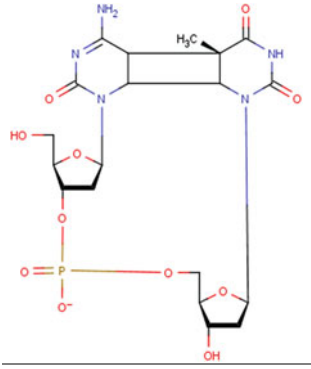
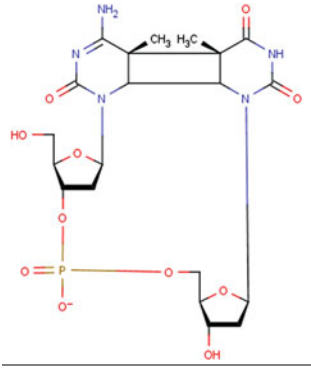
Exploiting the Grand Canonical Equivalent of the Dual Descriptor

Spontaneous Cyt deamination by hydrolysis occurs in living organisms. This hydrolytic reaction gives rise to Ura, the equation being shown in Fig. 7, whose abnormal presence in DNA is hopefully easily detected and removed by a dedicated repair enzyme.

Modified forms of canonical Cyt bases are also susceptible to deamination. This is, for example, the case of 5-methylcytosine and of 5,6-saturated derivatives of Cyt such as cytosine-containing pyrimidine dimers formed under UV radiation. These structural modifications affect strongly the deamination rate constants as reported in Table 14. Saturation of the C5–C6 bond accelerates considerably the deamination reaction. Incidentally, methylation at position 5 increases the deamination rate for 5,6-unsaturated cytosine derivatives, while it has exactly the opposite effect on 5,6-saturated derivatives.

A mechanism has been proposed and studied by DFT for cytosine deamination which is in agreement with the activation energy determined experimentally for the reaction (117 ± 4 kJ/mol [104]). The deamination pathway involves in a first step a nucleophilic addition of a water molecule to carbon C4 with the assistance of a second water molecule from the environment, followed by protonation of the amino

Table 14 Experimental deamination rate constant (k) of cytosine and some of its derivatives under various conditions

Bases	k	Experimental conditions	References
	2.1×10^{-10} s^{-1}	Single-stranded DNA pH = 7.4 – T = 37 °C	[100]
	9.5×10^{-10} s^{-1}	Single-stranded DNA pH = 7.4 – T = 37 °C	[100]
	6×10^{-5} s^{-1}	Free base pH = 7.4 – T = 37 °C	[101]
	10^{-3} min^{-1}	Dinucleoside monophosphate pH = 7 – T = 25 °C	[102]
	10^{-5} min^{-1}	Dinucleoside monophosphate pH = 7 – T = 25 °C	[103]

group and finally -NH_3 departure [98]. The same mechanism happens to hold for 5-methylcytosine deamination [105]. For 5,6-saturated derivatives, it appears that a tautomerization step leading to the imino form of the cytosine derivative is needed prior to the nucleophilic addition [106, 107]. In each case, the assisted nucleophilic

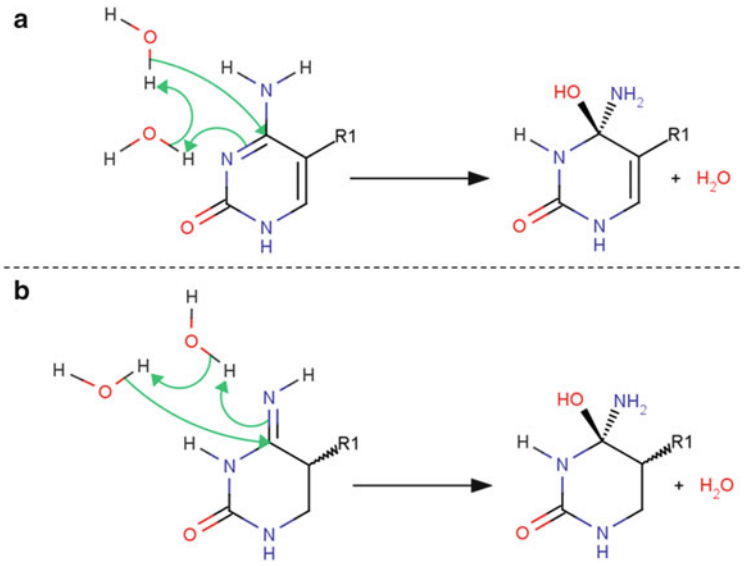


Fig. 8 Rate-controlling step of the mechanism proposed for deamination of 5,6-unsaturated (**a**) and 5,6-saturated (**b**) cytosine derivatives

addition, schematized in Fig. 8, has been identified as the rate-controlling step. It may be pointed out that the differences in the activation energies associated with this step for various investigated cytosine derivatives are in agreement with the deamination rate constants experimentally observed [107].

Since the nucleobases, either in their amino or imino form, constitute the reactant of the rate-determining step, the origin of the difference in sensitivity of the different forms of cytosine to the deamination reaction rate has to be searched among reactivity differences of the bases themselves, either global differences or local differences of their carbon C4 which undergoes the nucleophilic addition during the rate-determining step. In that purpose several global indexes and C4-local indexes computed for the different cytosine derivative are reported in Tables 15 and 16 respectively, both for the amino and imino forms of each base.

For 5,6-unsaturated forms of cytosine, methylation at position 5 tends to increase the chemical potential and to decrease, though only slightly for the imino form, the hardness. As a consequence the electrophilic power is quasi unchanged. It may be noted that this applies for both the amino and imino forms.

For 5,6-saturated forms of cytosine, methylation at position 5 has almost no effect, either on the chemical potential or on the hardness. And therefore this is also true for the electrophilic power. On the other hand, it has to be noted that 5,6-saturated derivatives of cytosine have higher chemical potentials than their 5,6-unsaturated counterparts and also much higher hardnesses, resulting in substantial lower electrophilic powers. This is in contradiction with the fact that methylation at position 5 and saturation of the C5–C6 bond both enhance the deamination

Table 15 Chemical potential, hardness, and electrophilic power of cytosine (Cyt), 5-methylcytosine (5mCyt), 5,6-dihydrocytosine (dhCyt), and 5,6-dihydro-5-methylcytosine (dh5mCyt) in their amino and imino tautomeric forms

Amino form				Imino form			
	μ (eV)	η (eV)	ω (eV)		μ (eV)	η (eV)	ω (eV)
Cyt	−3.70	5.38	1.28	Cyt	−3.80	5.41	1.34
5mCyt	−3.55	5.25	1.20	5mCyt	−3.67	5.38	1.25
dhCyt	−3.49	6.08	1.00	dhCyt	−3.56	6.72	0.94
dh5mCyt	−3.52	6.02	1.03	dh5mCyt	−3.60	6.82	0.95

Table 16 Electrostatic charge on carbon C4 (CHELPG method) [108] and grand canonical dual descriptor ($\frac{\Delta f_{C4}}{\eta^2}$ term) condensed (*through Chamorro and Perez scheme*) on carbon C4 for cytosine (Cyt), 5-methylcytosine (5mCyt), 5,6-dihydrocytosine (dhCyt), and 5,6-dihydro-5-methylcytosine (dh5mCyt) in their amino and imino tautomeric forms

Amino form			Imino form		
	q_{C4}	$\frac{\Delta f_{C4}}{\eta^2}$ (10^{-2} eV^{-2})		q_{C4}	$\frac{\Delta f_{C4}}{\eta^2}$ (10^{-2} eV^{-2})
Cyt	0.91	0.63	Cyt	0.73	0.29
5mCyt	0.74	0.73	5mCyt	0.63	0.40
dhCyt	0.86	1.03	dhCyt	0.65	0.61
dh5mCyt	0.70	1.02	dh5mCyt	0.56	0.58

rate. In conclusion, the difference in sensitivity of the different forms of cytosine to deamination must have a local origin.

Methylation at position 5 and saturation of the C5–C6 bond both tend to decrease the partial charge hold by carbon C4, indicating that the rate-controlling step is not under electrostatic control. Saturation of the C5–C6 double bond increases considerably the $\frac{\Delta f_{C4}}{\eta^2}$ value, in agreement with the fact that 5,6-saturated derivatives of cytosine undergo deamination much faster than the unsaturated forms.

As mentioned earlier, the grand canonical dual descriptor is a size consistent local index, suitable for comparing the reactivity of chemical systems of different sizes (as the different forms of cytosine) for a reaction under charge transfer control. This suggests that the nucleophilic addition of water to carbon C4 is under charge transfer control. This is supported by the fact that methylation at position 5 increases the value for 5,6-unsaturated forms of cytosine, while a decrease is observed for saturated compounds, in agreement with the effect of C5 methylation on the deamination reaction rate, which has opposite effects on saturated and unsaturated forms of cytosine.

Then, it can be concluded that the differences in sensitivity of the different forms of cytosine derivative to spontaneous deamination are closely related to modifications of the electrophilic power of their carbon C4 that is attacked during the reaction.

4 Conclusions

This chapter aims to review the use of DFT-based descriptors for understanding the chemical behavior of DNA bases. This approach can provide important information about the mechanism of formation of important DNA lesions and their occurrences. The usefulness of global DFT descriptors for investigating and predicting the reactivity of the bases is illustrated for several kinds of chemical situations. In parallel, local descriptors, specially the dual descriptor, help monitor the more reactive atomic sites within the bases. For instance, for the ground state, the dual descriptor gives a rationale for understanding the differences in reactivity of carbon C8 of adenine and guanine. Besides, the evolutions of the condensed values of the grand canonical dual descriptor of carbon C4 of Cyt derivatives clearly relate to the deamination rates of these compounds. Reactions between excited and ground state molecules are investigated by combining the dual potential and the dual descriptor respectively. This model accounts for the difference in formation efficiency between the formations of pyrimidine dimers. A general conclusion can be drawn from these investigations. Basically, the use of conceptual DFT descriptors is particularly efficient to explain qualitative differences of reactivity and selectivity between molecules, while quantum computations provide precise information but solely for a very specific set of chemical reagents. For an overall picture of a chemical reaction, it is interesting to combine the use of quantitative (computational reactivity studies) and qualitative (conceptual DFT approach) methods.

Acknowledgments All the authors thank INSERM: "This research has benefited from ITMO cancer of Aviesan within the framework Plan Cancer 2009–2013." Cette recherche a bénéficié de l'aide de l'ITMO cancer d'Aviesan dans le cadre du Plan Cancer 2009–2013.

References

1. Lindahl T (1993) Nature 362:709–715
2. Cadet J (1994) In: Hemminki K, Dipple A, Shuker DEG, Kadlubar FF, Segerbäck D, Bartsch H (eds) DNA adducts, identification and biological significance. IARC Scientific Publication, Lyon, pp 125–245
3. Gentil A, Cabrak-Neto JB, Mariage-Samson R, Margot A, Imbach J-L, Rayner B, Sarasin A (1992) J Mol Biol 227:981
4. Carell T, Brandmayr C, Hienzch A, Müller M, Pearson D, Reiter V, Thoma I, Thumbs P, Wagner M (2012) Angew Chem Int Ed 51:7110
5. Branco MR, Ficz G, Reik W (2012) Nat Rev Genet 13:7
6. Kkriaucionis S, Heintz N (2009) Science 324:929
7. Tahliani M, Koh KP, Shen YH, Pastor WA, Bandukwala H, Brudno Y, Agarawal S, Iyer LM, Liu DR, Aravind L, Rao A (2009) Science 324:930
8. Shapiro R, Klein RS (1966) Biochemistry 5:2358
9. Lindahl T, Lindahl B (1974) Biochemistry 13:3405
10. Shen JC, Rideout WM III, Jones PA (1994) Nucleic Acids Res 22:972
11. Lindahl T (1982) Ann Rev Biochem 51:61

12. Neddermann P, Jiricny J (1993) J Biol Chem 268:21218
13. Cadet J, Loft S, Olinski R, Evans MD, Bialkowski K, Wagner JR, Dedon PC, Moeller P, Greenberg MM, Cooke MS (2012) Free Radic Res 46:367
14. Murphy MP (2009) Biochem J 417:1
15. Ferguson LR (2010) Mutat Res 690:3
16. Park JB (2003) Exp Mol Med 35:325
17. von Sonntag C (2006) Free-radical-induced DNA damage and its repair – a chemical perspective. Springer, Berlin/Heidelberg/New York
18. Cadet J, Douki T, Ravanat J-L (2010) Free Radic Biol Med 49:9–21
19. Cadet J, Ravanat J-L, TavernaPorro M, Menoni H, Angelov D (2012) Cancer Lett 327:5
20. Dedon PC (2008) Chem Res Toxicol 21:206
21. Cadet J, Wagner J-R (2013) Cold Spring Harb Perspect Biol 5:a012559
22. Beckman JS, Beckman TW, Chen J, Marshall PA, Freeman BA (1990) Proc Natl Acad Sci U S A 87:1620
23. Denicola A, Freeman BA, Trujillo M, Radi R (1996) Arch Biochem Biophys 333:49
24. Medinas DB, Cerchario G, Trindale DF, Augusto O (2007) IUBMB Life 59:255
25. Crean C, Uvaydov Y, Geacintov N, Shafirovich V (2008) Nucleic Acids Res 36:742
26. Perrier S, Hau J, Gasparutto D, Cadet J, Favier A, Ravanat J-L (2006) J Am Chem Soc 128:5703
27. Cadet J, Douki T, Ravanat J-L (2008) Acc Chem Res 41:1075
28. Madugundu GS, Wagner JR, Cadet J, Kropachev K, Yun BH, Geacintov NE, Shafirovich V (2013) Chem Res Toxicol 26:1031–1033
29. Cadet J, Ravanat J-L, Martinez GR, Medeiros MH, Di Mascio P (2006) Photochem Photobiol 82:1219
30. Epe B (2012) Photochem Photobiol Sci 11:98
31. Cadet J, Sage E, Douki T (2005) Mutat Res 571:3
32. Cadet J, Mouret S, Ravanat J-L, Douki T (2012) Photochem Photobiol 88:1048
33. Courdavault S, Baudouin C, Charveron M, Canghilem B, Favier A, Cadet J, Douki T (2005) DNA Repair (Amst) 4:836
34. Wallace SS, Murphy DL, Sweasy JB (2012) Cancer Lett 327:73
35. Friedberg EC (2001) Nat Rev Cancer 1:22
36. Heil K, Pearson D, Carell T (2011) Chem Soc Rev 40:4271
37. Budden T, Bowden NA (2013) Int J Mol Sci 14:1132
38. Batista FZ, Kaina B, Menegheni R, Menck CFM (2009) Mutat Res 681:197
39. Wang Y (2008) Chem Res Toxicol 21:276
40. Belmadoui N, Boussicault F, Guerra M, Ravanat J-L, Chatgilialoglu C, Cadet J (2010) Org Biomol Chem 8:3211
41. Chatgilialoglu C, Ferreri C, Terzidis MA (2011) Chem Soc Rev 40:1368
42. Geerlings P, De Proft F, Langenaeker W (2003) Chem Rev 103:1793
43. Chermette H (1999) J Comput Chem 20:129
44. Ayers PW, Anderson JSM, Bartolotti JL (2005) Int J Quantum Chem 103:1793
45. Senet P (1996) J Chem Phys 105:6471
46. Parr RG, Donnelly RA, Levy M, Palke WE (1978) J Chem Phys 68:3801
47. Parr RG, Pearson RG (1983) J Am Chem Soc 105:7512
48. Parr RG, Yang W (1984) J Am Chem Soc 106:4049
49. Geerlings P, Sablon N, Fievez T, De Proft F (2010) Abs Papers Am Chem Soc 240:212
50. Fuentealba P, Parr RG (1991) J Chem Phys 94:5559
51. Morell C, Grand A, Toro-Labbé A, Chermette H (2013) J Mol Model 19:2893–2900
52. Iczkowski RP, Margrave JL (1961) J Am Chem Soc 83:3547
53. Mulliken RS (1934) J Chem Phys 2:782
54. Yang W, Parr RG (1985) Proc Natl Acad Sci U S A 82:6723
55. Pearson RG (1983) J Am Chem Soc 85:3533
56. Pearson RG (1987) J Chem Educ 64:561

57. Parr RG, Zhou Z (1993) Acc Chem Res 26:256
58. Parr R, von Szventpaly L, Liu S (1999) J Am Chem Soc 121:1992
59. Klopman G (1968) J Am Chem Soc 90:223
60. Salem L (1968) J Am Chem Soc 90:543
61. Fukui K (1957) J Chem Phys 27:1247
62. Fukui K (1987) Science 218:747
63. Yang W, Mortier WJ (1986) J Am Chem Soc 108:5708
64. Hocquet A, Toro-Labbé A, Chermette H (2004) J Mol Struct (THEOCHEM) 686:213
65. Cadet J, Grand A, Morell C, Letelier JR, Montcada JL, Toro-Labbé A (2002) J Phys Chem A 107:5334
66. Fievez T, Weckhuysen BM, De Proft F, Geerlings P (2009) J Phys Chem C 113:19905
67. Muya JT, De Proft F, Geerlings P, Nguyen MT, Ceulemans A (2011) J Phys Chem A 115:9609
68. Sablon N, De Proft F, Geerlings P (2010) J Chem Phys Lett 1:1228
69. Yang W, Cohen A, De Proft F, Geerlings P (2012) J Chem Phys 136:144110
70. Sablon N, De Proft P, Sola M, Geerlings P (2012) Phys Chem Chem Phys 14:3960
71. Morell C, Grand A, Toro-Labbé A (2005) J Phys Chem A 109:205
72. Morell C, Grand A, Toro-Labbé A (2006) Chem Phys Lett 425:342
73. Berkowitz M, Parr RG (1988) J Chem Phys 88:2554
74. Ayers PW, Morell C, De Proft F, Geerling P (2007) Chem Eur J 13:8240
75. Morell C, Labet V, Grand A, Ayers PW, Geerlings P, De Proft F, Chermette H (2009) J Chem Theory Comput 5:2274
76. Morell C, Labet V, Ayers PW, Genovese L, Grand A, Chermette H (2011) J Phys Chem A 115:8032
77. Tognetti V, Morell C, Ayers PW, Joubert L, Chermette H (2013) Phys Chem Chem Phys 15:14465–14475
78. Kumar V, Kishor S, Ramaniah LM (2012) J Mol Model 18:3969
79. Ciino P, Gomez-Paloma L, Barone V (2004) J Org Chem 69:7414
80. Sivanesan D, Subramanian V, Nair BU (2001) J Mol Struct (THEOCHEM) 544:123
81. Saha S, Wang F, MacNaughton JB, Moewes A, Chang DP (2008) J Synchrotron Radiat 15:151
82. Saha S, Roy RK (2007) J Phys Chem B 111:9664
83. Parthasarti R, Amutha R, Subramanian V, Nair BU, Ramasami T (2004) J Phys Chem A 108:3817
84. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09, Revision A.02. Gaussian, Wallingford
85. Chamorro E, Perez P (2005) J Chem Phys 123:114107
86. Box HC, Dawidzik JD, Budzinski EE (2001) Free Radic Biol Med 31:856
87. Cadet J, Berger M, Douki T, Ravanat J-L (1997) Rev Physiol Biochem Pharmacol 131:1
88. Grand A, Morell C, Labet V, Cadet J, Eriksson LA (2007) J Phys Chem A 111:8968–8972
89. Akin M (2010) Thesis of the Georgia Institute of Technology, Atlanta
90. Inostroza-Rivera R, Herrera B, Toro-Labbé A (2012) Comput Theor Chem 990:222
91. Wagner JR, Cadet J (2010) Acc Chem Res 43:564
92. Bellon S, Ravanat J-L, Gasparutto D, Cadet J (2002) Chem Res Toxicol 15:598

93. Dizdaroglu M, Jaruga P, Rodriguez H (2001) Free Radic Biol Med 30:774
94. Jaruga P, Birincioglu M, Rodriguez H, Dizdaroglu M (2001) Biochemistry 41:3703
95. Zhang RB, Eriksson LA (2006) Chem Phys Lett 417:303–308
96. Xerri B, Morell C, Grand A, Cadet J, Cimino P, Barone V (2006) Org Biomol Chem 4:3986–3992
97. Labet V, Morell C, Grand A, Cadet J, Cimino P, Barone V (2008) Org Biomol Chem 6:3300
98. Labet V, Grand A, Morell C, Cadet J, Eriksson LA (2008) Theor Chem Acc 120:429
99. Douki T, Cadet J (2001) Biochemistry 40:2495
100. Ehrlich M, Norris KF, Wang RY-H, Kuo KC, Gehrke CW (1986) Biosci Rep 6:387
101. Slae S, Shapiro R (1978) J Org Chem 43:1721
102. Douki T, Cadet J (1992) J Photochem Photobiol B 15:199
103. Douki T, Cadet J (1994) Biochemistry 33:11942
104. Frederico LA, Kunkel TA, Shaw BR (1990) Biochemistry 29:2532
105. Labet V, Morell C, Cadet J, Eriksson LA, Grand A (2009) J Phys Chem A 113:2524
106. Labet V, Morell C, Douki T, Cadet J, Eriksson LA, Grand A (2010) J Phys Chem A 114:1826
107. Grand A, Cadet J, Eriksson LA, Labet V, Jorge NL, Schreiber ML, Douki T, Morell C (2012) Theor Chem Acc 131:1187
108. Breneman CM, Wiberg KB (1990) J Comp Chem 11:361

Characterising Heterocyclic Rings Through Quantum Chemical Topology

Mark Z. Griffiths and Paul L.A. Popelier

Contents

M.Z. Griffiths • P.L.A. Popelier (✉)
Manchester Institute of Biotechnology (MIB), 131 Princess Street, Manchester M1 7DN, UK

School of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, UK
e-mail: Paul.Popelier@manchester.ac.uk

F. De Proft, P. Geerlings (eds.), *Structure, Bonding and Reactivity of Heterocyclic Compounds*, Topics in Heterocyclic Chemistry 38, DOI 10.1007/978-3-642-45149-2_3

Abstract In this chapter, ring structures are characterised through their calculated properties within the theory of quantum chemical topology (QCT). QCT properties of the atoms within a ring can predict the properties at a special point, a so-called ring critical point (RCP). Both the RCP properties and the atomic properties according to QCT successfully distinguish between different ring structures. There are four features of a ring that are responsible for its ring atom properties: (i) the number of heteroatoms within the ring, (ii) the heteroatom's element, (iii) the ring substituent, and (iv) the substituent site. Interestingly, the four features affect the ring's properties independently. Therefore, a change in a heteroatom's element will always affect the ring's properties the same amount, irrespective of other ring features. This is called *ring characteristic orthogonality*. For substituent types, it is the atom of the substituent connecting the substituent to the ring that dominates the effect on the ring, rather than the entire substituent itself. Using these relationships between ring features and QCT properties opens up the possibility of improving ring structures in areas such as drug design.

Keywords Molecular similarity • Quantum chemical topology • Quantum theory of atoms in molecules • Ring critical point • Atomic charge • Substituent effect

1 Introduction

A ring's structure determines how it behaves in different environments. Whether a ring is in aqueous solution, inside an organism, or in the presence of reagents for organic synthesis, it is the (electronic) structure of the ring that determines its behaviour. Therefore, understanding how the structure and behaviour of a ring are linked is an attractive prospect, with implications in areas such as drug design, organic synthesis and protein action. Prediction within the aforementioned areas becomes much more viable if a link between structure and behaviour can be explained through a theoretical viewpoint. Such a link can be established by successfully characterising a ring's structure through calculated properties. Once ring systems are successfully characterised, then alterations can be made to rings to create new molecules with more desirable properties. Such alterations to a molecule's structure are documented in an extensive literature [1–9]. The structure of a

ring refers to its electronic and three-dimensional structure. The three-dimensional structure of a ring includes features such as the number of atoms within a ring, the element of the atoms, their positions in the rings and their connectivity. The most complete representation of electronic structure is determined through the quantum chemical wave function.

The ring structures studied in this contribution are all five-membered heterocyclic rings found in a wide range of important molecules including fungicides [10], sedatives [11], biosynthetic precursors [12], homogeneous catalysts [13], antifungal drugs [11] and multiple COX-2 inhibitors [14–16].

Quantum mechanical characterisation of a ring provides a potential basis for quantitative ring similarity and prediction of behaviour. Characterisation can only be achieved if a relationship is found between structure and properties, which leads to the question: how do the features of a ring determine its properties? As the three-dimensional structure (or features) of a ring changes, which includes replacement of atoms within the ring, then the electronic structure must change with it. Therefore, changes to the features of a ring must be expressed by changes in the ring's electronic structure. A relationship between features and electronic structure can answer the question posed above because the ring's properties are dependent on its electronic structure.

When characterising the structure of a ring, individual atoms of the ring tend to be ignored and the ring is only considered as a whole. Isolating single atom properties is desirable when considering the effects of a ring's features on its properties. However, such a treatment can be difficult as the ring tends to be expressed as a single entity only. There are numerous studies on the electronic structures of rings where only the entire ring is considered such as calculation of aromaticity [17–21] and studies of delocalisation of electron density throughout the ring [22, 23]. However, these studies lose information regarding the contributions of individual atoms within a ring to its properties. This contribution aims to incorporate individual atom properties into the characterisation of rings to establish how the features of a ring affect its properties. This is achieved through quantum chemical topology (QCT), which partitions a molecule into its constituent atoms.

2 Background

2.1 Quantum Chemical Topology

QCT refers to a body of work that uses the idea of a gradient vector field to extract chemical insight from modern wave functions or high-resolution experimental electron densities. The name QCT was first proposed [24] in 2003 in an attempt to bundle results that had been obtained under separate headers, the oldest and most important being the quantum theory of atoms in molecules (QTAIM). This theory, proposed by Bader and co-workers, is in fact a generalisation of quantum mechanics to subspaces. Such a subspace can be identified with a topological atom (or QCT

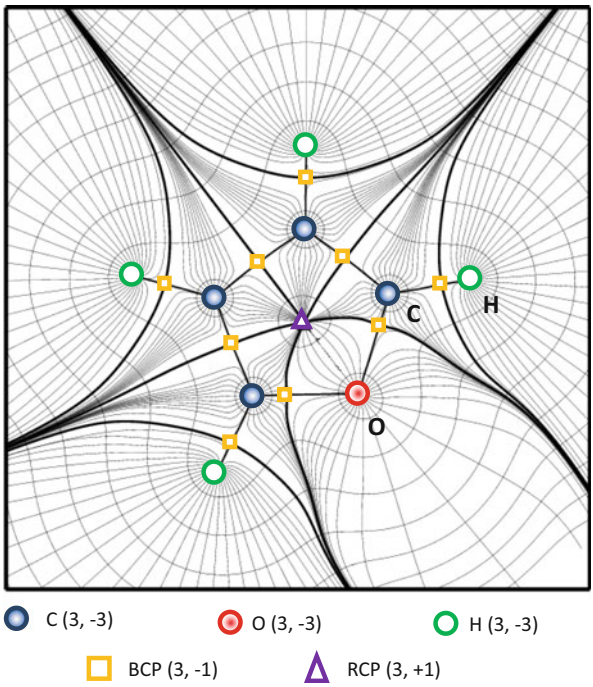


Fig. 1 Gradient vector field in the molecular symmetry plane of furan, showing interatomic surfaces, atomic interaction lines and critical points

atom), which naturally arises by the electron density partitioning itself via its gradient vector field. How this happens is discussed below. The application of QCT has grown into many fields, including Quantitative Structure Activity Relationship (QSAR) [25], molecular dynamics [26], free radicals [27], electron delocalisation [19] and hydrogen bonding [28, 29].

QCT partitions a molecule into (topological) atoms starting from an electron density, which in this work is calculated ab initio. It is best to refer to an actual example while explaining the concepts used in this contribution. Figure 1 shows the electron density in the molecular plane of furan, which is a planar five-membered heterocyclic aromatic containing a single oxygen atom. Plotting all points in space with the same electron density value gives isodensity envelopes shown as contour lines in Fig. 1. We can now trace a network of trajectories that are locally orthogonal to the set of constant electron density contour lines at any point. This network is called the gradient vector field, as shown in Fig. 1. This vector field consists of so-called gradient paths, which are trajectories of steepest ascent in the electron density. In other words, if one follows a gradient path, then one will reach, after each step, the highest local electron density in the quickest possible way. Note that a gradient path has a direction, i.e. it has a starting point and an end point.

The majority of gradient paths start at infinity and terminate at a point where the gradient of the density vanishes, known as a critical point. These stationary points are marked as squares, circles or a triangle in Fig. 1, depending on the type of critical point. The difference between types can be explained as follows. Taking the

second derivative of the electron density at a critical point gives a three-by-three matrix known as the Hessian. The Hessian is symmetric and after diagonalisation yields three eigenvalues denoted λ_1, λ_2 and λ_3, which are the local curvatures in the electron density. When summed, they give the value of the Laplacian of the electron density. The type of critical point can be categorised by the signs (positive or negative) of the three eigenvalues. Critical points are written in the form (ω, Z) where ω is the number of nonzero eigenvalues and Z is the sum of the signs of the eigenvalues where a negative value is represented by -1 and positive by +1. A critical point at the nucleus (circle in Fig. 1) has three eigenvalues that are all negative and therefore summarised as (3, −3). A bond critical point (BCP), shown as squares, exists at a saddle point in the electron density between two nuclei and is marked by (3, −1) because it has two negative eigenvalues and one positive eigenvalue. A ring critical point (RCP) (triangle) or (3, +1) critical point appears in the centre of the furan ring in Fig. 1. In general, there also exist cage critical points (3, +3), but they will not be discussed further because they do not feature in this work.

At an RCP, λ_2 and λ_3 are both positive while λ_1 is negative, and by convention the eigenvalues follow the pattern $\lambda_3 > \lambda_2 > \lambda_1$. The shape of the electron density at a given RCP can be calculated by an RCP ellipticity, which we define here. Note that this ellipticity is analogous to (but not to be confused with) the well-known ellipticity defined at a BCP. The BCP ellipticity is calculated as $(\lambda_1/\lambda_2) - 1$, whereas the ellipticity at the RCP is calculated as $(\lambda_3/\lambda_2) - 1$. The latter attains the minimum value of 0, when it is evaluated at the centre of a cylindrically symmetric electron density. The electron density of a benzene ring approaches such an electron density of this symmetry, as proven by the very low value of the RCP ellipticity of approximately 2×10^{-4}. In previous studies the BCP ellipticity was one of three properties at the BCP that successfully characterised different bonds [30]. In the same vein the RCP ellipticity is proposed in this work to characterise a ring.

2.2 Partitioning of the Electron Density

A molecule is partitioned into its constituent atoms through the gradient vector field, which is a collection of an infinite number of gradient paths. Figure 1 shows the gradient vector field of furan in its molecular symmetry plane. The vast majority consists of gradient paths that originate at infinity and terminate at a critical point. Figure 1 contains numerous special gradient paths that are not part of this majority. An example of a special gradient path is one that connects two critical points, such as the (3, −3) at the nucleus (circle) and a BCP (square). Another example is a gradient path that connects the RCP (triangle) to a BCP (square). Two gradient paths that each begin at the same BCP but terminate at two different nuclei form a so-called atomic interaction line (AIL). There are many examples in Fig. 1, and they collectively form the molecular graph of furan, which recovers the expected Lewis diagram for this molecule. While most typical gradient paths terminate at a

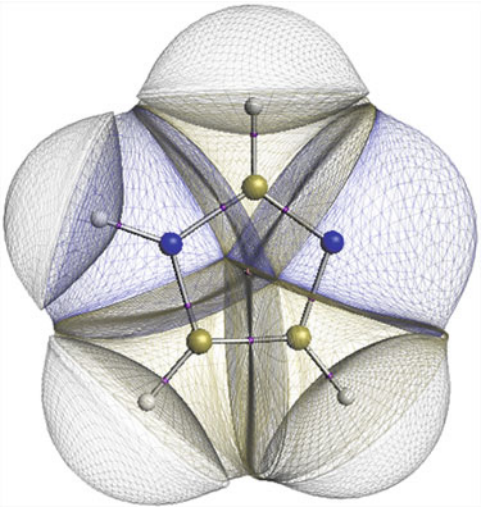


Fig. 2 Three-dimensional representation of topological atoms in imidazole with a 0.001 a.u iso-electron density boundary cut-off

nucleus, there are gradient paths that originate at infinity but terminate at a BCP. Collectively these gradient paths form the so-called interatomic surface (IAS). An intersection of such an IAS with the plotting plane is shown in Fig. 1, for example, as one of the bold lines that bisects the carbon oxygen AIL at the BCP. An IAS sharply divides the electron density between the two topological atoms that it bounds. In other words, the electron density at one of the IAS strictly belongs to one atom and the electron density at the other side to the other atom. Topological atoms do not overlap or penetrate each other. Moreover, they leave no gaps between them. From Fig. 1 it is clear that an atom is a subspace that consists of gradient paths that are all attracted to its nucleus.

Figure 2 displays the gradient vector field in three dimensions and shows that each atom has a distinct topological volume (known as the atomic basin) within the molecule. Note that, topologically, a QCT atom extends to infinity when it is not bound by other atoms and therefore a cut-off value of 0.001 a.u. iso-electron density is used to cap the atoms in Fig. 2.

The atomic basin (Ω) can be considered as a "volume of influence" of an atom within a given molecule. Volume integration of a property density over an atomic basin gives the atom's property within the molecule. For example, if this property density is the electron density itself, then the volume integration results in the atom's electronic population. If corrected for the nuclear charge, this population leads to the QCT charge of an atom. This is discussed in more detail below. Clearly, whichever the property density being integrated, the integral yields the atom's contribution towards the molecular property of interest. For example, if the property density is simply unity, as in Eq. 1,

$$v(\Omega) = \int_{\Omega} d\tau, \tag{1}$$

then the atom's volume is obtained, where $d\tau$ is an infinitesimal volume element. The volume of the molecule, or a fragment thereof, is then simply obtained by adding the relevant atomic volumes. This perfect additivity is very convenient in any QCT analysis of properties.

A number of other atomic properties feature in this contribution. One type of atomic kinetic energy is defined by Eq. 2,

$$\mathrm{K}(\Omega) = -\frac{1}{4} N \int_{\Omega} d\tau' \left[\psi^* \nabla^2 \psi + \psi \nabla^2 \psi^*\right] \tag{2}$$

where $N\int_{\Omega} d\tau'$ is a shorthand notation referring to an integration (over all space) over all electrons except one, which is integrated over the topological atom's volume. Note that this compact notation also sums over the spins of the electrons. Topological atoms have the desirable property that they have a well-defined kinetic energy, which cannot be said of arbitrary subspaces within a molecule.

Because a topological atom has its own virial theorem [31], which expresses a balance between its kinetic and potential energy, the *total* energy of an atom can be expressed using its kinetic energy only, as shown in Eq. 3,

$$E(\Omega) = -\int_{\Omega} d\tau \mathrm{K}(\mathbf{r}) \tag{3}$$

where $\boldsymbol{r}$ is a position vector of the kinetic energy density $\mathrm{K}(\mathbf{r})$. Again, summation of an atomic property over all atoms in a molecule produces the molecular property. Therefore, a property for a molecular fragment, such as a functional group or the set of ring atoms, is simply obtained by summing the constituent atoms. Any atomic property can be calculated by integrating the property function over the basin [32]; this is achieved through different QCT packages [33, 34].

2.3 Electrostatics

The molecular electrostatic potential is an important property, which has been widely utilised in studies on intermolecular interactions [35–37] and hydrogen bonding [38–41], for example. The molecular electrostatic potential can be physically understood by imagining a unit charge that moves around from one position to another, probing the molecular charge density. We assume that the charge

density is not polarised by this probing charge. For each position **r** of the charge, a new interaction energy will be obtained. This energy, divided by the unit charge, is the molecular electrostatic potential. It can be written as

$$U_{ele}(\mathbf{r}) = \sum_{i=1}^{N} \frac{Z_i}{|\mathbf{r} - \mathbf{R}_i|} - \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{4}$$

where $\rho(\mathbf{r}')$ is the electron density and the first term takes care of the nuclear potential, for a molecule of N atoms, with nuclear positions $\mathbf{R}_i$.

Evaluating the second term in Eq. 4 at every single point is time consuming. A Taylor series expansion can be used to reduce the cost of evaluating the electrostatic potential at every point. This expansion essentially factorises the expression $|\mathbf{r} - \mathbf{r}'|$ such that a volume integral in $\mathbf{r}'$ can be pre-computed independently of the **r** coordinate. The expansion creates terms called "multipole moments", which are capable of representing the electrostatics exactly, provided the multipolar series expansion converges. More precisely, one can place multipole moments on the nuclear positions and have them generate the molecular potential via the multipolar expansion. The only disadvantage of this multipolar expansion is possible convergence problems, but an advantage of topological atoms is that they occupy a finite volume, and hence formal convergence is possible [42]. The higher-order terms have a smaller effect on the result, so it is possible to accurately represent the electrostatics using only the first few multipole moments of the expansion [43]. Each multipole moment can be categorised by its rank "l", where $l = 0$ for a monopole moment, $l = 1$ for a dipole moment, $l = 2$ for a quadrupole moment and so on.

The first term of the expansion is the monopole moment given by

$$Q_{mono}(\Omega) = -\int_{\Omega} d\tau \rho(\mathbf{r}) \tag{5}$$

The monopole represents the electronic population of the atom, which when corrected for the nuclear charge is the atomic charge. The second term of the expansion involves the three dipole moments, one of which is given by

$$Q_{dip,x}(\Omega) = -\int_{\Omega} d\tau x \rho(\mathbf{r}) \tag{6}$$

where x is referring to a global frame. The third term of the expansion involves only five rather than six quadrupole moments because we use the spherical tensor formalism rather than the Cartesian one. The five components of the quadrupole moment correspond to the well-known d orbitals from quantum mechanics and read $3z^2 - r^2$, $x^2 - y^2$, xy, xz and yz. The quadrupole moment represents how the electron density deviates from a spherical distribution [44]. The next term involves seven (i.e. $2\ell+1 = 7$) octupole moments, but they do not feature in this work.

2.4 Partial Least Squares (PLS)

The large number of QCT properties makes linear correlations difficult. This problem is solved with partial least squares (PLS) regression [45, 46]. PLS is a multivariate linear regression method. A set of independent variables are correlated to dependent variables. The dimensionality of the independent variables is reduced by a linear combination of these variables creating so-called latent variables. The set of latent variables describes the maximum variance in the independent variables where the first latent variable explains the maximum variance in the data set, the second latent variable explains the remaining variance in the data set and so on. A predictive model is created through the combination of these latent variables and therefore reducing the dimensionality. Through PLS multiple independent properties can be correlated to dependent properties.

3 Quantum Chemical Topology of Cyclic Molecules

3.1 Ring Atom Properties and the Ring Critical Point

The QCT partitioning of a molecule lends itself to localise chemical information in rings and their constituent atoms because individual atoms within the ring are naturally defined. Previous QCT work on rings calculated properties such as the delocalisation of the electron density onto the ring [18, 23] or the extent of aromaticity calculated for the ring [17–21], but the QCT atoms themselves are rarely considered. The properties of a given QCT atom are influenced by other atoms in the molecule. Atomic properties respond to changes in environment and therefore the position of an atom within a ring and its neighbours determine the atom's properties.

Figure 3 visualises the changes observed in QCT atoms relative to their environment (or position in the ring). It is convenient to normalise the monopole moments of all five ring atoms, by setting the smallest and largest value to 0 and 1, respectively. The normalised monopole moments are represented by a colour on a scale that progresses through the rainbow from violet (most positive) to red (most negative).

The carbon in between the two nitrogens is found to be much more positive relative to the other carbons within the ring, while the -NH nitrogen is more negative than the = N- nitrogen. Interestingly, a difference in charges also appears between the two carbons adjacent to each other, although the difference is much more subtle. The carbon adjacent to the -NH nitrogen is marginally more negative than the carbon adjacent to the = N- nitrogen. This is because the more negative –NH nitrogen has a more pronounced effect on its adjacent atoms than the effect of the = N- nitrogen on its own adjacent atoms. These observations show that an atom's element and its environment both affect the atom's properties and these observations are reflected within the QCT atoms.

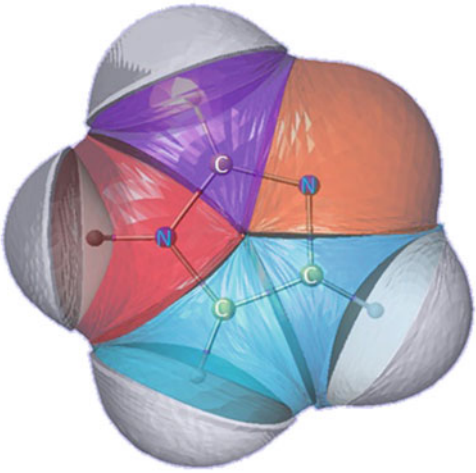


Fig. 3 Normalised monopole moments for the non-hydrogen atoms within imidazole, represented by *colours* in the rainbow spectrum where *violet* = most positive and *red* = most negative

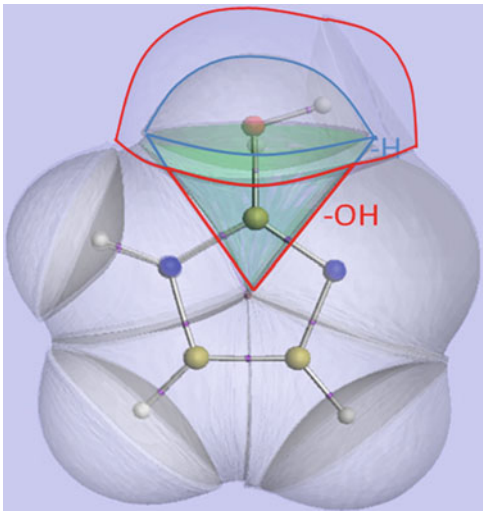


Fig. 4 Superposition of imidazole and imidazole-2-ol with the change in atomic basin highlighted for the carbon upon substitution of H (*blue* contour) with OH (*red* contour)

QCT atoms of a ring can also reflect any changes in substituents onto the ring atoms. A substitution onto a ring atom will change this atom's basin and therefore affect this atom's properties as well. As an example, Fig. 4 shows how the atomic basin of a carbon atom within an imidazole ring changes when the hydrogen bonded to the carbon is substituted for an alcohol group. The highlighted basin (green) is crushed upon substitution of the hydrogen with the alcohol, the basin changing from the blue outline to the red one. The oxygen of the alcohol has a much larger electronegativity than the hydrogen of the non-substituted imidazole, and, therefore, a larger portion of the electron density between the carbon and the oxygen is

Table 1 Non-substituted five-membered unsaturated rings in data set A

Furan	Isothiazolin	Phosphole	Thiazole
Furazan	Isoxazole	Pyrazole	Thiazoline
Imidazole	Isoxazoline	Pyrazoline	Thiophene
Imidazoline	Oxazole	Pyrrole	1,2,3-triazole
Isothiazole	Oxazoline	Tetrazole	1,2,4-triazole

attributed to the oxygen atom. The two nitrogen basins are also affected by the substitution as the carbon basin is pushed downwards, which causes a noticeable spread in the basin at either side.

4 Results and Discussion

4.1 Calculation of RCP and Ring Atom Properties

A series of five-membered heterocyclic ring structures were geometry optimised at the B3LYP/6-311 + G(2d,p) level of theory with GAUSSIAN03 [47]. The QCT properties were calculated from the resulting wave functions for the five atoms in a ring only. Note that the atomic ring properties listed below are always referring to the sum of these properties over all five atoms in the ring. Data set *A* consists of 20 molecules, which are shown in Table 1 as a series of five-membered ring structures without substituents. The structures in the set contain rings consisting of C, O, N, S and P atoms only, and all are unsaturated heterocyclics. The QCT atomic properties calculated were the volume, the total atomic energy *E*, the monopole moment and the magnitudes of the atomic dipole and quadrupole moments. The RCP properties calculated were the electron density ρ; its Laplacian $\nabla^2\rho$; the curvatures λ_1, λ_2, λ_3; and the RCP ellipticity ε.

4.2 Non-substituted Rings (Data Set A)

Finding a relationship between the ring atom properties and the RCP properties is achieved through partial least squares (PLS) regression. PLS reduces the complicated multivariable problem to a clearer picture, and the quality of the relationship is quantified through r^2 values for the predicted output variables correlated with the observed output variables. A series of PLS models were created, which differed in the selection of the input and output variables. For the first PLS model (Table 2), all five atomic properties were selected as inputs and all six RCP properties as the output variables.

Figure 5 focuses on the prediction of only one property, made by this first PLS model, which is the RCP electron density. This figure plots the actual observed RCP electron density against the predicted one, with an r^2 value of 0.91. This shows that

Table 2 r^2 values for PLS models with different selected input and output variables

Model	Input variables	Output variables	r^2 of data set A	r^2 of data set B
1	AP	RP	0.83	0.80
2	AP	RP except ellipticity	0.96	0.90
3	RP	AP	0.59	0.57
4	RP except ellipticity	AP	0.59	0.56
5	RP except ellipticity	Ellipticity	0.93	0.90

For clarity the RCP properties are highlighted. The third and fourth entries are the reverse of the first two entries

AP atom properties, *RP* RCP properties

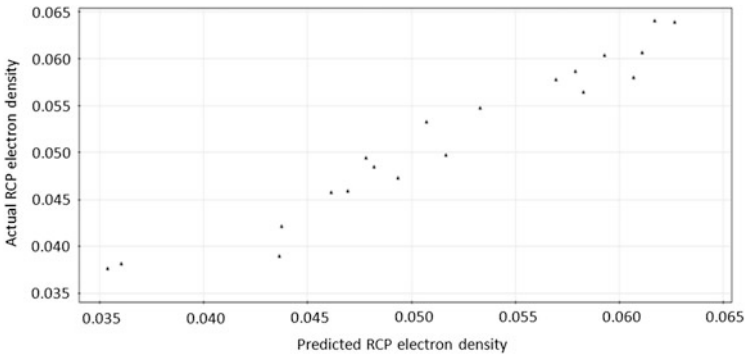


Fig. 5 Correlation ($r^2 = 0.91$) between the observed (actual) RCP electron density and the one predicted by the first PLS model (Table 2) for non-substituted five-membered heterocyclic molecules (data set A)

a clear relationship exists between the ring atom properties and the RCP properties. Therefore, it is possible to predict the RCP properties if the ring atom properties are known. The separate contributions towards the prediction of the output variables of the electrostatic and energy variables were determined through two different models, one model containing only electrostatic input variables and the other only energy input variables. For both models the output variables selected were all six RCP properties. The "electrostatic only model" (i.e. monopole, dipole and quadrupole moments) has an r^2 of 0.89. The "energy only model" has an r^2 of only 0.71. Therefore, it is the electrostatic variables that provide the greatest contribution towards predicting the RCP properties from the ring atom properties. This result is also confirmed by analysing the components of the first PLS model (combined electrostatic and energy model), which shows that the first component is dominated by the electrostatic terms.

4.3 Substituted Ring Systems (Data Set B)

Data set *B* is an expansion of data set *A*, which now includes mono-substituted rings. For each of the 20 molecules, the hydrogen atoms were sequentially replaced by a simple functional group. The substituents were –OH, $-NH_2$ and $-OCH_3$, creating a set of 178 mono-substituted ring molecules. As an example, Fig. 6 shows all six possible substituted rings (3 substitutions times 2 positions) for furan. The symmetry of furan makes substitution at the two remaining substitution sites unnecessary. The alcohol and amine groups were chosen because they are the simplest substituents that are found readily in many biologically relevant molecules. The methoxy group was chosen to elucidate both the effect of replacing the hydrogen of the alcohol with a methyl and the perception of this replacement by the ring atoms and RCP properties. Therefore, it is possible to determine the physical distance in 3D at which the atoms of the substituent cease to significantly affect the ring atom and RCP properties. This distance allows for truncation of the substituents and therefore reducing the number of substituents required for modelling. For example, if the effect, of a butyl and longer alkyl substituents, on the ring properties is comparable to the effects of a propyl, then the propyl is sufficient to represent all larger alkyl substituents. Note that the number of possible substituted sites varies from two to four, depending on the symmetry of the ring structure itself and the number of available substitution sites (determined by the atoms in the ring).

The PLS analysis was repeated for a new dataset (B) containing both non-substituted and substituted rings. As with the PLS models for data set *A*, variables included in the models were altered to assess the relationship of RCP properties to ring properties. The effect of reversing the ring atom and RCP properties as input and output variables of the model was also explored (PLS models 3 and 4).

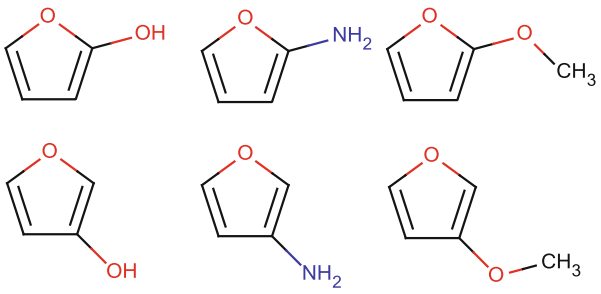


Fig. 6 The six possible substitution patterns in furan. Note that the symmetry of furan would create unwanted duplicates if it were substituted on the two remaining sites

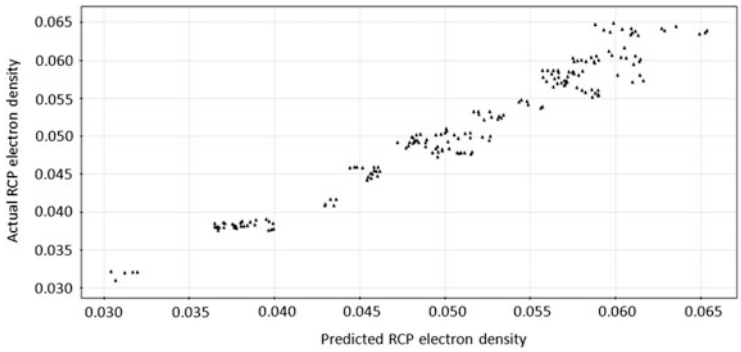


Fig. 7 PLS results for alcohol, ether, amine and non-substituted five-membered heterocyclic molecules

4.4 The Effects of a Ring's Atomic and RCP Properties

Figure 7, which is the counterpart of Fig. 5, shows the PLS model for all five atom properties as inputs and all six RCP properties as outputs for the data set *B* (PLS model 1).

The corresponding r^2 value is 0.80, whereas before, for dataset A (Fig. 5) it was 0.83. Table 2 shows the r^2 values for a variety of PLS models (constructed in sets *A* and *B*), which differ in their selected input and output variables. The r^2 values of the data set that includes substituted rings (B) are always lower correlation than those of the data set without substituents (A), but only by 0.03 on average. The models show that prediction of the ring atom properties from the RCP properties is much more difficult than the reverse, as the largest r^2 of these models is only 0.59. However, the r^2 value never falls below 0.80 for any model (PLS models 1 and 2) that predicts RCP properties from the ring atom properties. This suggests the existence of some relationship between the atomic properties of the ring atoms and the RCP properties, but the information contained in the RCP properties is not enough to predict the ring's atomic properties. Further discussions regarding the models will always refer to the r^2 values of data set *B* only because correlations for larger data sets are more reflective of the model of interest.

4.5 The Role of Ellipticity

Of the first four models that relate the atomic and RCP properties, the model with the largest r^2 is PLS model 2 (0.90) which has atomic properties as inputs and RCP properties as outputs, with ellipticity excluded. This suggests that the ellipticity at the RCP is the most difficult RCP property to predict. However, predicting the ring's atomic properties from the RCP properties both with and without the ellipticity produces similar correlations of 0.57 and 0.56, respectively. But then

again, an r^2 value of 0.57 is rather poor and prevents one to be sure about causality in the first place. The high r^2 value of 0.9 for the model with RCP properties (except ellipticity) as inputs and the ellipticity as an output (PLS model 5) is anticipated because the ellipticity is a nonlinear function of the curvatures λ_2 and λ_3 as explained in Sect. 2.1, that is, $\varepsilon_{RCP} = (\lambda_3/\lambda_2) - 1$. Although PLS is a linear correlator, it manages to capture most of this nonlinearity. In summary, the decrease in the r^2 value when ellipticity is included in the output variables suggests that it must be more sensitive, or receptive, to changes in the ring atoms.

To determine the relationship of the ellipticity and ring atom properties, a series of plots were constructed comparing a single RCP property to a single atomic property. Figure 8 compares each three (out of a total of six) RCP properties to two (out of a total of five) atomic properties, leading to $3 \times 2 = 6$ panels. Panels (a), (b) and (c) show the kinetic energy K(Ω) (which is more convenient than E(Ω); see Eq. 3) versus the electron density, its Laplacian and the ellipticity. Panels (d), (e) and (f) show the monopole against the same three RCP properties. Figure 8a–e shows that the electron density and its Laplacian have similar relationships to the ring atom properties. Indeed, the corresponding panels (8*a* and 8*b*; 8*d* and 8*e*) have the same general shapes. The general shape consists of three clusters: one small (top of panel), one medium (middle of panel) and one large (bottom of panel). The middle cluster has a subcluster directly beneath it and to the left of its centre. The small cluster (top) has the lowest RCP property (electron density or its Laplacian) followed by the medium cluster and finally the largest cluster. The increased spread in the ellipticities stretches the graphs along the horizontal axis and distorts the shape relative to the top four panels (8*a*, 8*b*, 8*d* and 8*e*). This increased spread marks a heightened sensitivity of the ellipticity towards the ring atom properties. In other words, if the ellipticity is predicted, it will demand a more accurate model than the electron density or its Laplacian will demand. This observation may explain why the r^2 value for a PLS model with the ellipticity included (model 1 in Table 2) is lower than a PLS model without it (model 2).

In order to discover if relationships between ring scaffolds can be extracted from the RCP properties and ring atom properties, the raw data were examined "locally", through four types of rings. Indeed, analysis of the whole of data set *B* can mask any local information, which may not be apparent when examining the data set in its entirety. The four types of rings are 1,2,3-triazole, 1,2,4-triazole, isoxazole and imidazole. The two triazoles were chosen because they differ by the position of their heteroatoms in the ring. Imidazole was chosen because it only has two nitrogens in the ring compared with the three in triazole and oxazole because it has two different heteroatoms in the ring (N and O).

Figure 9a–d shows the overall monopole moment of the whole ring (i.e. number of electrons) for the four ring types as a function of three possible RCP properties. Figure 9a shows that each ring type is discernible from the three others as there is no overlap between the four clusters. In other words, each ring type has a distinct signature of monopole moment and RCP electron density. For example, all 1,2,3-triazole molecules lie within a range 33.0–33.6 a.u. for their monopole moment and 0.059–0.063 a.u. for their electron density, while none of the other three rings have

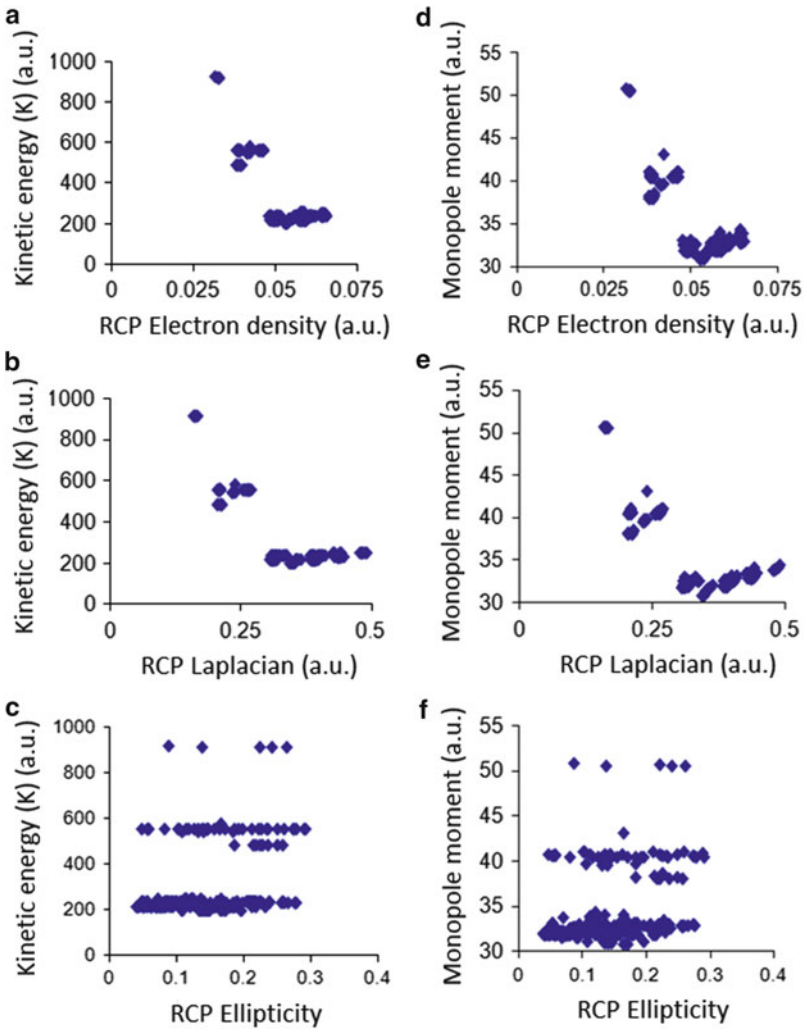


Fig. 8 Comparison of RCP properties with two different ring atomic properties (kinetic energy (K) and monopole) for data set B

any data points within this range. This unique range of properties for each ring also occurs for the Laplacian in Fig. 9b. However, there is an overlap between the 1,2,3-triazoles and 1,2,4-triazoles, which is not unexpected due to their close structural relationship. Again, the ellipticity recreates the general shape of the graphs for the other RCP properties but the shape is again stretched out (see Fig. 8). This stretching leads to a greater overlap between the ring clusters, due to the ellipticity's increased sensitivity to the variation of ring atoms, thus confirming the ellipticity's difficulty in discerning between ring types and also explaining why PLS models

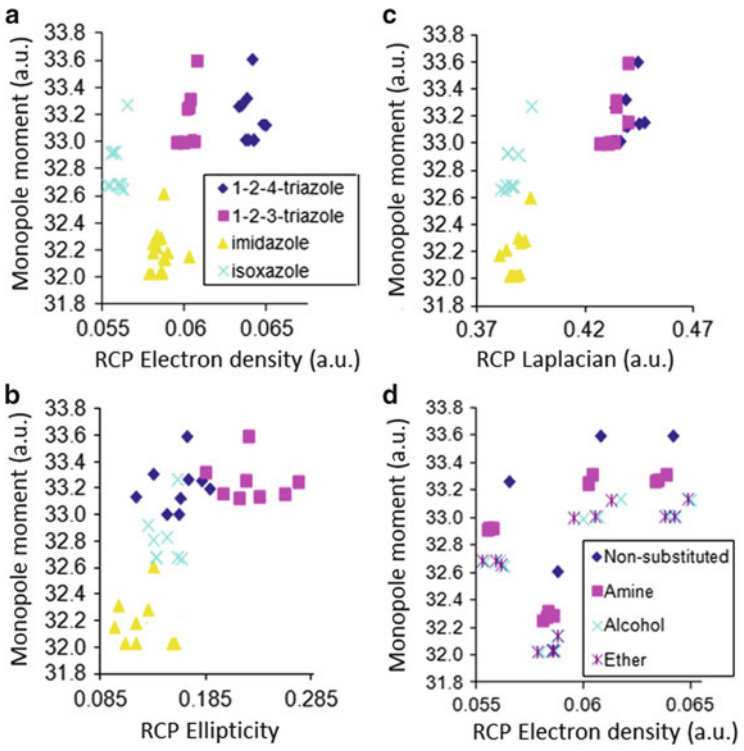


Fig. 9 Comparison of four different ring scaffolds with all substitution data points for each scaffold using ring atom monopole moments and RCP properties. The monopole moment and three different RCP electron density plot is drawn in two forms: (**a**) the data set divided into ring scaffolds, (**d**) and substitution types

struggle to predict the ellipticity. Figure 9d determines the effect of the substituents on a ring's properties. Figure 9d plots the same data as Fig. 9a but recoloured (or relabeled) such that data points are grouped by substituent type rather than by ring type. The patterns of the different substituents within each ring type are remarkably similar throughout all four ring types. This repeated pattern begins with a non-substituted ring at the top, followed by a set of amines just below it and followed again by the alcohol and methoxy substituents, further down. Interestingly, each alcohol data point is paired with a methoxy data point. These pairs are both substitutions onto the same available substitution site on the ring. Therefore, the RCP and ring atoms see the alcohols and ethers (i.e. methoxy) as very similar, suggesting that it is the atom of the substituent directly bonded to the ring that affects the properties more than the entire substituent.

5 Ring Atoms

5.1 Ring Atom Properties

In this section the ring's composition (marked by four indicators: number and type of heteroatoms, internal bonding structure and substitutions onto the ring) is related to the atomic ring properties. Only the contributions of the five ring atoms towards the atomic ring properties are considered. In other words, all properties were calculated using only the five atoms in the ring (excluding hydrogens and the atoms of the substituents). Confining the calculations to these five atoms enables one to explore the extent of distortion to a ring atom's basin upon change of its environment, as explained in Sect. 3.1.

Figure 10 shows the net charge for the ring atoms of a new selection of 16 heterocyclic rings each with eight different substituents, which were substituted onto the rings in the same manner as outlined in beginning of Sect. 4.2. The eight substituents were OH, OMe, NO_2, NH_2, NHMe, CN, F and COOH. The rings were selected as four groups of four, totalling 16. Within each group of four, the rings only differed in the element of one of the heteroatoms. For example, pyrrole (N), furan (O), thiophene (S) and phosphole (P) are the first group of four (panel a). The second group of four differs from the first group by an additional nitrogen at the second position of the ring (panel b). For the third group the additional hydrogen is now at the third position of the ring (panel c). The fourth group has two nitrogens in addition to the heteroatom at the first position of the ring. For this group, the two nitrogens are at the second and fifth positions in the ring.

Each profile line in any panel of Fig. 10 represents the average net charge of rings for the substituent of interest. For example, the average net charge for the nitro-substituted pyrrole rings is −0.2 a.u. (panel a). The average is over the three possible substitution sites. The ninth profile line in each panel refers to the non-substituted rings.

For each substituent the net charge follows the same profile as the heteroatom at the first position of the ring is changed, for all four panels. This shows that a change of the elements within the ring has the same systematic effect on the net charge, such that replacing a nitrogen with an oxygen increases the net charge by approximately 0.4 a.u. irrespective of the ring. Moreover, this systematic nature is also found for substituents because each substituent again has the same consistent effect on the net charge, regardless of the ring characteristics. This regularity expresses itself through the nine parallel profiles found in each panel, one profile for each substituent. For example, substituting an F with NO_2 decreases the net charge by approximately 0.2 a.u. for all rings.

In summary, there are two "characteristics" of the rings that are independently responsible for the ring's net charge. The first is changing the element of the ring, i.e. moving horizontally between rings, and the second is changing the substituent, i.e. moving vertically between substituents. This introduces the concept of *ring characteristic orthogonality,* where different characteristics of a ring affect the

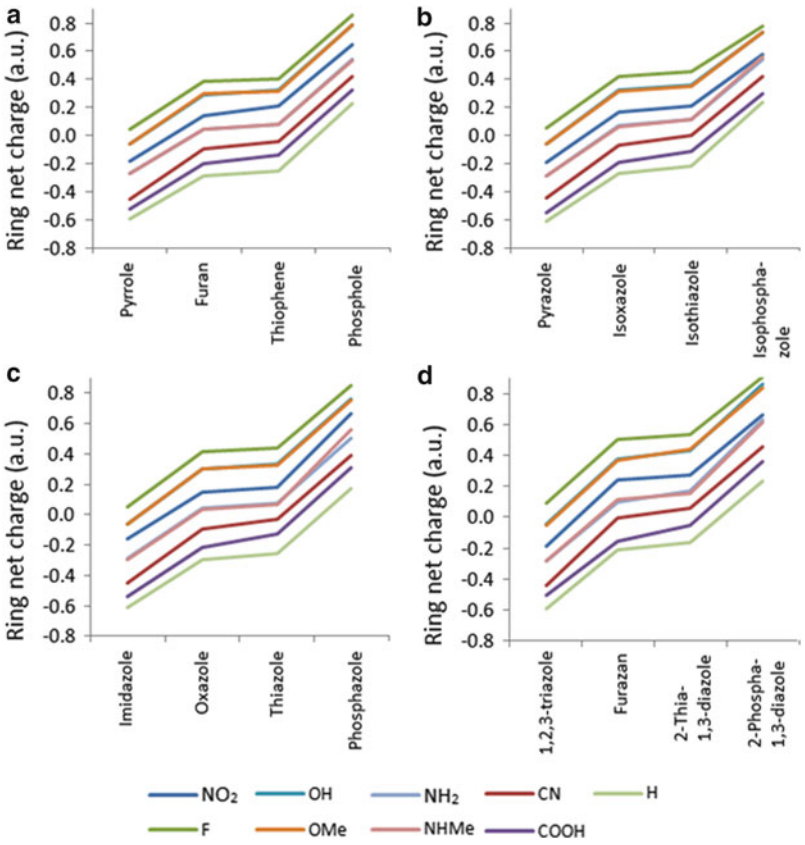


Fig. 10 Net charges for ring atoms of substituted five-membered heterocycles, where the heteroatom at the first position is varied for (**a**) single heteroatom ring scaffolds, (**b**) additional nitrogen in the second position of the rings in panel **a**, (**c**) additional nitrogen added to third position of the rings in panel **a**, (**d**) additional nitrogen's added to the second and fifth positions of rings in panel *a*

ring's properties orthogonally. In other words, each characteristic is independently responsible for the ring's properties. The concept of ring characteristic orthogonality is extended to four characteristics as the data set is expanded to include a larger variety of rings.

5.2 Characterising Ring Features

Figure 11 shows the net charges for 572 ring molecules separated into 9 substitutions (8 substitutions + H [non-substituted]). The net charges were obtained in the same manner as for Fig. 10, i.e. summed over the five ring atoms. The 572 consist of 20 different ring scaffolds, totalling 69 rings separated by possible substitution site.

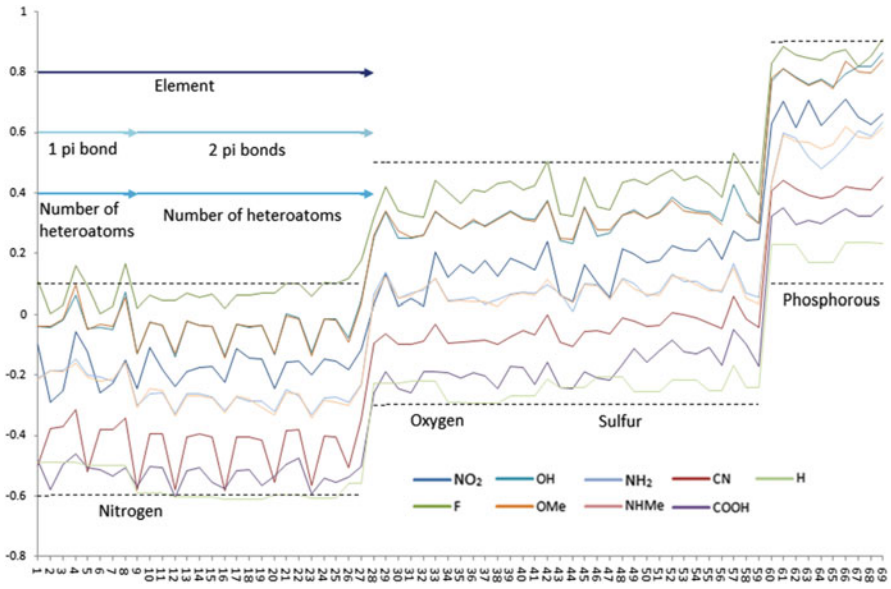


Fig. 11 Net charge of ring atoms for five-membered heterocyclics with respect to ring scaffolds and substitution types

For example, furan, from the set of 20 ring scaffolds, becomes furan-2 and furan-3 (see Fig. 6). Table 3 shows all 69 possible scaffolds. Each of the 69 rings is substituted with 8 substitution types, and the 20 non-substituted rings give $(69 \times 8) + 20 = 572$ unique ring structures. The rings were ordered along the x axis of Fig. 11 by a hierarchy of characteristics to maintain the orthogonality of the ring characteristics. The first separation in the hierarchy is into the elements of the rings, in a manner analogous to Fig. 10. The first rings contain nitrogen as their only heteroatom, followed by rings with an oxygen atom, then rings with sulfur and finally rings with phosphorous. The second separation of the hierarchy divides the rings of each element into single then double π-bond rings. The third hierarchy separation orders the rings within each π-bond group by increasing number of heteroatoms in the ring, from lowest to largest.

Each of the characteristics of the hierarchy is examined in turn to determine their effect on the net charge and if they remain within the scope of ring characteristic orthogonality. The first separation in the hierarchy is for the heteroatom elements. This separation is already known to be orthogonal with respect to substituents based on Fig. 10. Where Fig. 10 created a profile from four rings at a time, Fig. 11 recreates the same profile but for all 69 rings simultaneously. In Fig. 11 three bands appear, where each band corresponds to the element of the ring. For example, the range of net charges (of all ring atoms) for the nitrogen band is −0.6 to 0.2 a.u., that of the oxygen and sulfur band is −0.3 to 0.5 a.u. and that of the phosphorous band is 0.2 to 0.8 a.u. These ranges are the same as those found in Fig. 10. Therefore, if each band is considered as a single point on the profile, then the bands match up to the

Table 3 Data set of 69 unsaturated ring scaffolds separated into the different substitution sites possible for each ring scaffold

1	Pyrazoline-1	**24**	1,2,4-triazole-3	**47**	Isothiazoline-4
2	Pyrazoline-3	**25**	1,2,4-triazole-5	**48**	Isothiazoline-5
3	Pyrazoline-4	**26**	Tetrazole-1	**49**	Thiazole-2
4	Pyrazoline-5	**27**	Tetrazole-5	**50**	Thiazole-4
5	Imidazoline-1	**28**	Oxazoline-2	**51**	Thiazole-5
6	Imidazoline-2	**29**	Oxazoline-4	**52**	Isothiazole-3
7	Imidazoline-4	**30**	Oxazoline-5	**53**	Isothiazole-4
8	Imidazoline-5	**31**	Isoxazoline-3	**54**	Isothiazole-5
9	Pyrrole-1	**32**	Isoxazoline-4	**55**	Thiophene-2
10	Pyrrole-2	**33**	Isoxazoline-5	**56**	Thiophene-3
11	Pyrrole-3	**34**	Furan-2	**57**	1-thia-2,5-diazole-3
12	Pyrazole-1	**35**	Furan-3	**58**	1-thia-2,4-diazole-3
13	Pyrazole-3	**36**	Oxazole-2	**59**	1-thia-2,4-diazole-5
14	Pyrazole-4	**37**	Oxazole-4	**60**	Phosphole-1
15	Pyrazole-5	**38**	Oxazole-5	**61**	Phosphole-2
16	Imidazole-1	**39**	Isoxazole-3	**62**	Phosphole-3
17	Imidazole-2	**40**	Isoxazole-4	**63**	Phosphazole-2
18	Imidazole-4	**41**	Isoxazole-5	**64**	Phosphazole-4
19	Imidazole-5	**42**	Furazan-3	**65**	Phosphazole-5
20	1,2,3-Triazole-1	**43**	Thiazoline-2	**66**	Isophosphazole-3
21	1,2,3-Triazole-4	**44**	Thiazoline-4	**67**	Isophosphazole-4
22	1,2,3-Triazole-5	**45**	Thiazoline-5	**68**	Isophosphazole-5
23	1,2,4-Triazole-1	**46**	Isothiazoline-3	**69**	1-Phospho-2,5-diazole-3

profile of Fig. 10. The orthogonality of heteroatom element successfully extends from the 16 scaffolds of Fig. 10 to all 69 rings in Fig. 11. In other words, the first data set of four ring types (Fig. 10) is already representative for the full data set (Fig. 11).

The second separation in the hierarchy divides each element's band into rings with a single π bond and a double π bond. The single π bond rings occupy the first half of a band, while the double π-bond rings occupy the second half. There are no discernible differences between the first half (single π-bond) and the second half (double π bond) of this hierarchy separation. However, the first halves show erratic spikes within the net charge. These spikes are caused by substitutions onto nitrogens or sp^3 carbons. Note that the single π-bond rings also have sp^2 carbons and therefore some net charges are similar to the net charges of the double π bond rings, because substitutions occur onto the same type of carbon. This shows that the internal bonding structure of the ring indirectly affects the net charge profile as a single π-bond ring introduces sp^3 carbons that affect the net charge differently to sp^2 carbons. Even though the removal of one π-bond from the ring does not directly affect the ring's net charge, the effect of substituting onto an sp^3 carbon relative to an sp^2 carbon is still orthogonal to the other characteristics. The orthogonality

remains because the same spikes are found throughout all the element bands and are relatively parallel.

The third stage of the hierarchy is the ordering of rings within the single or double π-bond separations by increasing number of heteroatoms within the ring. For example, in the nitrogen double π-bond separation, the rings are ordered from a single nitrogen in the ring (pyrrole) to four nitrogens in the ring (tetrazole). The effect of the number of heteroatoms within the ring is best observed in double π-bond separations. This is because the erratic spikes found in the single π bond separation (discussed above) are greater than any changes in net charge due to the number of heteroatoms and therefore mask the effect. So, if the oxygen double π bond separation is examined, a small increase in net charge is found as the number of heteroatoms in the ring increases. Again, the effect of the number of heteroatoms within a ring on the net charge is orthogonal to changes due to other characteristics. However, the change in net charge due to the number of heteroatoms in a ring is occasionally masked by changes due to the π-bond structure of a ring.

It is also worth noting that for any ring pattern with an N-H, substitutions at this position tend to give very different ring net charges relative to substitutions at any other position in the ring. However, the change in net charge for substitutions onto this position is consistent.

Unexpected net charges occur when an NO_2 is substituted onto sp^3 carbons, producing net charges similar to or lower than NH_2 and NHMe where the NO_2's net charge is expected to be larger than both. This explains why the pink and blue profiles in Fig. 11 occasionally cross each other.

The three characteristics (hierarchy separations) discussed above and the parallel profiles of the substituents total four orthogonal characteristics that determine a ring's net charge. In other words, each characteristic independently affects the ring net charge. For example, replacing the F with an OH in a mono-substituted pyrrole, where either F or OH can be on any position of the ring, decreases the net charge by 0.1 a.u. Furthermore, if this new ring requires a net charge increase of 0.3 a.u., this is achieved through replacing the nitrogen with oxygen and forming furanol. This orthogonality of the four characteristics allows one to traverse the net charge landscape by manipulation of any of the ring's four characteristics.

5.3 Substituent Effects on π Electron Density

The order of the substituents (Fig. 11: highest profile line, F, to lowest profile line, H) expresses to what extent they alter the electronic charge of the ring. An interesting question is whether this order matches the order of functional groups of the corresponding Hammett constants. Starting with the most electron withdrawing, they are ordered as follows: $NO_2 > CN > F > H > OMe > NH_2$. On the other hand, the order of the substituents by decreasing ring net charge is $F > OMe > NO_2 > NH_2 > CN > H$. Clearly, the net charge variation of the substituents does not follow the order of their Hammett constants. A more informed

way of answering the question above realises that Hammett constants may correlate much better to the charge due to the rings' π electron density only [48]. In order to test this, the π molecular orbitals were determined for each molecule by visual inspection of the molecular orbitals using the program GAUSSVIEW. The program MORPHY was modified to recalculate the net charge of all ring atoms, but now from electrons occupying π orbitals only. This exclusion of σ molecular orbitals was tested for furan and imidazole. It turned out that the ring net charges for the substituents in both furan and imidazole now matched the order of the Hammett constants, when only the π electrons were considered. This confirmed that the substituents were having the correct effect on the ring electron density, which was not confirmed by initial calculations.

5.4 Ring Substitution Effects

The effect of substitutions on the net charge has already been explored with regard to the parallel profiles. However, the substituents themselves have so far been ignored. In other words, the significance of the positions of the profile lines with respect to each other demands attention. Examining the profiles with respect to each other can determine if a pattern between the substituents exists or which features of a substituent affect the change it imposes on the net charge. All substituents generally form parallel profiles, and, therefore, each substituent causes a systematic effect on the net charge specific to the substituent. However, strictly speaking, there are three other cases where the substituent breaks from this parallel arrangement. Fluorine remains parallel throughout but the changes in the net charge tend to be less pronounced, i.e. the F profile is flatter than that of other substituents. Conversely, rings substituted with a CN group tend to vary more than for other substituents, but this is only observed in the nitrogen band (first third of Fig. 11). The oxygen, sulfur and phosphorous ring patterns with a CN substituent behave in a parallel fashion similar to other substituents. The third case is NO_2, where its profile occasionally drops below the NH_2 and NHMe profiles, but this case has been discussed in the previous section.

Interestingly, the plots for OH and OMe almost follow each other exactly. The net charge of the ring does not seem to see the effect of a replacement of H by Me on the oxygen of the substituent. This case also occurs with NH_2 and NHMe. These observations call for further calculations attempting to generalise this phenomenon. The substitutions were chosen to explore three aspects of a substituent's effect on the ring's net charge:

1. *Replacement of a hydrogen atom with a methyl group* – This replacement determines the effect of extending the substitution fragment and has already been witnessed for the substitution of OH to OMe and NH_2 to NHMe.
2. *Inserting a methylene group between the ring and original substituent* – This addition determines if the methylene group acts as a "buffer" for the interaction

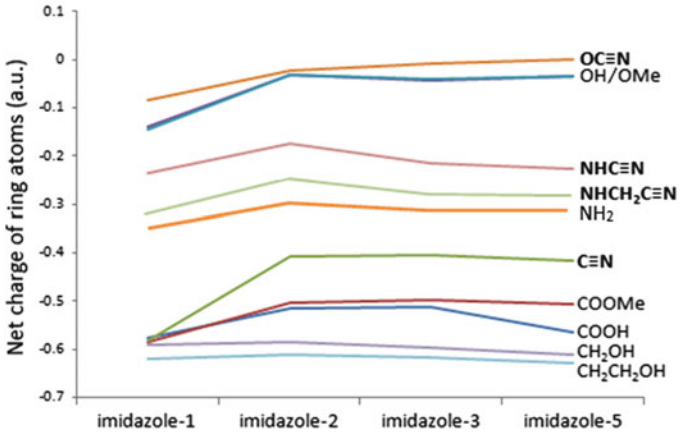


Fig. 12 Effects of substitution on ring atom properties. Substitutions were chosen to represent the effects of a functional group's proximity to the ring on the ring atom properties. All substituents with the cyano group are highlighted in *bold*

of the substituent with ring. In other words, the extent to which the effect of the original substituent permeates through the additional CH_2.

3. *Replacement of a hydrogen atom by CN* – This replacement determines the effect of a second functional group on the ring's net charge as this effect has to permeate through the original substituent, similarly to point 2. However, this replacement also determines if a second functional group will have a significant effect on the original substituents and alter how the ring's net charge responds to the original substituent.

The ten possible substituents are COOH, COOMe (applying point 1. to COOH), $C\equiv N$, OH, OMe (applying point 1. to OH), $OC\equiv N$ (applying point 3. to OH), $NHC\equiv N$ (applying point 3. to NH_2), $NHCH_2C\equiv N$ (applying points 2. and 3. to NH_2), CH_2OH (applying point 2. to OH) and CH_2CH_2OH (applying point 2. twice to OH).

Figure 12 shows the net charges of all the substituents onto all available substitution sites of imidazole. The substitutions are divided into three sets: the top set (between 0 and −0.15 a.u.), the middle set (between −0.17 and −0.35 a.u) and the lowest set (between −0.4 and −0.65 a.u.). The top set consists of OH, OMe and $OC\equiv N$, which are all the substituents connected to the ring through an oxygen. The middle set consists of $NHC\equiv N$, $NHCH_2C\equiv N$ and NH_2, which are all the substituents connected to the ring through a nitrogen. The bottom set consists of COOH, COOMe, $C\equiv N$, CH_2OH and CH_2CH_2OH, which are all the substituents connected to the ring through a carbon. This separation into sets based on connecting element rather than substituent becomes more remarkable when examining all four substituents that contain the cyano group: $C\equiv N$, $OC\equiv N$, $NHC\equiv N$ and $NHCH_2C\equiv N$. Naturally, one would expect the four CN substituents to form a set; however, they are separated into three sets by their connecting atom. Therefore,

it is the element of the atom connecting the substituent to the ring that dominates the effect of the substituent on the ring's net charge rather than the substituent itself.

The addition of the methyl onto a substituent (point 1.) has very little effect on the ring's net charge. The alterations of COOH to COOMe or OH to OMe cause little to no difference in the ring's net charge. Indeed, OH and OMe are almost identical. The effect of inserting a CH_2 between the ring and substituent (point 2.) is relatively small. The profile lines of CH_2OH and CH_2CH_2OH are almost identical. However, the OH substituent does differ greatly from its CH_2OH and CH_2CH_2OH counterparts. This large difference can be attributed to the connecting atom changing from oxygen to carbon. The OH belongs to the top set (oxygen connection to ring), but when the CH_2 is inserted between OH and the ring, the connecting atom changes to carbon. Therefore, CH_2OH and CH_2CH_2OH belong to the bottom set (carbon connection to ring). The large difference between the OH profile line and the CH_2OH and CH_2CH_2OH profile lines cannot be attributed to the effect of the CH_2 because it shows a small effect between CH_2OH and CH_2CH_2OH. While effects from the three separate alterations (points 1., 2. and 3.) are displayed in Fig. 12, it is the element of the atom connecting the substituent to the ring that dominates the substituent's effect on the ring's net charge.

6 Practical Use of Ring Properties: A Case Study

6.1 Imidazoleglycerol-Phosphate Dehydratase

Section 4 has shown how the QCT atom and RCP properties can reflect or capture a ring's characteristics. An important application of this characterisation is its use in ligand design. As all rings can be successfully characterised by the QCT properties, comparison of QCT properties between rings leads to powerful ligand design capabilities. The potential for ligand design through RCP and ring atom properties is explored with imidazoleglycerol-phosphate dehydratase (IGPD), an enzyme found only in plants. Figure 13 shows the structure of the strongest inhibitor of IGPD, which has the lowest inhibitor constant (K_i). The inhibitor constant is the concentration of the inhibitor required to displace 50 % of the agonist from the binding site. Therefore, if the K_i value of an inhibitor is small, then a small concentration is required to inhibit the target (by displacing the agonist), in which case the inhibitor is considered as stronger. IGPD was chosen because it is a well-studied system where it is known that the ring of the inhibitor is directly involved with binding to the protein [49]. Alterations to the ring dramatically change the K_i value of the molecule. Through crystal structures, it is known that the inhibitors that have the strongest binding have hydrogen bond acceptors at the second and fourth positions of the ring (Fig. 13).

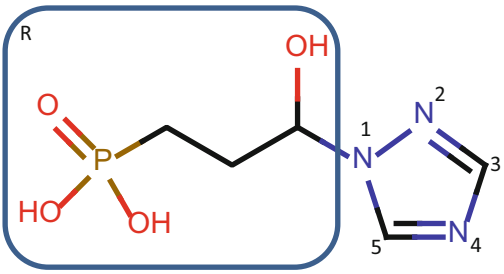


Fig. 13 Strongest inhibitor of IGPD. The ring positions of the ring of interest are shown and the rest of the molecule (R) is highlighted

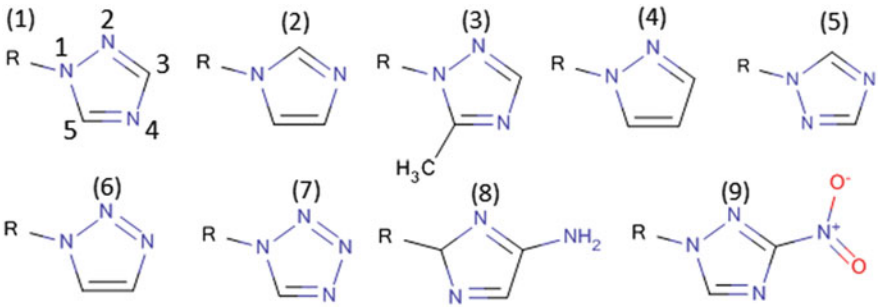


Fig. 14 Nine different rings for IGPD inhibitor molecule where R denotes the rest of the molecule as shown in Fig. 13. Ring positions are shown on molecule (1)

6.2 Method

The nine inhibitors in Fig. 14 were geometry optimised at the B3LYP/6-311 + G(2d, p) level. The optimised structures were verified by eye to ensure that in each molecule the ring retained its orientation relative to the rest of the molecule (R in Fig. 13). The ring's initial orientation is determined by the dihedral angle between ring atom 2, ring atom 1, the carbon (within R, i.e. outside the ring) adjacent to ring atom 1 and the oxygen of the alcohol. For a ring to retain its original orientation, this dihedral angle must lie in the interval $[-20°, +20°]$. If the ring rotated out of its original orientation during the geometry optimisation, then the structure was not considered and the calculation was restarted.

QCT properties were calculated with the program MORPHY01 [50, 51]. It should be noted that the multipole moments such as the dipole moment (and higher rank moments) are direction dependent. Therefore, it is important to express the moments of all the rings in a global axis system, common to all molecules, so that the moments are comparable between molecules. Such a comparison can be made conveniently through a single distance score. However, because the components contributing to this distance are physical properties with different units, all properties were normalised.

The global axis system in which the multipole moments are expressed in this work is fully defined by three atoms. First, the origin is defined as the carbon atom

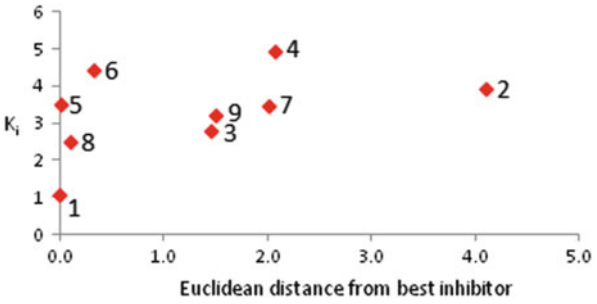


Fig. 15 Quantum Isostere Database (QID) distances from best IGPD inhibitor calculated with the electron density, its Laplacian and RCP ellipticity

Table 4 Euclidean distances from strongest IGPD inhibitor and their corresponding K_i values

Molecule	RCP distance	RCP and atom distance	K_i
1	0	0	1.08
2	4.10	8.91	3.95
3	1.46	2.73	2.78
4	2.07	4.89	4.93
5	0.02	7.26	3.5
6	0.32	5.16	4.42
7	2.01	4.46	3.46
8	0.11	2.65	2.48
9	1.50	3.22	3.21

adjacent to the ring. Next, the *x* axis is defined by the bond between the carbon atom at the origin and the ring atom (N1) it is bonded to. Finally, the *xy* plane is defined by these two atoms and a third atom, which is the oxygen of the alcohol group bonded to the origin carbon. From the *xy* plane the *z* axis protrudes perpendicularly forming a right-handed axis system. This global axis system is chosen because the ring plane now approximately lies in the *xy* plane of the global axis system, by virtue of the geometry optimisation characteristics mentioned just above.

6.3 Results

First, the Euclidean distance from the best inhibitor (1) was calculated for rings 2–9 using only three RCP properties, which are the electron density, its Laplacian and the RCP ellipticity. Figure 15 plots the calculated distances against the K_i values for all nine molecules. All values are also given in Table 4.

The best inhibitor (ring 1) has a K_i of 1.08 and, naturally, a Euclidean distance of 0. According to the Euclidean distance, the ring closest to ring 1 is ring 5, the K_i of which is 3.5 but the distance only 0.02. The distance of ring 5 is too small for its K_i value because it does not have hydrogen bond acceptors at positions 2 and 4 of the ring. However, the distance is small because it is the same ring as ring 1, except for a different orientation with respect to the rest of the molecule. The difference in

orientation is not seen by the RCP properties because they are expressed in a local axis system that "travels" with the ring in 3D space, as discussed in Sect. 2.1. Therefore, the RCP properties of rings 1 and 5 are almost identical, but not quite because the distance is not exactly 0. Therefore, the influence of the rest of the molecule on ring 5 is different to that on ring 1, due to the different orientations of the rings with respect to the rest of the molecule. Ring 6 is a similar case to ring 5, where the ring structure is similar to that of ring 1 but its orientation differs. Hence, ring 6 has a distance (from ring 1) of only 0.32 because of the similar ring structure, but the K_i value of ring 6 differs substantially from that of ring 1, by 3.34. The relatively small distance of 0.32 places ring 6 much closer to the best inhibitor (ring 1) than it should, according to its K_i value, because the ring lacks a hydrogen bond acceptor at its ring position 4 (see Fig. 14). However, the distance is small because it is an isomer of ring 1 (ring 1 is a 1,2,4-triazole while ring 6 is a 1,2,3-triazole), and, therefore, the RCP properties are similar. Excluding the cases (rings 5 and 6) where it is the ring's orientation that affects the K_i value (by not having the hydrogen bond acceptors at the correct position), there is a general increase in distance as the K_i increases.

6.4 A Combined RCP and Ring Atom Property Approach

QCT ring atom properties were incorporated into the distance calculation to determine if the ring atom properties can improve the relationship between K_i and distance. The following ring atom properties were calculated: the ring monopole, dipole, quadrupole moments and their respective magnitudes; the kinetic energy K; and the volume (0.001 a.u. cut-off). The ring properties were calculated by summing the individual atom properties for only the atoms within the ring. The moments for the atoms were taken with respect to the defined global axis system.

Figure 16 shows the distances for the ring atom and RCP properties combined (green triangles) compared to those of the RCP properties only (red square). The incorporation of the ring atom properties into the distance calculation increases the distance of ring 5 from 0.02 to 7.26. Now the new distance for ring 5 is not too small for its K_i value. Although rings 1 and 5 are very similar and their ring atom properties as well, expressing the multipole moments in the global axis system successfully differentiates between the two rings. Ring 5 can be seen as a "flipped" ring 1. Therefore, the moments flipped with the ring and their direction component is different. If the moments were expressed in a local axis system instead of the global axis system, this flip would not be observed and the moments would be almost identical for both rings. Although the ring pattern is identical, the flip is seen in the change of the moments with respect to the global axis system.

A similar case occurs with ring 6, where the distance increases from 0.32 to 5.16 upon incorporation of ring atom properties into the distance calculation. It is remarkable that the summed ring atom properties are able to create more suitable distance scores because the summed properties still have no information as to the

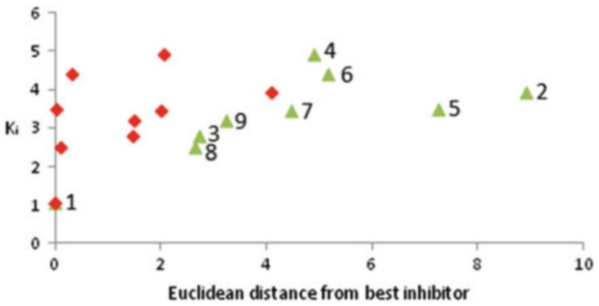


Fig. 16 Euclidean distances from best IGPD inhibitor calculated with electron density, its Laplacian and RCP ellipticity (*red squares*) and distances calculated with QCT ring atom and RCP properties (*green triangles*) for nine IGPD inhibitors

ideal positions of the hydrogen bond acceptors. Individual atom properties are not expressed, and it is through the directional dependence of the moments that it is possible to distinguish between rings 1 and 5.

Whether using only RCP properties or RCP and ring atom properties, ring 2 has the largest distance. Ring 2 differs most to ring 1 in ring pattern, with the exception of ring 8, but the amine substituent affects the properties of the ring as seen in Sect. 3.1, and so has the largest distance and the distance may be over-penalising. Rings 2, 4 and 8 are the only ring patterns with fewer than 3 nitrogens in the ring, and these rings are the three highest scoring distances using only RCP properties. Ring 4 has a better score than ring 2 because the two nitrogens are adjacent, and therefore the ring pattern matches more closely that of ring 1. Ring 8 has the influence of the amine on the ring, so the RCP properties decide that ring 2 is has the largest distance of any ring. When the ring atom properties are included in the distance calculation (rather than just the RCP properties), then ring 2 retains it extreme position, and its distance increases to 8.91.

In summary, distances calculated with only the RCP properties correlate quite reasonably with K_i values where rings with similar RCP properties exhibit similar activities. The few outliers that occur are expected and can be explained. When ring atom properties are added to the distance calculation, then the outlier better fits the correlation.

7 Conclusions

From a series of PLS models, it turned out that ring atom properties can be obtained from RCP properties but not the reverse. However, the ellipticity proved to be the most difficult to predict. This is because the ellipticity is the most sensitive to the changes in the ring atoms and therefore more difficult to predict accurately.

Four features were found to determine a ring's net charge. These features are (i) the element of the heteroatom within the ring, (ii) the number of heteroatoms within the ring, (iii) the substituent itself and (iv) the substituent site. Each feature causes a signature change to the ring's net charge, independent to any changes caused by other ring features. This phenomenon was called *ring characteristic*

orthogonality, where different characteristics of a ring affect the ring's properties orthogonally. It was found that the substituent's atom connecting this substituent is more important in determining the substituent's effect on the ring's net charge than the remaining atoms in the substituent.

Finally, the RCP properties and ring atom properties were tested against the IGPD inhibitor molecule. The RCP properties produced a mediocre correlation and struggled with certain cases where the ring's orientation altered its activity. When ring atom properties were included, then these outliers fell back into the correlation and the general trend was improved.

References

1. Lameijer EW, Kok JN, Back T, Ijzerman AP (2006) J Chem Inf Model 46:553–562
2. Ertl P (2003) J Chem Inf Comput Sci 43:374–380
3. Holliday JD, Jelfs SP, Willett P, Gedeck P (2003) J Chem Inf Comput Sci 43:406–411
4. Lima LMA, Barreiro EJ (2005) Curr Med Chem 12:23–49
5. Patani GA, LaVoie EJ (1996) Chem Rev 96:3147–3176
6. Wermuth C-G (1996) Molecular variations based on bioisosteric replacements. In: Wermuth C-G (ed) The practice of medicinal chemistry. Academic, London, pp 202–237
7. Olesen PH (2001) Curr Opin Drug Discov Dev 4:471–478
8. Devereux M, Popelier PLA, McLay IM (2009) J Chem Inf Model 49:1497–1513
9. Graham JE, Ripley DC, Smith JT, Smith VHJ, Weaver DF (1995) J Mol Struct THEOCHEM 343:105–109
10. Gisi U, Sierotzki H, Cook A, McCaffery A (2002) Pest Manag Sci 58:859–867
11. Katritzky A, Rees C (1984) Comprehensive heterocyclic chemistry, vol 5. Pergamon Press, Oxford
12. Lehringer A, Nelson DL, Cox MM (2008) Lehninger principles of biochemistry. W. H. Freeman, New York
13. Liu H, Du DM (2009) Adv Synth Catal 351:489–519
14. Talley JJ, Brown DL, Carter JS, Graneto MJ, Koboldt CM, Masferrer JL, Perkins WE, Rogers RS, Shaffer AF, Zhang YY (2000) J Med Chem 43:775–777
15. Yang CY, Meng CL, Liao CL, Wong PYK (2003) Prost Other Lipid Mediat 72:115–130
16. Steinbach G, Lynch PM, Phillips RK, Wallace MH, Hawk E, Gordon GB, Wakabayashi N, Saunders B, Shen Y, Fujimura T (2000) New Eng J Med 342:1946–1952
17. Krygowski TM, Ejsmont K, Stepien BT, Cyranski MK, Poater J, Sola M (2004) J Org Chem 69:6634–6640
18. Poater J, Sola M, Viglione RG, Zanasi R (2004) J Org Chem 69:7537–7542
19. Poater J, Fradera X, Duran M, Sola M (2003) Chem Eur J 9:400–406
20. Sjoberg P, Murray JS, Brinck T, Politzer P (1990) Can J Chem 68:1440–1443
21. Murray JS, Abu-Awwad F, Politzer P (2000) J Mol Struct THEOCHEM 501:241–250
22. Poater J, Duran M, Sola M (2004) Int J Quantum Chem 98:361–366
23. Matta CF, Hernandez-Trujillo J (2003) J Phys Chem A 107:7496–7504
24. Popelier PLA, Aicken FM (2003) ChemPhysChem 4:824–829
25. Popelier PLA, Smith PJ (2006) Eur J Med Chem 41:862–873
26. Liem SY, Popelier PLA, Leslie M (2004) Int J Quantum Chem 99:685–694
27. Singh NK, Popelier PLA, O'Malley PJ (2006) Chem Phys Lett 426:219–221
28. Poater J, Fradera X, Sola M, Duran M, Simon S (2003) Chem Phys Lett 369:248–255
29. Koch U, Popelier P (1995) J Phys Chem 99:9747–9754
30. Popelier PLA (1999) J Phys Chem A 103:2883–2890

31. Bader RFW (1990) Atoms in molecules. A quantum theory. Oxford University Press, Oxford
32. Biegler-Koenig FW, Bader RFW, Tang TH (1982) J Comput Chem 3:317–328
33. Popelier PLA (1998) Morphy98 – a program written by Popelier PLA with a contribution from Bone RGA, Manchester
34. Biegler-Koenig FW, Schoenbohm J (2002) J Comp Chem 23:1489–1494
35. Murray JS, Politzer P (1998) J Mol Struct THEOCHEM 425:107–114
36. Suresh CH, Koga N, Gadre SR (2000) Organometallics 19:3008–3015
37. Kroemer RT, Hecht P, Liedl KR (1996) J Comput Chem 17:1296–1308
38. Lamarche O (2003) Theoretical prediction and application of hydrogen bond and polarity/polarisability descriptors. PhD, Cardiff University
39. Platts JA (2000) Phys Chem Chem Phys 2:973–980
40. Platts JA (2000) Phys Chem Chem Phys 2:3115–3120
41. Murray JS, Politzer P (1991) J Org Chem 56:6715–6717
42. Kosov DS, Popelier PLA (2000) J Phys Chem A 104:7339–7345
43. Kosov DS, Popelier PLA (2000) J Chem Phys 113:3969–3974
44. Hofinger S, Wendland M (2002) Int J Quantum Chem 86:199–217
45. Wold S, Sjostrom M, Eriksson L (1998) Partial least squares projections to latent structures (PLS) in chemistry. In: Schleyer P (ed) Encyclopedia of computational chemistry, vol 3. Wiley, Chichester, pp 2006–2021
46. Wold S, Sjostrom M, Eriksson L (2001) Chemom Intell Lab Sys 58:109–130
47. GAUSSIAN03, Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JAJ, Vreven JT, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2003) Gaussian, Inc., Pittsburgh
48. Slee T, Larouche A, Bader RFW (1988) J Phys Chem 92:6219–6227
49. Mori I, Fonne-Pfister R, Matsunaga S, Tada S, Kimura Y, Iwasaki G, Mano J, Hatano M, Nakano T, Koizumi S, Scheidegger A, Hayakawa K, Ohta D (1995) Plant Physiol 107:719–723
50. Popelier PLA (1994) Chem Phys Lett 228:160–164
51. Popelier PLA (1996) Comput Phys Commun 93:212–240

Valence Bond Theory in Heterocyclic Chemistry

Zahid Rashid, Ria Broer, Joop H. van Lenthe, and Remco W.A. Havenith

Contents

Z. Rashid • J.H. van Lenthe
Theoretical Chemistry Group, Department of Chemistry, Debye Institute For Nanomaterials Science, Utrecht University, Princetonplein 1, 3584 CC, Utrecht, The Netherlands

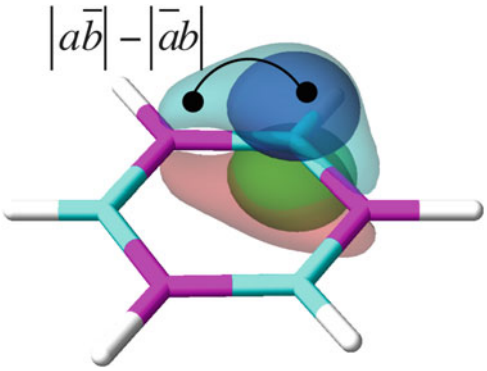
R. Broer
Theoretical Chemistry, Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

R.W.A. Havenith (✉)
Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

Ghent Quantum Chemistry Group, Department of Inorganic and Physical Chemistry, Ghent University, Krijgslaan 281 (S3), B-9000 Gent Belgium
e-mail: r.w.a.havenith@rug.nl

F. De Proft, P. Geerlings (eds.), *Structure, Bonding and Reactivity of Heterocyclic Compounds*, Topics in Heterocyclic Chemistry 38, DOI 10.1007/978-3-642-45149-2_4

Abstract This chapter deals with the application of valence bond theory in heterocyclic chemistry. A short introduction to the different valence bond methods is given, followed by illustrative results obtained by valence bond calculations. The illustrations show the applicability of the valence bond theory to obtain detailed information on the electronic structure of molecules in terms of chemical concepts.



Keywords Heterocyclic chemistry • Valence bond theory • Aromaticity

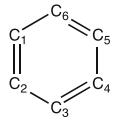
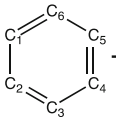
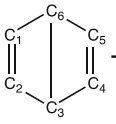
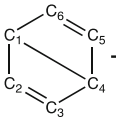
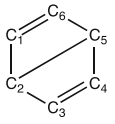
1 Introduction to Valence Bond Methods

Computational chemistry is a very powerful tool for the study of the electronic and geometric structure of molecules as it provides independent knowledge on structure, reactivity and properties. A main advantage of applying theory is that it does not only yield an accurate description of the electronic structure but also allows an interpretation in terms of chemical concepts, like chemical bonds. Currently, molecular orbital methods [1–5], such as Hartree-Fock theory, density functional theory and correlated wavefunction-based methods, such as the coupled cluster approach, are standard methods for studying the electronic structure of molecules. In these molecular orbital methods, the molecular orbitals (MOs) are expanded as linear combinations of atomic orbitals, i.e. of atom-centred basis functions. These MOs are orthogonal to each other, and each MO spans the entire molecule rather than being localised on particular atoms or in certain bonding regions. These delocal MOs do not give the intuitive chemist's picture of a bond. They may be

transformed to localised orbitals [6–9] without affecting the total wavefunction; however, there is no unique way to do that.

An alternative approach to find the solutions of the electronic Schrödinger equation for molecular systems is the use of valence bond (VB) theory (for reviews, see Refs. [10, 11]), which is closely related to the chemist's idea of molecules as being composed of atoms and held together by chemical bonds. Actually, the idea that atoms combine by sharing their valence electrons to form molecules predates modern quantum mechanics [12]. Heitler and London [13], using the principles of quantum mechanics, showed for the first time how the sharing of electrons holds the atoms together in a covalent molecule like H_2, thus providing the quantum mechanical basis for this viewpoint. Pauling and Slater introduced the idea of maximum overlap between the orbitals and the "changed quantisation" (hybridisation), which became very popular among the organic chemists. Further efforts by Heitler and Rumer [14], Slater [15], Pauling [16, 17] and many others [18] developed the Heitler-London methods into a general theory of electronic structure which is now known as the valence bond theory.

In the valence bond (VB) theory, the starting point is that atoms form bonds to make a molecule. Singly occupied atomic-like non-orthogonal orbitals are used to form a bond, by spin coupling the orbitals. The wavefunction for one particular bonding pattern is written as an antisymmetrised product of the spin-coupled bonds. This wavefunction consists of multiple determinants and is called a VB structure. In general, more than one structure can be formed, and the final VB wavefunction is written as a linear combination of such structures. These structures bear close resemblance to Lewis structures, and the wavefunction can thus be interpreted in terms of structures and resonances between structures. Delocalised bonding can be distinguished from localised bonding in this picture through the importance of multiple structures in the description of a molecule. For example, for the description of the π-system of benzene in VB terms, there are six singly occupied atomic p-orbitals, one on each carbon atom, which lead to five different singlet spin functions, thus to five different structures. These five structures correspond to the two Kekulé structures and the three Dewar structures. A VB wavefunction for benzene can thus be written as a linear combination of these five structures:

$\Psi =$ c_1  + c_2  + c_3  + c_4  + c_5 

In the VB calculation, the structure coefficients c_i are variationally optimised. Once the structure coefficients are known, a weight can be assigned to each of the structures. In the case of benzene, the dominant structures are *1* and *2*, so one expects $c_1^2 = c_2^2 > c_3^2 = c_4^2 = c_5^2$. However, as the atom-centred orbitals that are used to build the structures are mutually non-orthogonal, the structures are also not orthogonal. Thus, their sum $\sum_i c_i^2$ is not equal to 1, and this hampers the use of c_i^2 as a weight of structure *i* in the VB wavefunction. Weights for each structure that do

sum up to 1 can be assigned, according to different definitions. One definition is according to Chirgwin and Coulson [19]:

$$W_j = \sum_i c_i c_j S_{ij}\,; \quad \sum_j W_j = 1 \tag{1}$$

A disadvantage of these weights is that individual weights can be negative or larger than 1, which makes the interpretation rather cumbersome. Gallup and Norbeck (GN) [20] defined the weights of each structure according to

$$W_j = N\left|c_j\right|^2 / \left(\mathbf{S}^{-1}\right)_{jj}\,; N^{-1} = \sum_i \left|c_i\right|^2 / \left(\mathbf{S}^{-1}\right)_{ii} \tag{2}$$

An advantage of this definition for the weights of structures is that these weights do not only sum up to 1 but also are always positive.

A major application area for the VB theory is the study of systems with delocal electrons, for example, (poly)cyclic aromatic hydrocarbons. In such cases, it is not one or two Kekulé structures that dominate the wavefunction, but a qualitatively correct picture can only be obtained by using several VB structures. Then the weight of the different structures and the energy that is gained by resonance between the structures with respect to one (reference) structure, the resonance energy [17, 21], are of great importance. This resonance energy can be used as an indicator for aromatic character.

In this chapter, first, a short overview of the VB theory is given, followed by illustrative VB calculations on heterocyclic systems.

2 Valence Bond Methods

2.1 Classical Valence Bond Theory

The first application of the VB theory was published by Heitler and London in a paper treating the chemical bond in the H_2 molecule. In the approach that is nowadays called the Heitler-London method [13], a bond is described as a singlet-coupled pair of singly occupied atomic orbitals. For example, for the H_2 molecule, the $1s_a$ orbital of hydrogen atom A is singlet coupled to the $1s_b$ orbital of hydrogen atom B. The antisymmetrised Heitler-London wavefunction for the H_2 molecule can be represented by a linear combination of two Slater determinants:

$$\Phi_{HL} = N\left\{\left|1s_a(1)\overline{1s_b(2)}\right| - \left|\overline{1s_a(1)}1s_b(2)\right|\right\} \tag{3}$$

or, alternatively, as a product of a spatial and a spin wavefunction

$$\Phi_{HL} = N\{1s_a(1)1s_b(2) + 1s_b(1)1s_a(2)\}\{\alpha(1)\beta(2) - \beta(1)\alpha(2)\} \quad (4)$$

where N is the normalisation constant. The above wavefunction consisting of only one covalent structure gives a qualitatively correct picture of the bonding at equilibrium internuclear distance. For obtaining quantitative results, it must be improved. This can be done by adding configurations, which correspond to ionic structures:

$$\Phi_{ionic} = N\{1s_a(1)1s_a(2) + 1s_b(1)1s_b(2)\}\{\alpha(1)\beta(2) - \beta(1)\alpha(2)\} \quad (5)$$

The total VB wavefunction is then taken as a linear combination of the covalent Heitler-London structure and the ionic structures:

$$\Psi_{VB} = c_1\Phi_{HL} + c_2\Phi_{ionic} \quad (6)$$

and c_2/c_1 is variationally adjusted in a non-orthogonal CI (VBCI) (see also [22, 23]).

2.2 The Coulson-Fisher Approach

In the classical form of the VB method, the atomic orbitals are not allowed to distort freely as the atoms approach each other in forming a molecule. To compensate for this restriction, a large number of ionic structures are required to obtain an accurate VBCI wavefunction. Once the number of ionic structures becomes large, the simple predictive and interpretative power of the VB wavefunctions is obscured. Coulson and Fischer [24] suggested the use of optimal (slightly deformed atomic) orbitals. For the H_2 molecule, a Coulson-Fischer wavefunction can be represented as

$$\Psi_{CF} = N\{\phi_1(1)\phi_2(2) + \phi_2(1)\phi_1(2)\}\{\alpha(1)\beta(2) - \beta(1)\alpha(2)\} \quad (7)$$

where ϕ_1 and ϕ_2 are now deformed "atomic" orbitals:

$$\phi_1 = 1s_a + \lambda 1s_b \text{ and } \phi_2 = 1s_b + \lambda 1s_a \quad (8)$$

where λ is a variationally adjustable parameter. The resulting orbitals are predominantly atomic in character; however, they are slightly delocal. Since this deformation of the orbitals automatically includes the effect of the ionic structures, an obvious advantage of the Coulson-Fischer approach is that the number of structures in the wavefunction can be kept to a minimum. The total number of covalent structures is given by

$$f_S^N = \begin{pmatrix} N \\ \frac{1}{2}N + S \end{pmatrix} - \begin{pmatrix} N \\ \frac{1}{2}N + S + 1 \end{pmatrix} \tag{9}$$

where N is the number of spin-coupled electrons and S is the total spin of the system.

2.3 Generalised Valence Bond Method

The first computational VB method based on the Coulson-Fisher approach is the generalised valence bond (GVB) method [25–27]. In the GVB method, two non-orthogonal orbitals are used to describe a pair of electrons in a bond. Each pair of electrons in a bond is coupled to a singlet. The singlet-coupled pairs of orbitals are then combined to give the total wavefunction of a system with an overall singlet spin state. This is known as the GVB perfect pairing (GVB-PP) approach. Such two-electron two-orbital pairs are called geminal pairs. These orbitals are optimised variationally. The computational cost is reduced by forcing the orbitals from different geminal pairs to be orthogonal, i.e. by imposing strong orthogonality (SO) between different pairs. The SO-GVB-PP approach is a restricted form of VB where only one single spin-coupling pattern is allowed. In another form called unrestricted SO-GVB method [28], multiple spin couplings are allowed for a given number of singly occupied orbitals. The restriction on orbitals of different pairs to remain orthogonal, while computationally advantageous, leads in some cases to artefacts [29].

2.4 The Spin-Coupled Valence Bond Method

The spin-coupled valence bond (SCVB) method [30–33] uses a more generalised form of the Coulson-Fisher wavefunction for molecules. In the SCVB approach, the N electrons which are involved in the bonding are described by N singly occupied non-orthogonal orbitals. These orbitals are then spin coupled in all possible ways to give the overall spin S of the system. The general form of an SCVB wavefunction is

$$\Psi_{SCVB} = N \sum_{k=1}^{f_S^N} c_{Sk} \hat{A} \left[\phi_1 \phi_2 \phi_3 \cdots \phi_N \Theta_{S,M;k}^N \right] \tag{10}$$

where N is the normalisation factor, $\hat{A}$ is an antisymmetriser, $\Theta_{S,M;k}^N$ describes a particular spin-coupling pattern between the orbitals and c_{Sk} is the spin-coupling

coefficient. The different spin-coupling modes are generally referred to as the VB structures. The total number of these structures, f_S^N, is given by Eq. 9. In SCVB, both the orbitals and the spin-coupling coefficients are optimised. An inexpensive alternative to SCVB is the coupled cluster valence bond method [34, 35].

2.5 *The Valence Bond Self-Consistent Field Method*

The valence bond self-consistent field (VBSCF) method [36–38] is the most general form of the modern VB approaches. Mathematically, it can be viewed as the non-orthogonal equivalent of the MCSCF methods. In VBSCF approach, the VB wavefunctions can be constructed using any number of VB structures and non-orthogonality among all orbitals is allowed. Both the orbitals and the VB structure coefficients are variationally optimised. The resulting wavefunctions are compact and can be easily interpreted in terms of chemical concepts. The orbitals used to build the VB structures may be fully optimised as in the SCVB approach, or they may be restricted to a subspace of the full orbital space, e.g. centred on only one particular atom. The expressions for the analytical molecular energy gradients [39, 40] and second-order response properties [41] have also been developed for the VBSCF wavefunctions. Recently, a new approach, called the atoms in valence bond (AIVB) method [42] has been developed within the framework of the VBSCF method. In this approach, instead of the traditional VB structures, the wavefunction is constructed as a linear combination of all possible atomic states of the different atoms in a molecule. The advantage of this approach is that the VB wavefunction contains information about the atomic states of the atoms that form the molecule, without any preconceptions.

As has been noted, the VB orbitals may be restricted to a subspace or may be fully optimised (Fig. 1). Orbitals can thus be kept strictly centred on one atom (strictly atomic or "local" model), all atoms ("delocal" model, no orbital restrictions) or on a group of atoms ("partial delocal" model) [43]. Note that, for example, for H_2 in a minimal atomic basis set, the HL wavefunction including ionic contributions (Eq. 6) and the CF wavefunction (Eq. 7) are identical. The ionic structures in the strictly atomic wavefunctions are the so-called Brillouin structures (structures that are singly excited states with respect to the ground state wavefunction [37]) of the delocal wavefunction.

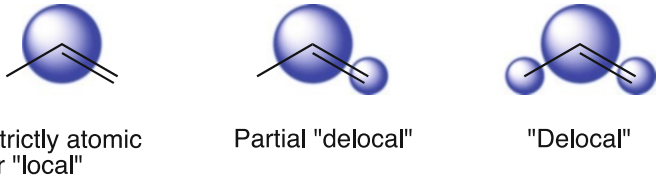


Fig. 1 Schematic representation of the orbital models used in VBSCF

The use of these orbital restrictions to keep orbitals strictly atomic has conceptual advantages: with strictly atomic orbitals, the interpretation of the structures is clear. A major disadvantage of these restrictions is that the wavefunction usually has not enough degrees of freedom to fully adopt to the molecular situation, and the danger exists that the outcome of the calculation is determined by the restrictions put into the input of the computation (except for the case that all possible (or many) structures are taken into account, but this is usually too expensive and due to the large number of structures also not interpretable). The delocal model has more freedom to adapt to a molecular situation; the orbitals are allowed to delocalise over the entire molecule. As a result, it might even happen that the final orbitals after orbital optimisation have lost all resemblance with the orbitals of the original structure, so that the meaning of the VB structures has also changed. Forgetting to include an important structure is lethal in the local model but less lethal in the delocal model, although orbital optimisation cannot recover everything! A second disadvantage of applying orbital restrictions is that the result becomes more basis set dependent: in the case of very large basis sets, the local or partial delocal orbital restrictions are not tenable, and the VB local answers converge with increasing basis sets to those of the delocal model.

An improvement of the standard VB model can be obtained by using the so-called breathing orbitals [44–47], i.e. each structure has its own optimised orbitals. For example, in ionic structures, the fragments bearing the negative charge probably have more diffuse orbitals than they have in the neutral structures. This makes the calculations more difficult, but higher accuracy is obtained.

2.6 Block Localised Wavefunction Approach

A special type of electronic structure theory which shows similarities with the VB theory and that has become rather popular recently is the block-localised wavefunction (BLW) approach [48–50]. In the VB methods, bonds are formed by spin coupling two singly occupied orbitals. In the BLW approach, a bond is formed by doubly occupying an orbital that is localised between two atoms (Fig. 2).

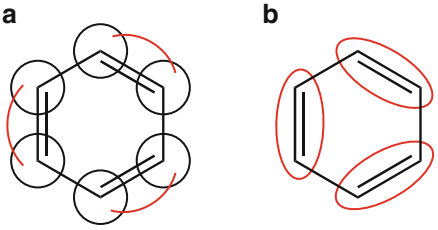


Fig. 2 Symbolic representation of 1,3,5-cyclohexatriene in (**a**) the VB method, where the three π-bonds are made by spin coupling six singly occupied p-orbitals, and (**b**) the BLW method, where the three bonds are described by three localised, doubly occupied orbitals

The similarity with the VB theory is that these doubly occupied orbitals are not required to be mutually orthogonal. They are only restricted to remain localised on a particular set of atoms. The wavefunction in this case is one single determinant, built from localised, non-orthogonal, doubly occupied orbitals. The BLW approach relates to the Hartree-Fock method as the spin-coupled and VBSCF methods relate to the CASSCF method. The BLW approach measures a delocalisation energy rather than a resonance energy, which are subtly different [51].

2.7 Orthogonal Valence Bond Methods

The close resemblance between the VB methods that use orthogonal optimised orbitals and the CASSCF method [52] makes the decomposition of the CASSCF wavefunction in terms of the orthogonal VB (OVB) structures rather straightforward [53, 54]. Within the OVB framework, the structures are built from orthogonal "atomic" orbitals, which can be deduced from a preceding CASSCF calculation [48, 49] or, in an alternative approach, be obtained by Löwdin orthogonalisation [55, 56] of non-orthogonal atomic orbitals. For H_2, it was found that in the orthogonal VB approach, the singlet neutral structure behaves similarly to the triplet state, which means that it is dissociative. Bond formation is due to mixing with the bound ionic state [53]. The advantages of this method are its flexibility and generality and the fact that one can think in terms of non-overlapping, exclusive bonding situations. A conceptual advantage is that the neutral singlet configuration is really neutral and close to the intrinsically pure neutral triplet state.

3 Illustrative Results

3.1 Inorganic Benzenes

The inorganic analogue of benzene, borazine ($B_3N_3H_6$), has been the subject of different VB studies [57–60]. In these studies, the central question is the aromatic character of borazine. Aromaticity is not a well-defined concept, and various mutually differing criteria for aromaticity exist; one of these criteria is the resonance energy [17, 21], which is the extra energetic stabilisation with respect to a hypothetical system [61]. Using VB, not only the localised charge distribution can be described, but also models may be constructed to describe the hypothetical localised charge distributions, which serve as references for further analysis. The resonance energy can be defined as the energy difference between the localised and the delocal description. This resonance energy is dependent on the model used to describe the localised hypothetical charge distribution.

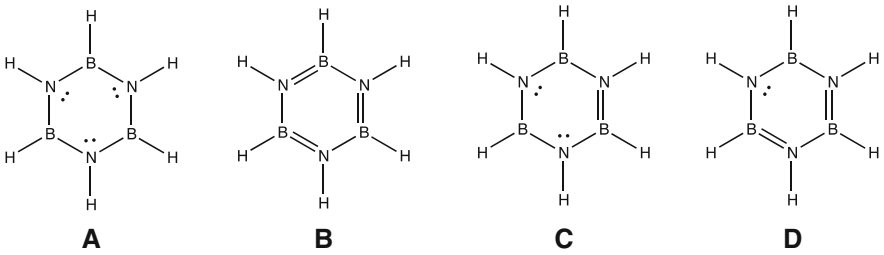


Fig. 3 VB structures considered for borazine in various VB calculations

For a VBCI wavefunction with strictly atomic orbitals, where structures with three (structure **A**, Fig. 3), two (structure **C**), one (structure **D**) and no (structure **B**) lone pairs are included among other (Dewar) structures (30 structures in total), a considerable resonance energy of 54 kcal/mol was found, relative to the energy obtained with a wavefunction consisting of localised borazine structures (8 in total) [57]. For benzene, with a similar treatment, a resonance energy of 46 kcal/mol was found, indicating that borazine has as much aromatic character as benzene. In borazine, the dominant structure **A** has a weight of 0.17. Each of the six symmetry equivalent structures **C** has a weight of 0.08. The two benzene-like structures **B** have only low weights of 0.02 each, and the structures **D** have a weight of 0.04 each. If only four structures are considered in the VBCI calculation (structures **A**, **B** and another structure with the three lone pairs on the B atoms), a resonance energy of 15 kcal/mol is found [58]; structure **A** has a weight of 0.90 in this calculation, while the structures **B** have each a weight of 0.05. The major difference between these two calculations that can explain the differences in resonance energies and weights is the number of structures included in the two VB wavefunctions. The former calculation included 30 structures, while in the latter, only 4 structures were considered in the VB wavefunction. At first sight, these two sets of results suggest that either or both are inadequate to give a qualitatively valid picture of the aromaticity of borazine.

Improving the 4-structure VB calculation by including orbital optimisation without any restrictions, either by using the spin-coupled VB method [60] or by using the VBSCF delocal method [58, 59], yields another set of substantially different results: the resonance energy now becomes negligible (0.2 kcal/mol), and the orbitals distort significantly (Fig. 4). The two sets of p_π-orbitals are now both centred on nitrogen, one p_π-orbital is very localised, while the other p_π-orbital is delocal onto the neighbouring boron centres and is in fact a correlating orbital. In the wavefunction, only one structure is important, which corresponds to **A**. In contrast to the findings with local orbitals, borazine does not exhibit any appreciable stabilisation through resonance and should therefore be considered as nonaromatic. This last conclusion, based on the spin-coupled VB calculations, is also in line with computational studies of the magnetic properties of borazine: according to ring current calculations, the lone pairs of borazine are intrinsically localised [58, 59, 62]. The results described above also demonstrate the danger of

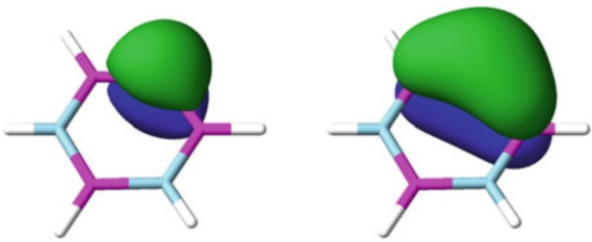


Fig. 4 The delocal VBSCF orbitals on nitrogen in borazine

applying orbital restrictions; constrained calculations may lead to the wrong conclusions.

Other six-membered heterocycles $B_3P_3H_6$, $Al_3N_3H_6$, $Al_3P_3H_6$, $B_3O_3H_3$ and $B_3S_3H_3$ have also been considered [58, 60]. All but $B_3P_3H_6$ showed similar behaviour as borazine: very small resonance energies were found using the VBSCF delocal method, and only one VB structure contributed significantly to the wavefunction, and all can be best described as being essentially nonaromatic. $B_3P_3H_6$ has more resemblances to benzene: with local orbitals, the benzene-like structures **B** have a considerable weight of 0.21 each, while structure **A** has a weight of 0.58. The resonance energy is 32 kcal/mol. In a spin-coupled description, the resonance energy drops to 13 kcal/mol. The spin-coupled VB orbitals distort not as severely as they did in the case of borazine: the p_π-orbitals on phosphorus are very localised, but there are also p_π-orbitals localised on boron centres, though these are delocal onto the neighbouring phosphorous centres. These findings are again in agreement with the calculated current density maps: localised currents around the formal lone pairs are found for all considered inorganic benzenes, except for $B_3P_3H_6$, which shows a diatropic ring current [58]. It should be noted that these VB calculations for $B_3P_3H_6$ were performed on the D_{3h} symmetric geometry, which is not a real minimum (1 imaginary frequency of $96i$ cm^{-1}, a_2'' symmetry). Current density calculations on the local C_{3v} minimum geometry showed that current changes from a global circulation to a set of localised islands, corresponding to the axial lone pairs of the phosphorus atoms [58].

Four ($B_2N_2H_4$) and eight ($B_4N_4H_8$) membered azabora-heterocycles showed similar characteristics as borazine. The most important structure in the VB calculations is the structure with two and four lone pairs on nitrogen for $B_2N_2H_4$ and $B_4N_4H_8$, respectively. The spin-coupled wavefunction can best be described in terms of lone pairs on nitrogen, each with a correlated pair of π-electrons. There is neither resonance nor resonance energy in these molecules.

3.2 Borabenzene

Borabenzene (C_5H_5B, Fig. 5), which has six π-electrons, just like benzene, has also been studied using the spin-coupled VB method [63]. Here, not only the π-bonds have been studied using VB but also the σ-bonds around the boron atom.

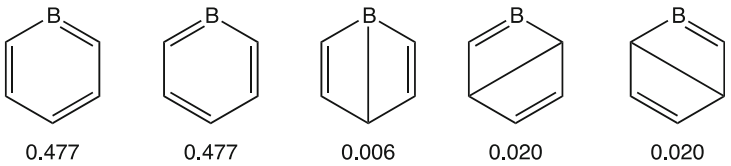


Fig. 5 The VB structures of borabenzene (C_5H_5B) together with the Gallup and Norbeck weights

The spin-coupled wavefunction accounts for *ca*. 90 % of the correlation energy recovered by the π-electron CASSCF(6,6) calculation [63], indicating that the spin-coupled wavefunction provides a reliable source of information about the π-electron system. The spin-coupled π-orbitals resemble those of benzene. There are, however, small differences; especially the boron p_π-orbital is more diffuse and more delocal onto its neighbouring carbon atoms, whereas the carbon p_π-orbitals centred on the carbon atoms next to the boron atom are more localised and have only delocalisation tails to their neighbouring carbon atoms.

The spin-coupled wavefunction also provides information about the importance of the different spin-coupling patterns: according to the Gallup and Norbeck weights [20], the two Kekulé-benzene-like structures contribute the most to the wavefunction (Fig. 5). The structure weights and orbital shapes of the π-system of borabenzene are similar to those of benzene, leading to the conclusion that the π-system of borabenzene should be considered as aromatic and the high reactivity of this molecule is therefore not caused by the π-system.

To further inquire about the high reactivity of this molecule, the σ-B-C bonds have been considered by the spin-coupled method. Surprisingly, it was found that the orbitals that form the B-C bond are not oriented along the B-C axes but tend to bend to the outer side of the hexagon formed by the B and C atoms. The bending of these orbitals leads to the formation of an electron-deficient site on the ring, which is the preferred attachment site for Lewis bases.

3.3 Nitrogen-Containing Aromatic Heterocycles

The spin-coupled VB description of the nitrogen-containing analogues of benzene (Fig. 6) is remarkably similar to each other [64] and similar to that of benzene [65]. For all but **2**, the two Kekulé structures are equally important (Table 1). For **2**, the Kekulé structure with the double bond between the two nitrogen atoms is less important, and the structure with C = N bonds is preferred. The resonance energy of **2** is also considerably lower than that of **1**, **3** and **4** (Table 1), which have a similar resonance energy as benzene. The spin-coupled orbitals on carbon and nitrogen of **1**, **3** and **4** are predominantly atomic in nature (2p-orbitals) but have small delocalisation tails on neighbouring atoms. The $2p_C$-orbitals of **2** are slightly differently distorted: the $2p_C$-orbitals next to a nitrogen atom are more distorted

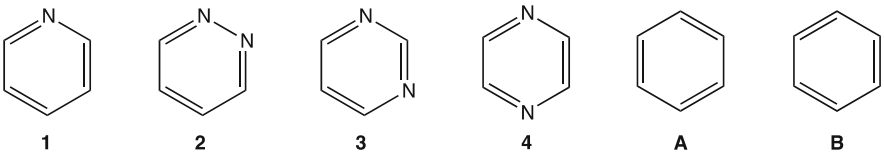


Fig. 6 The nitrogen-containing six-membered rings **1**–**4**, together with the spin-coupling schemes **A** and **B**

Table 1 The weights of the Kekulé structures **A** and **B** (Fig. 6) and the resonance energies (kcal/mol) for **1**–**4**; for comparison, the values for benzene are also included

Compound	W_A	W_B	$\sum W_{Dewar}$	E_{res}
1	0.400	0.400	0.200	22
2	0.205	0.540	0.255	13
3	0.400	0.400	0.200	23
4	0.400	0.400	0.200	23
Benzene	0.405	0.405	0.190	22

towards their neighbouring nitrogen atoms, and the $2p_N$-orbitals are more localised on nitrogen.

In this work [64], it had been concluded that the special stability of these aromatic molecules arises from the mode of coupling of the electron spin, as shown by the resonance energies, and not from any supposed delocalisation of the orbitals. The small distortions of the orbitals have the further consequence that the contribution of ionic structures disappears almost completely.

Pyrazine (**4**) has also been studied using the BLW approach, together with other azaacenes (Fig. 7) and the corresponding dihydroazaacenes [66]. The azaacenes all have $4n + 2$ π-electrons, while the dihydroazaacenes have $4n$ π-electrons. In Table 2, the BLW resonance energies (defined as the energy difference between the fully conjugated compound and its most stable resonance contributor; thus actually, it is a delocalisation energy) are listed together with the extra cyclic resonance energies (ECREs). The latter is defined as the difference between the resonance energies of a cyclic conjugated compound and an acyclic polyene either with the same number of double bonds or with the same number of diene conjugations. The ECRE measures the aromatic stabilisation (destabilisation) energy for cyclic conjugated compounds [50]. Positive ECREs indicate aromaticity and negative ECREs indicate anti-aromaticity.

Using the BLW procedure, large resonance energies are found for **4**–**8**. Note that the BLW resonance energy of **4** is substantially higher than the resonance energy obtained using the spin-coupled approach (Table 1), but also note that these two resonance energies cannot be directly compared, due to the different natures of the approaches. The BLW resonance energies of the dihydro counterparts are also considerable. The ECREs, which measure only the aromatic stabilisation, are considerably lower and, for **4**, more in line with the spin-coupled resonance energy. For all considered compounds, except **4-H$_2$**, the ECREs are positive, indicating aromatic character. This would imply that the dihydro derivatives, even though they

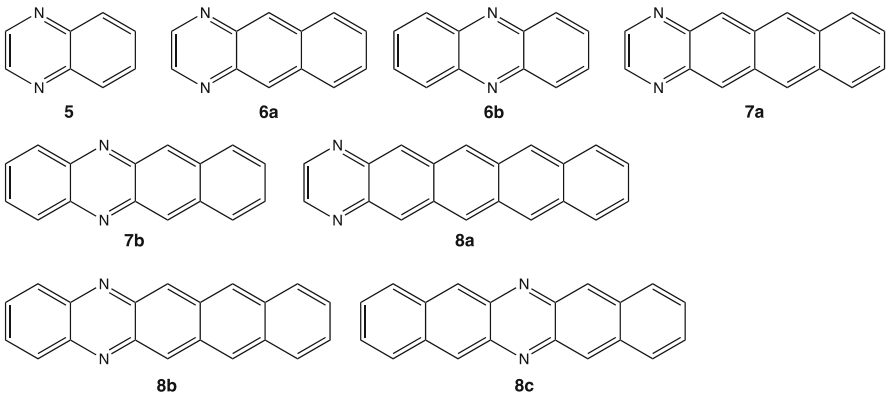


Fig. 7 Structure of the studied azaacenes [66]

Table 2 Block-localised-wavefunction resonance (delocalisation) energies (RE, kcal/mol) and extra cyclic resonance energies (ECRE, kcal/mol) for compounds **4**–**8** and their (planarised) dihydro counterparts [66]

Compound	RE	ECRE	Compound	RE	ECRE
4	60	30	**4-H$_2$**	57	−12
5	111	49	**5-H$_2$**	124	20
6a	157	59	**6a-H$_2$**	179	39
6b	154	59	**6b-H$_2$**	191	52
7a	202	68	**7a-H$_2$**	227	52
7b	200	70	**7b-H$_2$**	244	68
8a	246	71	**8a-H$_2$**	274	65
8b	243	70	**8b-H$_2$**	291	83
8c	242	77	**8c-H$_2$**	296	85

have an anti-aromatic ($4n$) π-electron count, are stabilised by aromaticity [66]. Unfortunately, the systems **5**–**8** have not yet been investigated by the spin-coupled VB or VBSCF delocal method.

3.4 Heterocyclic Five-Membered Rings

The π-electron systems of five-membered ring molecules **9**–**12** (Fig. 8) have been the subject of a spin-coupled study [67]. In **9**–**11**, five of the spin-coupled orbitals are localised, one on each carbon atom and one on the heteroatom; the sixth spin-coupled orbital is also associated with the heteroatom, but this orbital showed significant delocalisation onto the neighbouring carbon atoms. Thiazole (**12**) shows, like pyrrole, five localised orbitals, one on each atom. The sixth orbital, mainly associated with the sulphur atom, shows a very large distortion and is significantly delocal onto the neighbouring carbon atoms. In the series, it was found that only two of the five possible Rumer structures are important, viz. the

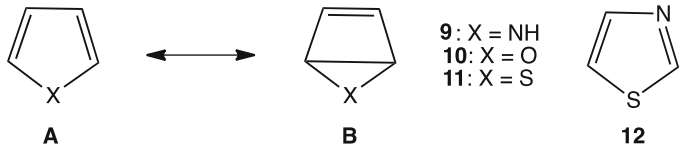


Fig. 8 The resonance structures **A** and **B** considered for the five-membered rings **9–12**

Table 3 Weights of the structures **A** and **B**, the resonance energy (E_{res}, kcal/mol), the BLW resonance energy (RE, kcal/mol) and the extra cyclic resonance energy (ECRE[a], kcal/mol) for compounds **9–12** and other selected five-membered rings

Compound	W_A	W_B	E_{res}	RE	ECRE	Compound	RE	ECRE
9	0.80	0.20	7	78	18	X = BH	16	−8
10	0.85	0.15	3	62	13	X = PH	73	18
11	0.88	0.12	6	51	13	$X = SiH^+$	23	−16
12	0.85	0.11	8			$X = SiH^-$	96	17
X = AlH				13	−4	$X = SiH_2$	17	−1

[a]The ECRE values are the ECRE2 values of reference [50]

structures **A** and **B** (Fig. 8). In all four cases, structure **A** has the highest weight (Table 3). These five-membered rings are also stabilised by resonance; resonance energies between 3 kcal/mol and 8 kcal/mol are found. Furan (**10**) showed the lowest resonance energy. The orbitals associated with the oxygen atom are more localised than those of pyrrole (**9**). The orbitals associated with the sulphur atom of thiophene (**11**) are more delocal than those of pyrrole. Although structure **A** is more dominant in thiophene, its resonance energy is somewhat higher. It seems that there is no simple relation between the resonance energy and the weights of the structures; apparently, the interaction matrix elements between the structures are different for different compounds. In conclusion [67], the spin-coupled descriptions of **9–12** are very similar to those of benzene and for other six-membered heterocycles. The π-system is adequately described by six non-orthogonal orbitals that turn out to be rather localised, and the stability results from the mode of spin coupling. The orbitals on the heteroatoms show very significant distortions towards its neighbours; this distortion depends on the electronegativity of the atom. Furthermore, the ionisation potentials of these molecules were accurately predicted using the spin-coupled theory.

The five-membered rings **9–11**, together with other five-membered heterocycles (Table 3), were also the subject of a study of the resonance energy, calculated with the BLW approach [50]. The BLW resonance energies (RE, Table 3) are substantially higher than the resonance energies calculated using the spin-coupled VB method. The calculated stabilisation due to aromaticity (ECRE, Table 3) is considerably smaller, and the latter value relates better to the aromaticity of the molecules. The ECRE of the five-membered compounds showed satisfactorily a correlation with nucleus-independent chemical shift (NICS) [68] values. Note that neither the RE nor the ECRE show the same trends as the spin-coupled resonance energies.

3.5 Three-Membered Rings

The σ-bonds of a series of three-membered rings (Fig. 9) have also been studied using the spin-coupled VB approach [69] in order to explain the unusual properties, such as a low strain energy and short carbon-carbon bond lengths. Two different models to describe the bonding in cyclopropane (**13**) have been proposed. In the model proposed by Walsh [70], the carbon atoms are sp^2 hybridised, with one of the three sp^2 hybrids pointing towards the centre of the ring, and these orbitals form three molecular orbitals, all concentrated within the interior of the ring (Fig. 10a) of which one is doubly occupied. Each carbon atom has in addition to the sp^2 hybrids a p-orbital oriented tangentially with respect to the ring (Fig. 10a); these orbitals form three molecular orbitals of which the lowest two are occupied. This type of configuration has been described as a closed-shell system with Hückel and Möbius aromatic subshells [71]. In the second model, suggested by Coulson and Moffitt [72], three bent C-C bonds are formed from six equivalent hybrids (Fig. 10b). Molecular orbital theory cannot distinguish the different models. The spin-coupled VB study would give the definite orbital description of the bonding in these strained three-membered rings.

It was demonstrated that the spin-coupled description, based on a single orbital product, accounts for the most important correlation effects [69]. The optimised spin-coupled orbitals resemble hybrid orbitals. The bending is visible as these hybrids are not directed along the C-C bond axes. The degree of bond bending is

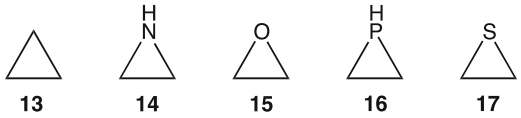


Fig. 9 The three-membered rings under study. Cyclopropane (**13**), aziridine (**14**), ethane oxide (**15**), phosphirane (**16**) and thiirane (**17**)

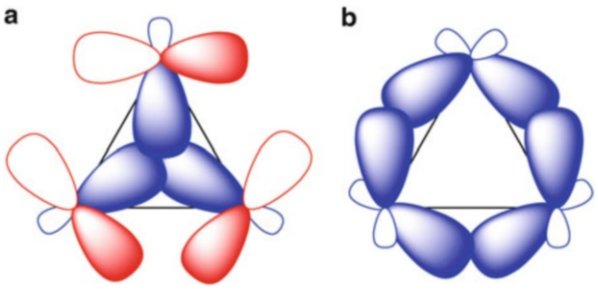


Fig. 10 Proposed bonding models for cyclopropane. (**a**) The bonding model of Walsh [70], with an inner two-electron Hückel subshell and an outer four-electron Möbius aromatic subshell. (**b**) The model of Coulson and Moffitt [72], with three bent carbon-carbon bonds constructed from six equivalent hybrid orbitals

dependent on the heteroatom: for **16** and **17**, the deformation is not discernible by eye. This observation can be related to the geometries of **16** and **17**, in which the heteroatom is placed at a longer distance from the C-C moiety. The shapes of the orbitals of the heteroatoms differ from those of carbon: the heteroatom hybrids have less s-character and have predominantly p-character. The weights of the different spin functions revealed that the wavefunctions of all molecules **13**–**17** were dominated by one spin-coupling pattern, viz. the one that corresponds to "perfect pairing". These spin-coupled calculations show that the bonding in the heterocycles is very similar to the bonding in cyclopropane. The σ-bonding in these three-membered rings does not differ much from ordinary σ-bonds, except that, due to the geometry, the orbitals have to bend in order to reduce strain and to be able to enhance the overlap.

3.6 Four-Membered Nitrogen-Containing E_2N_2 (E = S, Se, Te) Rings

The 6π-electron four-membered heterocycles E_2N_2 (E = S, Se, Te) have been studied by Braïda, Lo and Hiberty using the standard VB and breathing orbital VB (BOVB) methods to settle the controversies on their diradical character and to elucidate the nature of the radical sites [73]. In their calculations, they use strictly atomic orbitals, in contrast to the previous spin-coupled VB [74] and CASVB [75] calculations of Cooper and co-workers. The considered resonance structures are depicted in Fig. 11. Two of these six structures are diradical structures. According to the spin-coupled calculations [74], the diradical structure **b** best describes the bonding in S_2N_2.

In contrast, in the VBSCF calculations [73] for the series of molecules, the largest weight is obtained for structure **a** (Table 4), with the radical centres on the nitrogen atoms. The structure **a** becomes more dominant with increasing atomic number down the series. The structure **b**, with the radical centres on E, is less important. Upon inclusion of dynamic electron correlation via the breathing orbital VB approach, the weights of structure **a** decrease, the weights of the covalent structures **c**–**e** increase, and the weights of structure **b** increase.

The most stable structure in the series is structure **a**. The molecules possess significant resonance energies, which is largest for S_2N_2, confirming the aromatic nature of these species. The stability in these molecules comes from the resonance

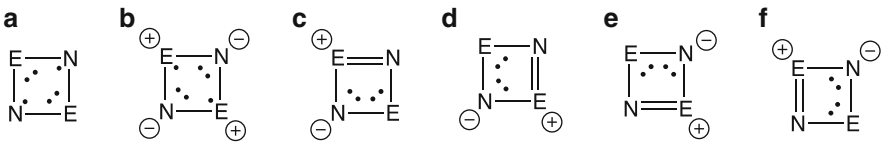


Fig. 11 The VB structures for the four-membered E_2N_2 (E = S, Se, Te) rings

Table 4 The weights of the VB structures for E_2N_2 and the resonance energy (kcal/mol) calculated with VBSCF and with BOVB

	VBSCF				BOVB			
Compound	a	b	c–e	E_{res}	a	b	c–e	E_{res}
S_2N_2 (**18**)	0.473	0.075	0.113	53	0.349	0.097	0.139	73
Se_2N_2 (**19**)	0.567	0.052	0.095	37	0.415	0.085	0.125	53
Te_2N_2 (**20**)	0.649	0.034	0.079	27	0.468	0.074	0.115	40

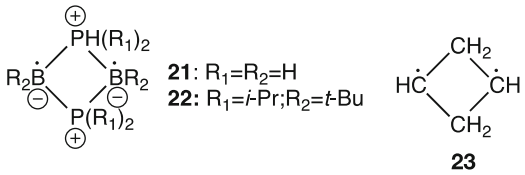


Fig. 12 The structure of 1,3-diphospha-2,4-diboretane diradicals (**21** and **22**) and of 1,3-cyclobutanediyl (**23**)

between all VB structures, the diradical and covalent ones. This study shows that aromaticity and diradical character can coexist. The interpretation of the VBSCF and BOVB studies seems to contradict the interpretation of the spin-coupled results: however, the authors argue that the obtained VBSCF/BOVB wavefunction does not differ much from the spin-coupled wavefunction but that the delocal nature of the spin-coupled orbitals obscured the interpretation [73]. Note that the resonance energies of the VBSCF and BOVB studies show the same trend in the series.

3.7 1,3-Diphospha-2,4-Diboretane Diradicals

The isolation of stable singlet diradicals containing boron and phosphorus (**22**, Fig. 12) [76] prompted theoretical studies of the electronic structure of these species [77–79]. It was found that **21** has much less diradical character than the carbon analogue **23**, based on the natural orbital populations. Whereas the natural orbital occupation numbers of **23** were found to be 1.14 and 0.86, indicating substantial diradical character, for **21**, the natural orbital occupation numbers were only 1.78 and 0.22 [77, 78], which is considerably smaller than would be obtained for a pure diradical. The larger interaction between the radicals on the boron centres was explained in terms of through-bond interactions.

The stabilisation of the diradicals by the through-bond interactions was further investigated using local (strictly atomic) and delocal VBSCF [80] calculations on **21** and **23**. For the local model, delocalisation of the singly occupied p-orbitals centred on BH/CH to the PH_2/CH_2 is prohibited. The analysis of the natural orbital populations showed that both **21** and **23** are essentially pure diradicals (Table 5). The overlap between the p-orbitals for both compounds is similar. This situation reflects the case when the through-bond interaction is "switched off". Removing the restriction of strictly atomic orbitals, thus allowing for delocalisation of the

Table 5 The total energies, natural orbital occupation numbers and the overlap between the singly occupied p-orbitals, calculated at the delocal VB/6-31G, and strictly atomic VB/6-31G levels of VBSCF

21	Delocal VBSCF		Strictly atomic VBSCF	
Total energy (Hartree)	−734.071659		−734.027702	
Occupation numbers	1.6522	0.3478	1.1553	0.8447
Overlap	0.37		0.14	
23				
Total energy (Hartree)	−154.762964		−154.738316	
Occupation numbers	1.0596	0.9404	1.0937	0.9063
Overlap	0.03		0.12	

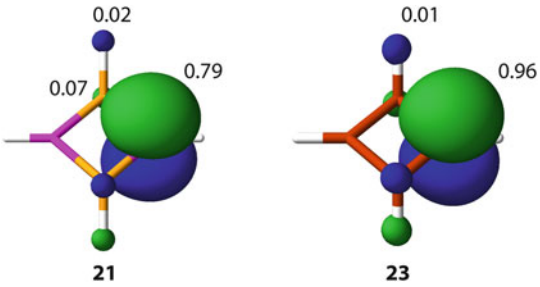


Fig. 13 Contour plots of the delocal VBSCF optimised singly occupied p-orbitals, together with the Mulliken population on the atoms for **21** and **23** [80]

p-orbital to the neighbouring atoms, of course gives an energy lowering for both compounds. Compound **23** lowers 15.5 kcal/mol in energy, while for **21**, the energy lowering is 27.6 kcal/mol. The overlap between the final VBSCF p-orbitals is substantially higher for **21** than for **23**; for **23** even a decrease in overlap is found compared to the local VB calculation. According to the natural orbital occupations, **23** remains a pure diradical, but the closed-shell character of **21** re-emerges. The final delocal VBSCF orbitals of **21** are more delocal onto the PH_2 groups (Fig. 13) than those of **23** onto the CH_2 groups. This is also corroborated by a Mulliken population analysis: the p-orbital of CH in **23** is much more localised than the p-orbital of BH in **21**.

3.8 Bis(nitronyl) Nitroxide: Ullman's Biradicals

In the field of molecular magnets, the organic nitroxide radicals have attracted considerable attention due to their stability. In a theoretical study of the magnetic coupling between the radical centres in coupled nitronylnitroxide radicals (Fig. 14), the CASSCF wavefunction was interpreted in terms of the orthogonal VB structures [81]. First, a study of the radical **24** was performed. The CASSCF(3,3) orbitals were localised, and the CASSCF(3,3) wavefunction was written in Slater determinants

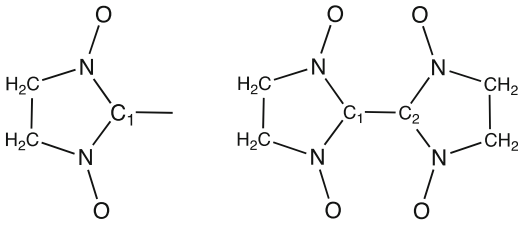


Fig. 14 The structure of the nitronylnitroxide radical (**24**) and the simplest variant of Ullman's biradicals (**25**)

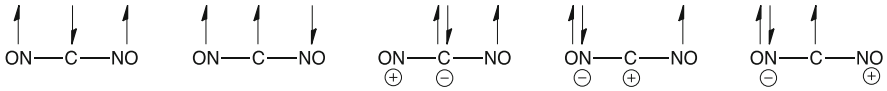


Fig. 15 Graphical representations of the VB structures of **24**

expressed in these localised orbitals. The localised orbitals could be interpreted as two antibonding NO π^*-orbitals and a $2p_\pi$-orbital centred at C_1 (Fig. 14). The analysis showed that the neutral structures (Fig. 15) are lowest in energy, and their combined weight in the wavefunction is 0.66. This analysis has been used to obtain model Hamiltonian parameters for the study of magnetic interactions.

Both the singlet and triplet states of the biradical **25** were analysed. The triplet state can be described as an antisymmetrised product of the two fragment wavefunctions: structures that do not result from combining the structures of the fragments turn out to have a negligible weight. The singlet state was, surprisingly, found to be quite similar to the triplet state: the energy and the weights of the structures are well approximated by the antisymmetrised product of the wavefunctions. A correlation between the spin density on C_1/C_2 and Heisenberg exchange parameter J was found, indicating that the interaction of the spin excess on these carbon atoms is responsible for a large part of the magnetic coupling. The results have been used to extract a set of parameters for the fragment which can be of use for larger magnetic systems in studies based on model Hamiltonians.

3.9 Hyperconjugation and Anomeric Effects

The so-called anomeric effect is the preference for electronegative substituents bonded to the carbon atom to occupy the axial position (Fig. 16) in the chair conformation of a monosaccharide. This preference for the axial position can be explained by the invocation of the hyperconjugation model: delocalisation from the oxygen lone pair to the empty antibonding σ_{CY}^*-orbital provides extra stabilisation.

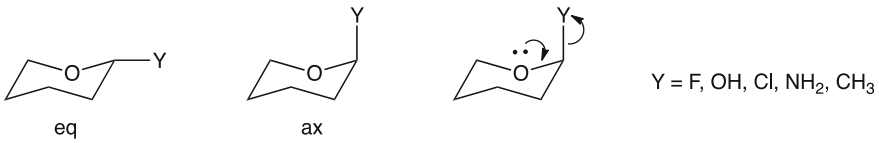


Fig. 16 The equatorial (*eq*) and axial positions (*ax*) in the considered tetrahydropyrans. In the *right* figure, the hyperconjugation model that is based on the stabilising interaction between the lone pair of the oxygen and the antibonding σ_{CY}^{*} is depicted

However, an alternative explanation can be given in terms of an electrostatic model and dipole-dipole interactions. An extensive block-localised wavefunction study was undertaken to discriminate between these two models [82]. In this study, the BLW approach was used to calculate the electron delocalisation effects and steric effects for substituted tetrahydropyrans and dimethoxymethane and the reference molecules dimethylether and substituted cyclohexanes. The reference molecules have no anomeric effects.

For different conformers of dimethoxymethane, the BLW resonance energy (i.e. the delocalisation energy) was calculated, and it was found that the overall stabilisation by delocalisation in the *trans* conformer was slightly larger than in the *gauche* conformer, while the *gauche-gauche* conformer is the most stable one. This finding suggests that the stability of the *gauche-gauche* conformer is not caused by hyperconjugation.

The BLW study of the substituted tetrahydropyrans showed that the relative stabilities of the axial and equatorial conformers were not changed by the exclusion of the delocalisation effects, even though the molecules have a substantial delocalisation energy. Thus, the anomeric effect is still present when intramolecular delocalisation is quenched: the hyperconjugation model is unable to explain the anomeric effect [82]. A detailed energy analysis revealed that the steric effect dominates the conformational preferences, and an alternative explanation such as the electrostatic model must therefore be introduced to interpret the anomeric effect.

4 Concluding Remarks

The illustrative results presented here show that the valence bond theory is able to give detailed information about the electronic structure of heterocyclic systems and is therefore a strong tool for the interpretation of the electronic structure in terms of chemical concepts like bond bending, hyperconjugation and aromaticity. These concepts can then be used to explain observed properties and reactivity. An example discussed here is the anomeric effect.

Advances in modern supercomputing allow the treatment of larger molecules using the valence bond method with optimised orbitals. Advances in valence bond

theory, such as the incorporation of electron correlation in valence bond methods [83–85], the Atoms in Valence Bond method [42] and quadratically convergent VBSCF methods [86], enhance the applicability of the valence bond methods to solve new, electronic structure-related, chemical problems.

In the heart of the valence bond theory, it is the spin coupling of orbitals that yields bond formation. Stabilisation by resonance in the valence bond theory is a result of the lowering of the total energy due to the interaction of different spin-coupled structures. As such, valence bond wavefunctions are multi-determinant wavefunctions and bear a close resemblance to MCSCF wavefunctions. Although one refers to the block-localised wavefunction approach as the simplest form of the valence bond theory, the BLW approach measures the effect of delocalisation of the orbitals in a single Slater determinant. This is an effect caused by limiting the one-electron space. The BLW approach thus relates to the Hartree-Fock method.

Acknowledgement RWAH and RB acknowledge the Zernike Institute for Advanced Materials for the financial support ("Dieptestrategie" programme).

References

1. Coulson CA (1961) Valence. Oxford University Press, London
2. Hall GG (1991) The Lennard-Jones paper of 1929 and the foundations of molecular orbital theory. In: Löwdin P-O, Sabin JR, Zerner MC (eds) Advances in quantum chemistry, vol 22. Academic Press, San Diego, pp 1–6
3. Hund F (1927) Zur Deutung der Molekelspektren. I. Z Physik 40:742–764
4. Lennard-Jones JE (1929) The electronic structure of some diatomic molecules. Trans Faraday Soc 25:668–686
5. Mulliken RS (1928) The assignment of quantum numbers for electrons in molecules. I. Phys Rev 32:186–222
6. Boys SF (1960) Construction of some molecular orbitals to be approximately invariant for changes from one molecule to another. Rev Mod Phys 32:296–299
7. Edmiston C, Ruedenberg K (1965) Localized atomic and molecular orbitals. II. J Chem Phys 43:S97–S116
8. Pipek J, Mezey PG (1989) A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. J Chem Phys 90:4916–4926
9. von Niessen W (1972) Density localization of atomic and molecular orbitals. I. J Chem Phys 56:4290–4297
10. Hiberty PC, Shaik S (2007) A survey of recent developments in ab initio valence bond theory. J Comput Chem 28:137–151
11. Shaik S, Hiberty PC (2004) Valence Bond, its history, fundamentals, and applications: a primer. In: Lipkowitz KB, Larter R, Cundary TR (eds) Reviews in computational chemistry. Wiley-VCH, New York, pp 1–100
12. Lewis GN (1916) The atom and the molecule. J Am Chem Soc 38:762–785
13. Heitler W, London F (1927) Wechselwirkung neutraler Atome und homöopolare Bindung nach der Quantenmechanik. Z Physik 44:455–472
14. Heitler W, Rumer G (1931) Quantentheorie der chemischen Bindung für mehratomige Moleküle. Z Physik 68:12–41
15. Slater JC (1931) Directed valence in polyatomic molecules. Phys Rev 37:481–489

16. Pauling L (1931) The nature of the chemical bond. Application of results obtained from the quantum mechanics and from a theory of paramagnetic susceptibility to the structure of molecules. J Am Chem Soc 53:1367–1400
17. Pauling L (1960) The nature of the chemical bond. Cornell University Press, Ithaca
18. Gallup GA (2002) Valence bond methods: theory and applications. Cambridge University Press, Cambridge
19. Chirgwin BH, Coulson CA (1950) The electronic structure of conjugated systems. 6. Proc R Soc Lond Ser A 201:196–209
20. Gallup GA, Norbeck JM (1973) Population analyses of valence-bond wavefunctions and BeH2. Chem Phys Lett 21:495–500
21. Pauling L, Wheland GW (1933) The nature of the chemical bond. V. The quantum-mechanical calculation of the resonance energy of benzene and naphthalene and the hydrocarbon free radicals. J Chem Phys 1:362–374
22. Broer R, Hozoi L, Nieuwpoort WC (2003) Non-orthogonal approaches to the study of magnetic interactions. Mol Phys 101:233–240
23. Hozoi L, de Vries AH, Broer R, de Graaf C, Bagus PS (2006) Ni 3s-hole states in NiO by non-orthogonal configuration interaction. Chem Phys 331:178–185
24. Coulson CA, Fischer I (1949) XXXIV. Notes on the molecular orbital treatment of the hydrogen molecule. Philos Mag 40:386–393
25. Goddard WA, Dunning TH, Hunt WJ, Hay PJ (1973) Generalized valence bond description of bonding in low-lying states of molecules. Acc Chem Res 6:368–376
26. Hay PJ, Hunt WJ, Goddard WA (1972) Generalized valence bond description of simple alkanes, ethylene, and acetylene. J Am Chem Soc 94:8293–8301
27. Hunt WJ, Hay PJ, Goddard WA (1972) Self-consistent procedures for generalized valence bond wavefunctions. Applications H3, BH, H2O, C2H6, and O2. J Chem Phys 57:738–748
28. Moss BJ, Bobrowicz FW, Goddard WA (1975) The generalized valence bond description of O2. J Chem Phys 63:4632–4639
29. Cooper DL, Gerratt J, Raimondi M, Wright SC (1987) The electronic structure of 1,3-dipoles: spin-coupled descriptions of nitrone and diazomethane. Chem Phys Lett 138:296–302
30. Cooper DL, Gerratt J, Raimondi M (1984) Studies of molecular states using spin-coupled valence-bond theory, Faraday Symp. Chem Soc 19:149–163
31. Cooper DL, Gerratt J, Raimondi M (1987) Modern valence bond theory. In: Lawley KP (ed) Advances in chemical physics: Ab initio methods in quantum chemistry, vol 69. Wiley, London, pp 319–397
32. Gerratt J (1971) General theory of spin-coupled wavefunctions for atoms and molecules. In: Bates D, Esterman I (eds) Advances in atomic and molecular physics, vol 7. Academic, New York, pp 141–221
33. Gerratt J, Raimondi M (1980) The spin-coupled valence bond theory of molecular electronic structure. I. Basic theory and application to the 2Σ+ states of BeH. Proc R Soc Lond A 371:525–552
34. Small DW, Head-Gordon M (2011) Post-modern valence bond theory for strongly correlated electron pairs. Phys Chem Chem Phys 13:19285–19297
35. Small DW, Head-Gordon M (2009) Tractable spin-pure methods for bondbreaking: local many-electron spin-vector sets and an approximate valence bond model. J Chem Phys 130:084103
36. van Lenthe JH, Balint-Kurti GG (1980) The valence-bond SCF (VB SCF) method: synopsis of theory and test calculation of OH potential energy curve. Chem Phys Lett 76:138–142
37. van Lenthe JH, Balint-Kurti GG (1983) The valence-bond self-consistent field method (VB-SCF): theory and test calculations. J Chem Phys 78:5699–5713
38. van Lenthe JH, Dijkstra F, Havenith RWA (2002) TURTLE – a gradient VBSCF program. Theory and studies of aromaticity. In: Cooper DL (ed) Valence bond theory, vol 10, Theoretical and computational chemistry. Elsevier, Amsterdam, pp 79–112

39. Dijkstra F, van Lenthe JH (1999) Gradients in valence bond theory. Chem Phys Lett 310:553–556
40. Dijkstra F, van Lenthe JH (2000) Gradients in valence bond theory. J Chem Phys 113:2100–2108
41. Havenith RWA (2005) Coupled valence bond theory. Chem Phys Lett 414:1–5
42. Zielinski ML, van Lenthe JH (2010) Atoms in valence bond – AiVB. Synopsis and test results. Chem Phys Lett 500:155–160
43. Rashid Z, van Lenthe JH, Havenith RWA (2012) Resonance and aromaticity: an ab initio valence bond approach. J Phys Chem A 116:4778–4788
44. Hiberty PC (1997) Reconciling simplicity and accuracy: compact valence bond wave functions with breathing orbitals. J Mol Struct (Theochem) 398–399:35–43
45. Hiberty PC, Flament JP, Noizet E (1992) Compact and accurate valence bond functions with different orbitals for different configurations: application to the two-configuration description of F2. Chem Phys Lett 189:259–265
46. Hiberty PC, Humbel S, Byrman CP, van Lenthe JH (1994) Compact valence bond functions with breathing orbitals: application to the bond dissociation energies of F2 and FH. J Chem Phys 101:5969–5976
47. Hiberty PC, Shaik S (2002) Breathing-orbital valence bond method – a modern valence bond method that includes dynamic correlation. Theor Chem Acc 108:255–272
48. Mo Y (2009) The resonance energy of benzene: a revisit. J Phys Chem A 113:5163–5169
49. Mo Y, Peyerimhoff SD (1998) Theoretical analysis of electronic delocalization. J Chem Phys 109:1687–1697
50. Mo Y, von R. Schleyer P (2006) An energetic measure of aromaticity and antiaromaticity based on the Pauling-Wheland resonance energies. Chem Eur J 12:2009–2020
51. Zielinski M, Havenith RWA, Jenneskens LW, van Lenthe JH (2010) A comparison of approaches to estimate the resonance energy. Theor Chem Acc 127:19–25
52. Roos BO (1987) The complete active space self-consistent field method and its applications in electronic structure calculations. Adv Chem Phys 69:399–445
53. Angeli C, Cimiraglia R, Malrieu J-P (2008) On the relative merits of nonorthogonal and orthogonal valence bond methods illustrated on the hydrogen molecule. J Chem Educ 85:150–158
54. Malrieu J-P, Guihéry N, Calzado CJ, Angeli C (2007) Bond electron pair: its relevance and analysis from the quantum chemistry point of view. J Comput Chem 28:35–50
55. Löwdin P-O (1950) On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. J Chem Phys 18:365–375
56. Löwdin P-O (1967) Group algebra, convolution algebra, and applications to quantum mechanics. Rev Mod Phys 39:259–287
57. Benker D, Klapötke TM, Kuhn G, Li J, Miller C (2005) An ab initio valence bond (VB) calculation of the π delocalisation energy in borazine, B3N3H6. Heteroat Chem 16:311–315
58. Engelberts JJ, Havenith RWA, van Lenthe JH, Jenneskens LW, Fowler PW (2005) The electronic structure of inorganic benzenes: valence bond and ringcurrent descriptions. Inorg Chem 44:5266–5272
59. Soncini A, Domene C, Engelberts JJ, Fowler PW, Rassat A, van Lenthe JH, Havenith RWA, Jenneskens LW (2005) A unified orbital model of delocalised and localised currents in monocycles, from annulenes to azabora-heterocycles. Chem Eur J 11:1257–1266
60. Cooper DL, Wright SC, Gerratt J, Hyams PA, Raimondi M (1989) The electronic structure of heteroaromatic molecules. Part 3. A comparison of benzene, borazine and boroxine. J Chem Soc Perkin Trans 2:719–724
61. Cyrañski M, von R. Schleyer P, Krygowski TM, Jiao H, Hohlneicher G (2003) Facts and artifacts about aromatic stability estimation. Tetrahedron 59:1657–1665
62. Steiner E, Fowler PW, Havenith RWA (2002) Current densities of localized and delocalized electrons in molecules. J Phys Chem A 106:7048–7056

63. Karadakov PB, Ellis M, Gerratt J, Cooper DL, Raimondi M (1997) The electronic structure of borabenzene: combination of an aromatic π-sextet and a reactive σ-framework. Int J Quantum Chem 63:441–449
64. Cooper DL, Wright SC, Gerratt J, Raimondi M (1989) The electronic structure of heteroaromatic molecules. Part 1. Six-membered rings. J Chem Soc Perkin Trans 2:255–261
65. Cooper DL, Gerratt J, Raimondi M (1986) The electronic structure of the benzene molecule. Nature 323:699–701
66. Wu JI, Wannere CS, Mo Y, von R. Schleyer P, Bunz UHF (2009) 4n π electrons but stable: N, N,-dihydrodiazapentacenes. J Org Chem 74:4343–4349
67. Cooper DL, Wright SC, Gerratt J, Raimondi M (1989) The electronic structure of heteroaromatic molecules. Part 2. Five-membered rings. J Chem Soc Perkin Trans 2:263–267
68. von R. Schleyer P, Maerker C, Dransfeld A, Jiao H, van Eikema Hommes NJR (1996) Nucleus-Independent chemical shifts: a simple and efficient aromaticity probe. J Am Chem Soc 118:6317–6318
69. Karadakov PB, Gerratt J, Cooper DL, Raimondi M (1995) Modern valence bond description of bonding in strained three-membered rings: cyclopropane, aziridine, ethene oxide, phosphirane and thiirane. J Mol Struct (Theochem) 341:13–24
70. Walsh AD (1949) The structures of ethylene oxide, cyclopropane, and related molecules. Trans Faraday Soc 45:179–190
71. Cremer D, Gauss J (1986) Theoretical determination of molecular structure and conformation. 20. Reevaluation of the strain energies of cyclopropane and cyclobutane carbon-carbon and carbon-hydrogen bond energies, 1,3 interactions, and σ-aromaticity. J Am Chem Soc 108:7467–7477
72. Coulson CA, Moffitt WE (1949) I. The properties of certain strained hydrocarbons. Philos Mag 40:1–35
73. Braïda B, Lo A, Hiberty PC (2012) Can aromaticity coexist with diradical character? An ab initio valence bond study of S2N2 and related 6π-electron four membered rings E2N2 and E4 2 + (E= S,Se,Te). Chem Phys Chem 13:811–819
74. Gerratt J, McNicholas SJ, Karadakov PB, Sironi M, Raimondi M, Cooper DL (1996) The extraordinary electronic structure of N2S2. J Am Chem Soc 118:6472–6476
75. Thorsteinsson T, Cooper DL (1998) Nonorthogonal weights of modern VB wavefunctions. Implementation and applications within CASVB. J Math Chem 23:105–126
76. Scheschkewitz D, Amii H, Gornitzka H, Schoeller WW, Bourissou D, Bertrand G (2002) Singlet diradicals: from transition states to crystalline compounds. Science 295:1880–1881
77. Jung Y, Head-Gordon M (2003) How diradicaloid is a stable diradical? Chem Phys Chem 4:522–525
78. Jung Y, Head-Gordon M (2003) Controlling the extent of diradical character by utilizing neighboring group interactions. J Phys Chem A 107:7475–7481
79. Seierstad M, Kinsinger CR, Cramer CJ (2002) Design of 1,3-diphospha-2,4-diboretane diradicals. Angew Chem Int Ed 41:3894–3896
80. Havenith RWA, van Lenthe JH, van Walree CA, Jenneskens LW (2006) Orbital interactions expressed in resonance structures: an approach to compute stabilisation of cyclobutanediyl diradicals. J Mol Struct (Theochem) 763:43–50
81. Angeli C, Calzado CJ, de Graaf C, Caballol R (2011) The electronic structure of Ullman's biradicals: an orthogonal valence bond interpretation. Phys Chem Chem Phys 13:14617–14628
82. Mo Y (2010) Computational evidence that hyperconjugative interactions are not responsible for the anomeric effect. Nat Chem 2:666–671
83. Wu W, Song L, Cao Z, Zhang Q, Shaik S (2002) Valence bond configuration interaction: a practical ab initio valence bond method that incorporates dynamic correlation. J Phys Chem A 106:2721–2726
84. Chen Z, Song J, Shaik S, Hiberty PC, Wu W (2009) Valence bond perturbation theory. A valence bond method that incorporates perturbation theory. J Phys Chem A 113:11560–11569

85. Song J, Wu W, Zhang Q, Shaik S (2004) A practical valence bond method: a configuration interaction method approach with perturbation theoretic facility. J Comput Chem 25:472–478
86. Rashid Z, van Lenthe JH (2013) A quadratically convergent VBSCF method. J Chem Phys 138:54105

Aromaticity of Organic and Inorganic Heterocycles

Ferran Feixas, Jordi Poater, Eduard Matito, and Miquel Solà

Contents

Abstract Heteroaromatic rings are present in many organic molecules. They can be found as part of the core of drugs or agrochemicals and in many important biochemical molecules. The last three centuries have brought important advances in heteroaromatic chemistry. Indeed, the first organic molecules with heteroaromatic rings were already synthesized in the middle of the nineteenth century. Then, the

F. Feixas
Department of Pharmacology, University of California, San Diego, La Jolla, CA, USA

J. Poater • E. Matito • M. Solà (✉)
Institut de Química Computacional i Catàlisi (IQCC) and Departament de Química, Universitat de Girona, Campus de Montilivi, Girona, Catalonia, Spain
e-mail: miquel.sola@udg.edu

F. De Proft, P. Geerlings (eds.), *Structure, Bonding and Reactivity of Heterocyclic Compounds*, Topics in Heterocyclic Chemistry 38, DOI 10.1007/978-3-642-45149-2_5

twentieth century witnessed the first inorganic heteroaromatic compound produced in the laboratory. And at the beginning of the present century, the first all-metal heteroaromatic cluster was detected. Here, we discuss the aromaticity of some of these heteroaromatic compounds using different descriptors of aromaticity, with special emphasis in those measures based on electron delocalization properties of the aromatic rings.

Keywords Aromaticity • Heteroaromaticity • Nucleus-independent chemical shifts • Electronic indices • Multicenter electronic indices • Harmonic oscillator model of aromaticity • Aromatic fluctuation index • Para delocalization index

1 Introduction

In 1856, August W. von Hofmann used for the first time the term aromatic applied to a chemical compound [1]. Ten years later, after his celebrated dream of the snake that seizes its own tail, August Kekulé proposed the molecular structure of benzene consisting in a six-membered carbon ring with alternating single and double bonds [2]. The first heteroaromatic molecules were synthesized already at that time. Thus, pyridine was synthesized by Thomas Anderson [3] in 1868 through studies on the distillation of bone oil and other animal matter. For many years, the study of aromaticity was limited to the realm of purely hydrocarbon chemistry, the so-called polycyclic aromatic hydrocarbons (PAH), and of heteroaromatic organic compounds. Interestingly, this latter group of compounds provides the main source of aromatic species. Balaban et al. reported that among the 20 million chemical compounds known to date, approximately half are heteroaromatic [4]. Some of these heterocyclic aromatic species are very important in biochemical processes, drugs, and agrochemicals.

The first synthesized inorganic heteroaromatic compound was borazine, $B_3H_6N_3$ [5]. It was obtained in 1926 by a reaction of diborane with ammonia. It results from the triple substitution of C = C bonds in benzene by isoelectronic and isostructural B - N units and, consequently, it is considered the inorganic analogous of benzene. Interestingly, single C = C by B - N substitution in benzene to form 1,2-dihydro-1,2-azaborine, a hybrid organic/inorganic benzene, was elusive and was achieved only 5 years ago [6]. On the other hand, metallabenzenes are heteroaromatic analogues of benzene or PAHs where one (or more) of the CH groups is (are) formally replaced by a transition metal with its ancillary ligands. In 1982 Roper et al. [7] synthesized the first metallabenzene, an osmabenzene, thus initiating a new group of heteroaromatic rings that contain metals in their structure, the so-called metalloaromatic compounds.

Boldyrev, Wang, and coworkers [8] detected in 2001 a series of bimetallic clusters containing Al_4^{2-}, face capped by an M^+ cation (M = Li, Na, Cu) having

all-metal aromaticity. These all-metal aromatic clusters were synthesized using a laser vaporization technique and characterized with their experimental and theoretical photoelectron spectra. The last decade has brought tremendous advancement in the field of all-metal and semimetal aromatic compounds [9–12]. In most of the cases, the aromatic rings of these new all-metal clusters are constituted by a single type of atom, like in Al_4^{2-}. However, some all-metal heteroaromatic rings were also described. In fact, the first detected all-metal heteroaromatic cluster, XAl_3^-(X = Si, Ge, Sn, and Pb), was also reported by Boldyrev, Wang, and coworkers [13] in the same year of the discovery of Al_4^{2-}. Other interesting examples of recently observed all-metal heteroaromatic compounds are $M_3O_9^-$ and $M_3O_9^{2-}$ with M = W and Mo (the first compounds showing d-orbital aromaticity) [14], $Ta_3O_3^-$ (the first metallic cluster showing δ-aromaticity) [15], and the more recent $Te_2As_2^{2-}$ species with conflicting aromaticity [16]. From a theoretical point of view, Tsipis et al. found f orbital aromaticity in heteroaromatic rings $U_n(\mu_2-X)_n$ (n = 3, 4; X = C, CH, NH) as well as in E@$[c-U_4(\mu_2-C)_4]$ (E = H(+), C, Si, Ge) and U@$[c-U_5(\mu_2-C)_5]$ clusters [17]. The discovery of the Al_4C^- is also remarkable, a heteroaromatic cluster with a tetracoordinated planar carbon atom [18, 19].

The research on heteroaromatic compounds started about 150 years ago, and it has continued with important contributions to science up to this date. There are several reviews devoted to organic heterocyclic chemistry where one can find quite complete revisions of the organic heteroaromatic chemistry [4, 20]. The aim of this chapter is much less ambitious. Here we simply summarize the computational works carried out in the last years in our group in relation to the application of different indicators used to quantify heteroaromaticity, both in organic and inorganic aromatic compounds.

2 The Problem of Measuring Aromaticity

We cannot unambiguously define the term aromaticity. However, it is generally accepted that aromaticity originates from the cyclic electron delocalization of electrons in a close circuit. Interestingly, aromaticity is often measured from the consequences of this peculiar electronic feature. For instance, the nucleus-independent chemical shift (NICS) measures the shielding, which originates from the circulation of electrons along the aromatic ring. The harmonic oscillator model of aromaticity (HOMA) [21] compares the bond distances in the ring with those of an aromatic molecule taken as a reference. Our group has been intensively involved in the design and calculations of several electronic aromaticity indices. In particular we have designed (a) FLU [22] that measures the electron fluctuation differences with respect to aromatic molecules by comparing contiguous electron delocalization indices [23] along the ring structures, (b) the para-delocalization index (PDI) which measures the electron delocalization across the six-membered ring structures [24], and (c) I_{NG} and I_{NB} that are the normalized versions of I_{ring} and MCI,

respectively. I_{ring}, which was developed by Giambiagi and coworkers [25], measures the multicenter delocalization index [26] which includes all the atoms in the ring, as they are connected in the ring. On the other hand, the MCI, which was developed by Bultinck et al. [27], sums up all the possible I_{ring} values arising from the permutations of the atoms in the ring.

As compared to classical organic aromatic compounds, the evaluation of aromaticity in heterocyclic compounds is much more complex. This complexity is due to many different factors. First, at variance with aromatic organic compounds, which only exhibit $\pi-$ electron delocalization, inorganic compounds present $\sigma-, \pi-$ and $\delta-$ or even $\phi-$ electron delocalization, thus giving rise to multifold aromaticity. This feature comes from the different chemical bonds arising in inorganic compounds that include *s, p, d*, and *f* orbitals.

Second, the concept of aromaticity was originally invoked (and used for many years exclusively) on the context of organic molecules. As a result, most of the methods to quantify aromaticity were developed for aromatic organic molecules, and they cannot be employed straightforwardly in inorganic clusters. For instance, all those methods that are based on reference values taken from organic molecules, such as HOMA and FLU, cannot be used in inorganic molecules, simply because the bonding patterns (e.g., the distance or the electron sharing) of many aromatic inorganic molecules have not been established. Furthermore, anyone would agree that the quintessential organic aromatic molecule is benzene and thus this molecule is taken as the paradigmatic example of how an organic aromatic should be in terms of geometry, electronic structure, or representative properties. Unfortunately there is no such a reference aromatic molecule in the inorganic realm. Other methods such as resonance energies (RE) or aromatic stabilization energies (ASE) are also very difficult to compute because of the lack of appropriate reference systems [28]. Altogether, there are very few methods to quantify aromaticity for inorganic compounds.

Among inorganic aromatic compounds, the most interesting compounds recently discovered are all-metal clusters. A few years ago, when the first all-metal clusters were synthesized, the aromaticity was assessed by means of simple electron counting based on $4n+2$ Hückel's rule or the calculation of NICS. Unfortunately the electron counting alone does not always provide evidence of aromaticity. For instance, Al_4^{2-} contains four σ electrons in spite of the fact that this molecule is σ- aromatic. In addition, NICS values are also known to provide the incorrect answer for some particular systems [29]. Recently, our group has been working on the use of multicenter aromaticity indices, which turn out to be the most reliable indicators of aromaticity both in the organic and inorganic realms [29, 30].

For the indices of aromaticity discussed in this chapter, the more negative the NICS, the lower the FLU or $FLU^{1/2}$, and the higher the PDI and MCI values, the more aromatic the ring is.

3 Computational Details

The calculations of the systems studied in this chapter have used many different computational methods and basis sets and the details are thus given in each section. However, all the electronic structure calculations were carried out using Gaussian 09 package [31] or the Amsterdam Density Functional (ADF) software [32, 33].

On the other hand, the calculation of the aromaticity indices has employed a plethora of programs. The NICS values have been calculated using the gauge-including atomic orbital (GIAO) [34, 35] implemented in Gaussian 09. The computation of the aromaticity indices calls for the calculation of the atomic overlap matrices (AOM) for the atomic partition employed. Two different atomic partitions have been used in these works: the quantum theory of atoms in molecules (QTAIM) partition due to Bader [36] and the Becke-rho partition that uses Becke's fuzzy partition scheme. The former has been calculated using AIMPAC software, while the latter uses the APOST-3D code [37, 38]. Finally, these AOMs are read and used by ESI-3D program [22, 39, 40] to generate the delocalization indices, the FLU, the PDI, and the various multicenter indices calculated in these works.

On the other hand, the Morokuma-like energy decomposition analysis (EDA) [41–43] was performed using the ADF package.

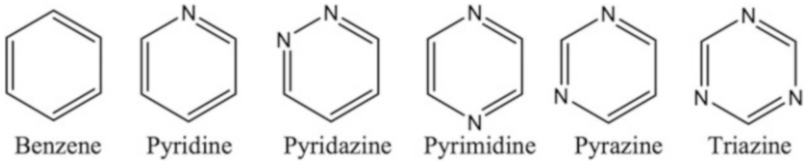
4 Aromaticity in Heterocyclic Organic Compounds

4.1 Six-Membered Heterocycles

In this section we analyze the performance of a set of well-known aromaticity indices to predict the aromatic character of a series of six heterocyclic rings that include benzene and other six-membered rings (6-MRs) that replace one, two, or three carbon atoms by nitrogen with respect to benzene's structure [30]. We anticipate a decrease of the aromaticity of these species as the number of replacements increase, i.e., the order given in Scheme 1.

This is exactly the trend obtained using the resonance energy (RE) criterion except for triazine. For the set of molecules with the same number of nitrogen atoms, the most aromatic are those with the largest number of N–N bonds [44]. Among the different tests suggested in Ref. [30] to assess the performance of aromaticity indices, reproducing the order given in Scheme 1 was the most difficult one for the aromaticity indices. None of the aromaticity indices was actually able to pass this test (Table 1).

The FLU fails only on finding pyrimidine more aromatic than pyridazine. The PDI fails to reproduce the expected aromaticity trend in the 6-MR heteroaromatic series. It considers pyrazine and pyridazine more aromatic than benzene itself. Since PDI measures the electron delocalization across the ring, the lone pairs of nitrogen, which do not account for the aromaticity of the ring but are part of



Scheme 1 Six-membered ring heteroaromatic species

Table 1 PDI, $FLU^{1/2}$, HOMA, NICS(0), NICS(1), and $NICS(1)_{zz}$ results. NICS values in ppm

System	PDI	$FLU^{1/2}$	HOMA	NICS (0)	NICS (1)	$NICS(1)_{zz}$	RE[a]
Benzene	0.103	0.000	0.989	−8.05	−10.22	−29.25	48.7
Pyridine	0.103	0.022	0.995	−6.82	−10.12	−28.59	44.1
Pyridazine	0.105	0.041	0.981	−5.35	−10.52	−28.28	43.9
Pyrimidine	0.101	0.032	0.999	−5.53	−10.00	−27.36	40.1
Pyrazine	0.108	0.042	0.996	−5.36	−10.24	−28.37	38.8
Triazine	0.096	0.043	1.000	−4.07	−9.66	−25.54	44.9

[a]From Ref. [21] in kcal mol^{-1}

nitrogen's basin, may increase the delocalization of C–N pair and give artificially larger PDI values. The HOMA, which uses only geometrical parameters, fails completely on giving a reasonable trend for the series. Apparently, the electronic effects on this series are more important than the geometrical features in order to describe the aromaticity of these species.

The NICS indices also fail on reproducing the aromaticity of these 6-MR heteroaromatic species. NICS(0) and $NICS(1)_{zz}$ find benzene and triazine the most and the least aromatic species of the series, but they fail to give the correct trend for the species with one or two nitrogen atoms. NICS(1) only identifies triazine as the least aromatic compound.

Similarly to FLU, the multicenter indices (MCI, I_{ring}, I_{NG}, I_{NB}) fail only on finding pyridazine more aromatic than pyridine. This result was obtained at the B3LYP/6-311++G(d,p) level of theory; however, the ordering gets completely messed up if MP2/6-311++G(d,p) is used [30]. On the other hand, if we use CCSD/6-311++G(d,p), the trend is perfectly reproduced, although some compounds in the series are found equally aromatic (see Table 2) [45]. The multicenter indices perform better than other aromaticity indices, but they are quite sensible to the level of theory for small differences as those reported in this heteroaromatic series [45].

4.2 Five-Membered Heterocycles

In this section we assess the performance of a set of well-known aromaticity indices to predict the aromatic character of a series of six heterocyclic 5-MRs,

Table 2 Multicenter indices (MCI, I_{ring}) at various levels of theory (HF, B3LYP, MP2, and CCSD with 6-311++G(d,p) basis sets) for the heterocyclic species in Scheme 1

	MCI				I_{ring}			
	HF	B3LYP	MP2	CCSD	HF	B3LYP	MP2	CCSD
Benzene	0.068	0.072	0.049	0.048	0.0451	0.0477	0.0339	0.0325
Pyridine	0.058	0.068	0.046	0.044	0.0392	0.0456	0.0317	0.0304
Pyridazine	0.056	0.069	0.045	0.044	0.0384	0.0469	0.0315	0.0304
Pyrimidine	0.049	0.065	0.046	0.042	0.0333	0.0440	0.0315	0.0291
Pyrazine	0.054	0.065	0.043	0.042	0.0349	0.0435	0.0291	0.0285
1,3,5-triazine	0.040	0.063	0.046	0.040	0.0275	0.0427	0.0315	0.0275

X = CH^-, NH, O, CH_2, BH, CH^+

Scheme 2 5-MR heteroaromatic species C_4H_4X (X = CH^-, NH, O, CH_2, BH, CH^+)

C_4H_4X (X = CH^-, NH, O, CH_2, BH, CH^+) [30] depicted in Scheme 2. According to the electron counting based on molecular orbital occupation, for X = CH^-, NH, O we have heterocycles with six π-electrons and, thus, these three rings should be considered aromatic. Moreover, since the nitrogen atom is less electronegative than the oxygen atom, a lower reduction of the electron delocalization for the NH group than for the O with respect to the totally delocalized cyclopentadienyl anion (CH^-) [4] is expected. Therefore, pyrrole is more aromatic than furan. On the other hand, the system with X = CH_2 is predicted to be a nonaromatic ring, whereas X = BH and CH^+ are both four π-electron antiaromatic compounds, with X = CH^+ being the one that presents the strongest paratropic ring currents and, thus, the most antiaromatic of the series. Thus, the expected order of aromaticity for these 5-MR heterocyclic systems is $CH^- > NH > O > CH_2 > BH > CH^+$. All the calculations have been carried out with Gaussian package [31], at the B3LYP/6-311++G(d,p) level of theory. The performance of ten descriptors of aromaticity based on different properties, i.e., FLU, MCI, I_{NB}, I_{ring}, I_{NG}, HOMA, ASE, NICS, NICS(1), and $NICS(1)_{zz}$, has been evaluated.

All the results obtained are gathered in Table 3. As can be seen, electronic multicenter descriptors of aromaticity, i.e., MCI, I_{NB}, I_{ring}, and I_{NG}, perform in general remarkably well to reproduce the aromaticity trends along this series. The FLU index fails predicting CH_2 as the most antiaromatic of the series instead of CH^+, while HOMA values incorrectly predict that pyrrole ring is more aromatic than the cyclopentadienyl anion (CH^-). Among the three magnetic descriptors of aromaticity used in this work, $NICS(1)_{zz}$ is the only one that reproduces the expected order, whereas NICS(0) and NICS(1) assign the stronger aromatic character of the series to the pyrrole ring. Finally, the ASE calculations of Cyrański and coworkers also succeeded in giving the anticipated order [46]. According to these

Table 3 $FLU^{1/2}$, MCI, I_{NB}, I_{ring}, I_{NG}, HOMA, NICS(0), NICS(1), and $NICS(1)_{zz}$ results for the C_4H_4X ($X = CH^-$, NH, O, CH_2, BH, CH^+)

Substituents	$FLU^{1/2}$	MCI	I_{NB}	I_{ring}	I_{NG}	HOMA	ASE[a]	NICS(0)	NICS(1)	$NICS(1)_{zz}$
CH^-	0.021	0.068	0.0467	0.0475	0.0447	0.810	22.1	−12.61	−9.63	−33.68
NH	0.071	0.043	0.0426	0.0306	0.0410	0.854	20.6	−13.64	−10.12	−30.99
O	0.091	0.028	0.0392	0.0224	0.0385	0.198	14.8	−11.91	−9.42	−27.14
CH_2	0.237	0.003	0.0251	0.0079	0.0312	−0.903	0.0	−3.12	−4.95	−12.64
BH	N/A[b]	−0.003	−0.0379	0.0043	0.0416	N/A[b]	−22.5	19.99	11.29	37.30
CH^+	0.218	−0.038	−0.0625	0.0001	0.0211	−1.346	−56.7	86.04	64.91	199.70

[a]From Ref. [46] in kcal mol^{-1}

[b]There is no reference parameter for the B–C bond

results, it is recommended to use either electronic multicenter indices or $NICS(1)_{zz}$ when studying the aromaticity of heterocyclic 5-MRs [30]. It is worth noting that the Euclidean distance between neurons in a self-organizing map also gives the expected order of aromaticity for this series [47].

4.3 Substituted Pyrazoles and Imidazoles

In the following work we studied the aromaticity of a series of 1,2-azoles (pyrazoles) and 1,3-azoles (imidazoles) with N-substituents, in order to gain a better comprehension on the influence of the N-substitution on the aromaticity of azoles, a subject that has not been addressed much in detail yet, in contrast to the good understanding of the corresponding C-substitution [48].

For such purpose, the influence of different N-substituents (NH_2, Cl, NO_2, CN, OH, SH, SO_2F, SO_2CF_3, and $N(SO_2CF_3)_2$) on the aromaticity of both pyrazoles and imidazoles (see Scheme 3) was analyzed through the geometrical HOMA index, the magnetic NICS, and the electronic FLU and MCI aromaticity criteria.

First, in all cases, the imidazole appears to be more stable than the pyrazole, with ΔG ranging from 6.5 ($X = NH_2$) to 11.5 kcal mol^{-1}($X = CN$). In addition, the energy difference decreases for π-donor substituents that suffer steric hindrance with the H atom attached to the C2 atoms of the imidazole ring.

The reason for the higher stability of the imidazole ring could first be attributed to its possible higher aromaticity. However, as can be observed from Tables 4 and 5, which enclose the aromaticity criteria values for imidazole and pyrazole, respectively, both rings appear to present similar aromaticities, with the pyrazole system being somewhat more aromatic than the imidazole one according to HOMA, $NICS(1)_{zz}$, $NICS(1)_{\pi,zz}$, and $FLU^{1/2}$ (but not MCI). This behavior is equivalent to that pyrazine or pyrimidine with respect to pyridazine, the two former being more stable than the latter, even though their aromaticities are quite similar [30, 44, 49]. In such case, the lower stability of the pyridazine structure is usually related to the weakness of the N–N bond present in this molecule (vide infra). And this weakness of the N–N bond can also be assigned to the pyrazole ring [49].

It is also observed how the reduction of aromaticity in both rings due to N-substitution is larger for the substituents having larger π-acceptor character ($X = NO_2$, CN, SO_2F, and SO_2CF_3; see Scheme 4b), which subtract π-electron density from the ring and produce a clear decrease in aromaticity. This is not the case for $X = N(SO_2CF_3)_2$ for which the π-acceptor capacity is partially blocked (Scheme 4c). On the other hand, π-donor substituents ($X = NH_2$, OH, Cl, and SH; see Scheme 4a) have a minor influence in the aromaticity of the rings, as π-electron donation is disallowed in these rings because the lone pair of N cannot be delocalized in the ring.

Finally, for comparison purposes, the equivalent series of substituted benzenes was also included (see Table 6). Compared to the significant change upon substitution of imidazole and pyrazole rings, it is again proven that the aromaticity in

Scheme 3 The imidazole and pyrazole rings (the X substituent is at position 1). The substituents examined are shown according to the minimum-energy structures obtained for the two rings

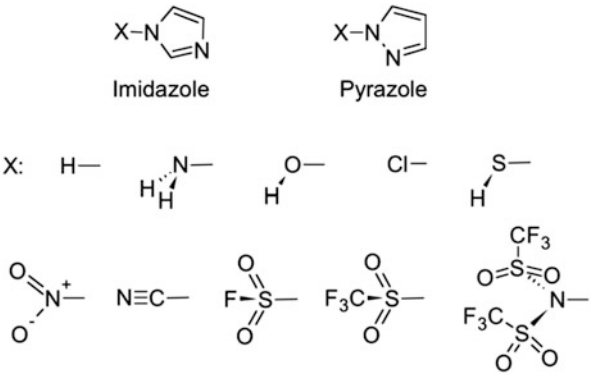


Table 4 HOMA, $NICS(1)_{zz}$ and $NICS(1)_{\pi,zz}$ (in ppm), MCI (in au), and $FLU^{1/2}$ (in au) values for substituted imidazole systems calculated at the B3LYP/6-311++G(d,p) level

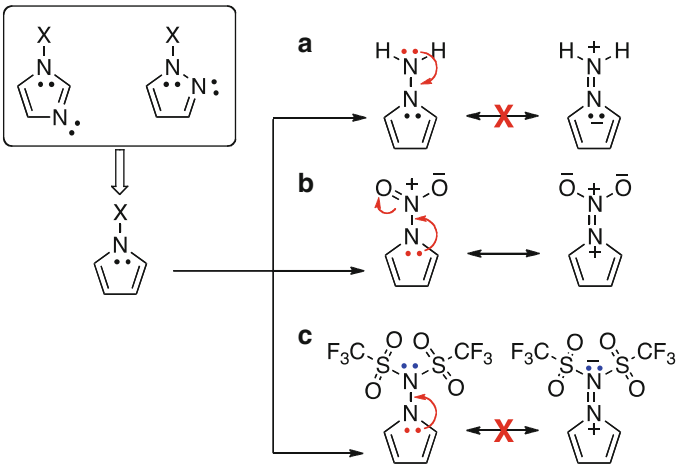
X	HOMA	$NICS(1)_{zz}$	$NICS(1)_{\pi,zz}$[a]	MCI	$FLU^{1/2}$
NH_2	0.873	−28.45	−22.76	0.049	0.115
OH	0.893	−29.03	–	0.049	0.118
Cl	0.863	−28.87	−22.74	0.048	0.123
SH	0.821	−28.72	–	0.046	0.127
H	0.880	−31.68	−23.98	0.051	0.109
NO_2	0.757	−22.24	−17.01	0.034	0.165
CN	0.729	−26.87	−19.72	0.036	0.159
SO_2F	0.709	−25.86	–	0.034	0.163
SO_2CF_3	0.718	−26.63	–	0.035	0.160
$N(SO_2CF_3)_2$	0.749	−25.97	–	0.035	0.163

[a]$NICS(1)_{\pi,zz}$ values reported only for substituted imidazoles with C_s symmetry

Table 5 HOMA, $NICS(1)_{zz}$ and $NICS(1)_{\pi,zz}$ (in ppm), MCI (in au), and $FLU^{1/2}$ (in au) values for substituted pyrazole systems

X	HOMA	$NICS(1)_{zz}$	$NICS(1)_{\pi,zz}$[a]	MCI	$FLU^{1/2}$
NH_2	0.917	−29.83	−23.82	0.045	0.109
OH	0.966	−29.64	–	0.047	0.088
Cl	0.931	−29.95	−23.84	0.044	0.104
SH	0.859	−29.77	–	0.041	0.122
H	0.913	−32.79	−24.97	0.046	0.100
NO_2	0.839	−25.24	−19.64	0.032	0.142
CN	0.756	−28.21	−20.88	0.029	0.154
SO_2F	0.781	−27.95	–	0.029	0.147
SO_2CF_3	0.793	−28.47	–	0.030	0.142
$N(SO_2CF_3)_2$	0.802	−27.65	–	0.030	0.154

[a]$NICS(1)_{\pi,zz}$ values reported only for substituted pyrazoles with C_s symmetry



Scheme 4 The mechanism of substituent delocalization in *N*-substituted five-membered rings: (*a*) π-donation partially blocked, (*b*) π-acceptation allowed, and (*c*) π-acceptation for an N-amino acid substituent partially blocked

Table 6 $NICS(1)_{zz}$ and $NICS(1)_{\pi,zz}$ (ppm), HOMA, MCI, and $FLU^{1/2}$ (au) values for benzene derivatives

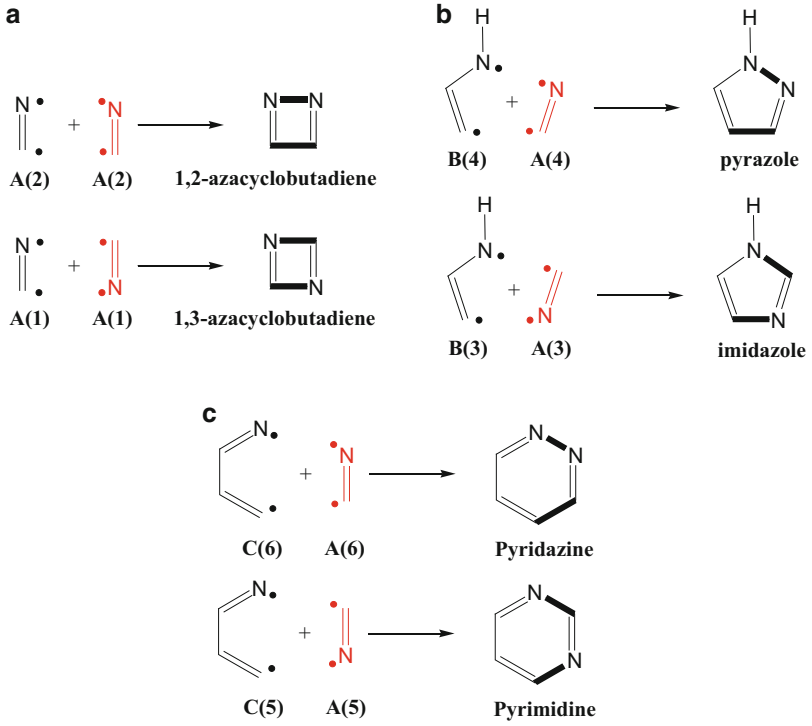
X	HOMA	$NICS(1)_{zz}$	$NICS(1)_{zz,\pi}$[a]	MCI	$FLU^{1/2}$
NH_2	0.976	−25.02	−24.87	0.060	0.040
OH	0.989	−26.76	−26.26	0.063	0.039
Cl	0.993	−27.40	−27.32	0.068	0.014
SH	0.985	−26.05	−25.95	0.065	0.017
H	0.989	−29.25	−29.07	0.072	0.000
NO_2	0.994	−27.51	−26.90	0.064	0.029
CN	0.977	−27.80	−27.44	0.063	0.031
SO_2F	0.992	−28.51	–	0.066	0.016
SO_2CF_3	0.991	−28.78	–	0.065	0.017
$N(SO_2CF_3)_2$	0.992	−27.77	–	0.062	0.032

[a]$NICS(1)_{zz\pi}$ values reported only for substituted benzene molecules with C_s symmetry

benzene is a robust property that remains almost unchanged upon ring substitution. Indeed, aromaticity changes due to substitution are much larger in imidazole and pyrazole than in benzene [50–53].

4.4 1,2 and 1,3-Diazasubstituted Four-, Five-, and Six-Membered Rings

This study is related to the previous systems [48], but now the heteroaromaticity of these systems is considered from another point of view; instead of geometric,



Scheme 5 Formation of (**a**) 1,2- and 1,3-diazacyclobutadiene, (**b**) pyrazole/imidazole, and (**c**) pyridazine/pyrimidine structural isomers from two triplet biradical fragments with the turn-upside-down approach

magnetic, and electronic aromaticity criteria, we focused on the isomerization energies. In particular, our aim was to better understand the different stability of the referred isomers pyrazole/imidazole and pyridazine/pyrimidine as well as that of 1,2- and 1,3-azacyclobutadiene isomers. Thus, first, we must remind that in many cases, both structural and stereoisomers can be constructed from the same fragments, but just connecting the two fragments in different ways. Let us consider, for instance, the 1,2- and 1,3-azacyclobutadiene isomers. They can be constructed from the same NC fragments just rotating one of the fragments (red fragment in Scheme 5a). In this case, we can analyze the bonding energy for the formation of each isomer from the same fragments, by means of a Morokuma-like energy decomposition analysis (EDA), to discuss the origin of the energy difference between the two isomers. This method is known as the turn-upside-down approach [54], and through it we get a deeper insight into the origin of isomerization energies (see Scheme 5) [55].

The ΔG isomerization energies, calculated at the BP86/TZ2P level, confirm that 1,3-diazacyclobutadiene, imidazole, and pyrimidine are more stable than their corresponding 1,2-isomers by 10.6, 9.4, and 20.4 kcal mol^{-1}, respectively

[55]. Previously, the lower stability of the latter was attributed to the lone-pair repulsion in N–N bonds in these systems [56, 57], although lone-pair protonation and deprotonation of pyridazine and pyrimidine barely change its isomerization energy [49]. Therefore, it seems that lone-pair repulsion cannot be the only cause that explains the higher stability of 1,3-isomers. The values of the turn-upside-down approach of these isomers are enclosed in Table 7.

For the particular case of 5-MR isomers (for the 4- and 6-MR isomers the conclusions are similar), imidazole/pyrazole (see Table 7), their formation is analyzed from azaethendiyl **A** and 1-azapropendiyl **B** fragments (triplet states). The fragments **A** and **B** that can be obtained from imidazole (**3**) and pyrazole (**4**) are slightly different, and thus **3** and **4** can be generated using fragments derived from **3** (**A(3)** and **B(3)**) or originated from **4** (**A(4)** and **B(4)**). The 9.4 kcal mol^{-1} higher stability of **3** can be mainly attributed to the better electrostatic interactions (9.0 kcal mol^{-1}), due to the better orientation of the dipole moments of the corresponding fragments, but especially to the σ-component of the orbital interaction term (ΔE_σ), which gives 30.3 kcal mol^{-1} stabilization due to the larger energy release in the formation of two C–N bonds as compared to the constitution of C–C and N–N bonds.

As a whole, from the rest of the values of Table 7 above, we observe how although the large Pauli repulsion explains the weaker N–N bond, it does not account for the higher stability of the 1,3-isomers. The 1,2-isomers have in fact a lower Pauli repulsion than the 1,3-forms, which are more stable, not because of a lower Pauli repulsion of the two C–N bonds as compared to the C–C and N–N bonds, but due to the more favorable electrostatic and σ-orbital interactions involved in the formation of two C–N bonds in comparison with the generation of C–C and N–N bonds in 1,2-isomers.

4.5 Aromaticity in the Guanine-Cytosine Base Pair Interacting with Metal Cations ($M = Cu^+$, Ca^{+2}, and Cu^{+2})

In this work [58], we analyzed the aromaticity in the rings of the guanine-cytosine (GC) base pair and its changes due to the interaction of GC with the cations Cu^+, Ca^{2+}, and Cu^{2+}. The analysis was carried out with the NICS, PDI, and HOMA indicators of aromaticity using the B3LYP functional with a mixed basis set of double-ζ quality with polarization functions.

The results for the GC complex **1** shown in Table 8 indicated that the 5-MR of guanine is clearly aromatic, while the 6-MRs of guanine and cytosine are only partially aromatic. To have a reference, it is worth noting that the values of benzene at the same level of theory are NICS = −9.7 ppm, PDI = 0.101 e, and HOMA = 0.981. The results obtained for the free GC base pair are in line with previous results of Cyrański et al. [59] on the same system and was also confirmed

Table 7 Analysis of the bonding (in kcal/mol) between two triplet azaethendiyl fragments (**A**) in 1,3-diazabutadiene (**1**), 1,2-diazabutadiene (**2**); azaethendiyl fragment (**A**) and triplet 1-azapropendiyl (**B**) in imidazole (**3**), pyrazole (**4**); and triplet azaethendiyl fragment (**A**) and triplet 1-azabutendiyl (**C**) in pyrimidine (**5**), pyridazine (**6**)

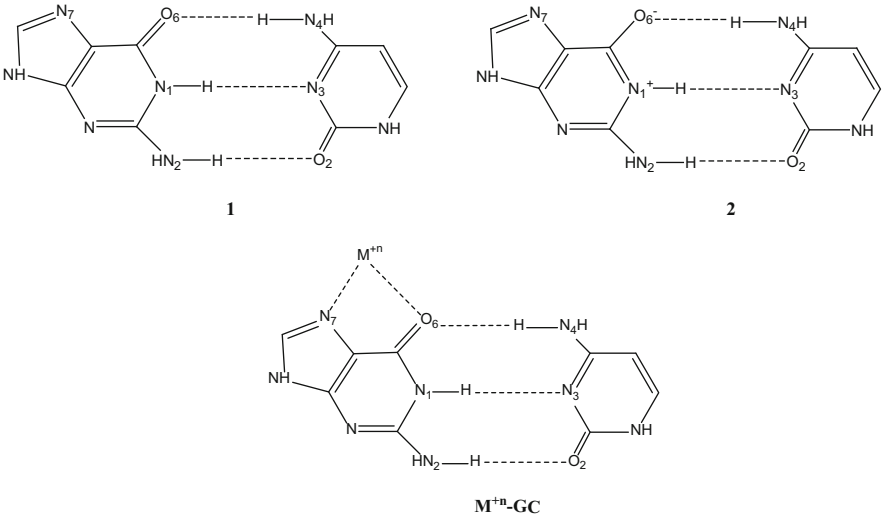
	1	2	3	4	5	6
	A(1) + A(1)	A(2) + A(2)	A(3) + B(3)	A(4) + B(4)	A(5) + C(5)	A(6) + C(6)
ΔE_{Pauli}	425.81	344.71 (−81.10)	621.48	593.31 (−28.17)	814.08	735.75 (−78.33)
ΔV_{elstat}	−220.89	−189.47 (31.41)	−338.50	−329.53 (8.97)	−405.60	−373.10 (32.50)
ΔE_{σ}	−364.37	−305.75 (58.63)	−499.06	−468.81 (30.25)	−580.98	−519.59 (61.39)
ΔE_{π}	−13.58	−10.04 (3.54)	−66.49	−67.10 (−0.61)	−79.03	−72.31 (6.72)
ΔE_{oi}	−377.95	−315.79 (62.16)	−565.55	−535.92 (29.63)	−660.01	−591.91 (68.11)
ΔE_{int}	−173.02	−160.55 (12.47)	−282.58	−272.13 (10.45)	−251.53	−229.26 (22.27)
ΔE_{def}	3.54	1.69 (−1.85)	21.72	20.63 (−1.09)	16.44	15.06 (−1.38)
ΔE	−169.48	−158.86 (10.62)	−260.8	−251.50 (9.36)	−235.09	−214.20 (20.89)

Computed at BP86/TZ2P. **A(1)** and **A(2)** refer to **A** in the geometry it adopts in **1** and **2**, respectively; **A/B(3)** and **A/B(4)** refer to **A** and **B** in the geometry they adopt in **3** and **4**, respectively; **A/C(5)** and **A/C(6)** refer to **A** and **C** in the geometry they adopt in **5** and **6**, respectively. Values in parentheses show the difference of the energy term with the corresponding one for **1** from **A(1) + A(1)**, for **3** from **A(3) + B(3),** and for **5** from **A(5) + C(5)** (see Scheme 5 for labels)

Table 8 NICS, PDI, and HOMA aromaticity measures of the five- and six-membered rings of guanine (G), and the six-membered ring of cytosine (C)

	NICS			PDI		HOMA		
System	G-5	G-6	C-6	G-6	C-6	G-5	G-6	C-6
GC	−11.94	−4.10	−1.86	0.036	0.040	0.848	0.795	0.703
$[GC]^{\cdot+}$	−5.41	−0.31	−2.49	0.023	0.042	0.829	0.550	0.773
Ca^{2+}-GC	−10.67	−4.76	−2.53	0.044	0.045	0.843	0.886	0.797
Cu^{+}-GC	−10.64	−4.59	−2.25	0.040	0.043	0.869	0.898	0.761
Cu^{2+}-GC	−7.37	−2.00	−3.07	0.022	0.040	0.915	0.760	0.822

NICS in ppm and PDI in electrons



Scheme 6 The two main resonant structures of the guanine-cytosine base pair (*top*) and schematic representation of the interaction of the GC complex with the metal cation (*bottom*)

by some of us in a subsequent study [60]. Let us now focus on the 6-MR of guanine. All indices point out an increase in the aromatic character of this 6-MR when Cu^{+} and Ca^{2+} interact with the GC pair through N_7 and O_6 atoms. This increase in aromaticity is consistent with the fact that metal cations increase the weight of the resonant structure **2** in the GC pair (see Scheme 6). In this resonant structure, the 6-MR of guanine has six π-electrons and, therefore, matches the Hückel's 4n + 2 rule [61]. This 6-MR in resonant structure **2** is more aromatic than the same ring in the resonance structure **1**. Because interaction with the metal cation increases the importance of resonant structure **2**, the guanine 6-MR of the GC pair is more aromatic when GC interacts with a metal cation. In contrast, there is a strong decrease of aromaticity of this 6-MR for the Cu^{2+}-GC structure. Interaction of GC with Cu^{2+} causes the removal of a π-electron from the guanine ring, generating $[GC]^{+}$ and reducing the aromaticity of its 6-MR [62]. As can be seen in Table 8, the

guanine 6-MR in the cation radical GC base pair has a low or nonaromatic character.

According to HOMA and NICS, the 6-MR of the cytosine base is less aromatic than the 6-MR of guanine. However, the PDI index indicates that the 6-MR of cytosine is somewhat more aromatic than that of guanine. In this situation, it is unsafe to reach definite conclusions on the relative aromatic character of these two rings [63]. What is important, however, is that the aromaticity of the cytosine 6-MR does almost not change due to the interaction of guanine with metal cations. Finally, for the 5-MR of guanine, both NICS and HOMA (PDI cannot be applied to 5-MRs) consider that this ring is more aromatic than the 6-MRs of guanine and cytosine. This 5-MR has six π-electrons and, therefore, follows the Hückel's $4n + 2$ rule. The aromaticity of this 5-MR remains also almost unaffected by the interaction of the GC base pair with the metal cations.

4.6 Porphyrins and Metalloporphyrins

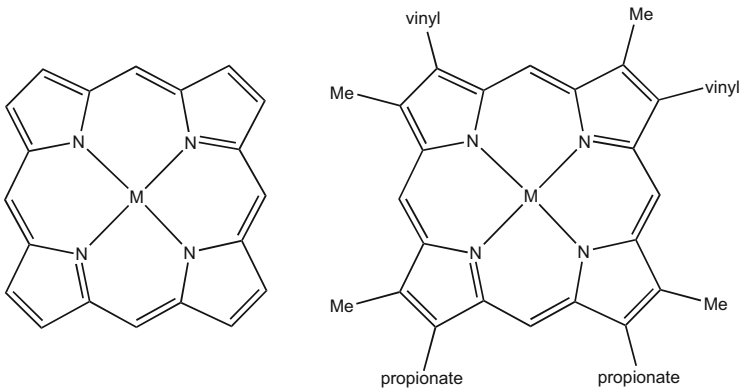
Metalloporphyrins are heterocyclic macrocycles that play an important role in biology and catalysis [64–67]. A large variety of porphyrins with different metal centers and peripheral substituents are found in nature displaying a wide range of biochemical functions. One of the most popular metalloporphyrins is formed by the heme group usually complexed to iron(II) or iron (III) metal ion which takes part in several biochemical processes such as transport, catalysis, and electron transfer reactions [68]. In this section we study the aromaticity patterns of a large group of porphyrins complexed to distinct metal ions with multiple spin states (see Table 9) [69]. As a metal center, we have considered all first-row transition metals (Sc-Zn), alkaline earth metals (Mg, Ca), and several second-row transition metals (Ru, Pd, Ag, Cd). This series of metalloporphyrins has been observed experimentally [64–67, 70]. Here, we focus on the "naked" metalloporphyrins, i.e., without peripheral substituents (see Scheme 7). As we previously said, measuring the aromatic character of big macrocycles like porphyrins or metalloporphyrins is a complex task, because some of the most popular aromaticity descriptors cannot be used [30, 71]. Geometry optimizations of these systems were carried out with the ADF suite of programs [32] using OPBE functional and basis set of triple-ζ quality containing diffuse functions and one set of polarization functions (TZP). OPBE functional has been shown to perform remarkably well for spin states [72].

First, most of the geometry optimizations led to metalloporphyrins with D_{4h} symmetry. In some cases, however, the most stable configuration was with the metal atom slightly out of the porphyrin plane (see Table 9). In particular, scandium- and calcium-porphyrins show this distortion in the ground state (doublet and singlet, respectively), while in chromium- and nickel-porphyrins, this displacement occurs in their respective triplet states, which for these two molecules are significantly less stable than their corresponding ground states (see Table 9). The most significant out-of-plane distortion is observed on the calcium system where the ion

Table 9 Metal-nitrogen (M-N) and out-of-plane distances (Å), and relative spin-state energies ($kcal \cdot mol^{-1}$) for metalloporphyrins studied

	Metal-nitrogen distance			Out-of-plane distance[a]			Spin-state energy		
Metal	Low	Interm.	High	Low	Interm.	High	Low	Interm.	High
^{2}Ag	2.115	–	–	–	–	–	–	–	–
^{1}Ca	2.274	–	–	0.857	–	–	–	–	–
^{1}Cd	2.154	–	–	–	–	–	–	–	–
^{2,4}Co	1.972	2.043	–	–	–	–	0	19.5	–
^{1,3,5}Cr	2.001	2.044	2.038	–	0.119	–	83.5[b]	40.3	0
^{2}Cu	2.021	–	–	–	–	–	–	–	–
^{1,3,5}Fe	1.975	1.983	2.055	–	–	–	30.2[b]	0	6.4
^{1}Mg	2.065	–	–	–	–	–	–	–	–
^{2,4,6}Mn	1.997	2.006	2.087	–	–	–	47.9	0.6	0
^{1,3}Ni	1.963	2.029	–	–	0.043	–	0	20.6	–
^{1,3}Pd	2.041	2.056	–	–	–	–	0	45.0	–
^{1,3,5}Ru	2.034	2.043	2.042	–	–	–	19.5[b]	0	33.5
^{2}Sc	2.107	–	–	0.270	–	–	–	–	–
^{1,3}Ti	2.051	2.064	–	–	–	–	12.4[b]	0	–
^{2,4}V	2.031	2.050	–	–	–	–	33.2[b]	0	–
^{1}Zn	2.055	–	–	–	–	–	–	–	–

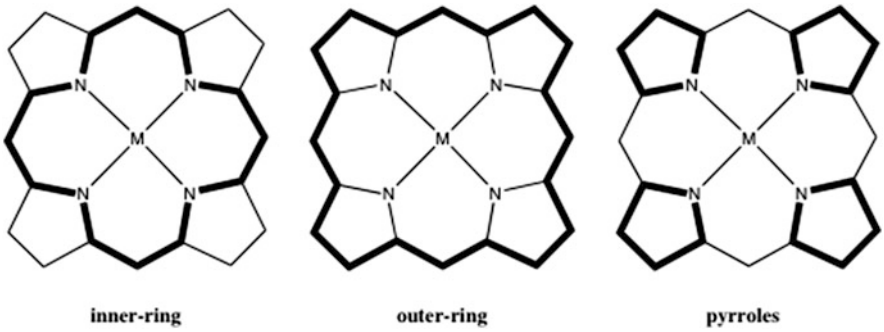
[a]Out-of-plane distance only given for those systems with C_{4v} symmetry, within D_{4h} this distance is zero because of symmetry
[b]Corrected for spin contamination



"naked" metalloporphyrin **heme-group**

Scheme 7 The naked and with the heme-group metalloporphyrins

is significantly displaced 0.86 Å from the porphyrin ring plane. In addition, the calculated spin ground states correspond to those observed experimentally, i.e., a quintet state for chromium, a triplet state for iron, a sextet for manganese, and a doublet for cobalt [68]. In most cases, the spin density was mainly localized on the



Scheme 8 Inner ring, outer ring, and pyrrole patterns

metal center; however, we found that the quintet state of ruthenium-porphyrin is effectively a coupling between Ru(III)-quartet state and porphyrin^{3-} doublet. To understand the nature of the interaction between the metal center and the heterocyclic porphyrin ring, we carried out the energy decomposition analysis of this series of metalloporphyrins. EDA showed that the interaction between the metal (M^{2+}) and the ligand (Por^{2-}) is largely electrostatic in nature. However, in all cases with the only exception of the calcium-porphyrin, the contribution of covalent interaction that arises from the orbital interaction term is crucial to stabilize the metalloporphryin. The purely electrostatic contribution obtained from the EDA analysis help to explain why the calcium atom is displaced almost 1 Å out of the porphyrin plane.

Second, we investigated the aromatic character of these complex heterocyclic compounds coordinated to a metal with different spin states. To this end, we explored the aromatic patterns of metalloporphyrins using three well-known aromaticity indices such as HOMA, FLU, and MCI. As can be seen in Scheme 8, we analyzed a number of different aromaticity patterns, i.e., the inner ring made by 16 members including the N atoms of the pyrrole groups (N16), the outer ring with 20 carbon atoms (N20), and finally the pyrrole rings (N5). HOMA and FLU have been calculated for all three N20, N16, and N5, while MCI could only be computed for N5 because of the elevated computational cost of generating up to 16! and 20! permutations of the members of the N16 and N20 rings.

To understand the aromaticity patterns of metalloporphyrins, we first started assessing the aromatic character of the free-based porphyrin ($PorH_2$) where two out of four pyrrole rings are protonated (see Table 10). In previous studies, it was shown that the aromaticity of the two protonated pyrrole rings in $PorH_2$ is higher than in the remaining two deprotonated rings [73]. This observation is attributed to the fact that when the nitrogen atom is protonated, the pyrrole ring possesses 6 π-electrons and, thus, the system is aromatic according to the $(4n+2)$ π-electron rule, while with the nitrogen deprotonated, the ring has formally 5 - π-electrons. As Table 10 shows, HOMA, FLU, and MCI predict higher aromatic character for the NH pyrrole rings of the $PorH_2$ system. These trends were also observed by aromatic ring current shielding (ARCS) analysis [74] and induced ring

Table 10 HOMA, FLU, and MCI measures for pyrrole, porphyrin($PorH_2$), Por, and Por^{-2}

	HOMA (N16)	HOMA (N20)	HOMA (N5)	FLU (N16)	FLU (N20)	FLU(N5)	MCI (N5)
Pyrrole			0.860			0.006	0.046
$PorH_2$	0.923	–	0.723(H)	0.006	–	0.017(H)	0.029(H)
			0.421			0.023	0.020
Por	0.701	0.118	0.255	0.018	0.035	0.034	0.018
Por^{-2}	0.897	0.330	0.399	0.005	0.027	0.024	0.029

current calculations [75, 76]; both studies indicated that protonated pyrrole rings present a stronger local diamagnetic ring current than pyrrole rings without the H atom. Interestingly, both FLU and HOMA values point to a large aromatic character for the inner N16 ring of PorH2 that can be related to strong electron delocalization through the N16 atoms leading to bond length equalization typical of aromatic systems. These results agree with previous studies based on NICS calculations and ring currents [73, 75, 76]. The metalloporphyrins studied in this work contain an M^{2+} center and a dianionic porphyrin ligand (Por^{2-}). Thus, deprotonation of $PorH_2$ leads to its dianion (Por^{2-}), where the charge is distributed over the whole porphyrin, and the molecule adopts a D_{4h} symmetry. HOMA and FLU values gathered on Table 10 for inner N16, and pyrrole N5 rings are similar to those observed for $PorH_2$, pointing out the aromatic nature of these cycles, while the outer N20 ring is clearly nonaromatic according to HOMA and FLU indices. Both $PorH_2$ and Por^{2-} have 18 π-electrons contributing to the inner N16 ring, consistent with $(4n + 2)$ rule for aromatic monocycles. To obtain the neutral porphyrin, we removed two electrons from Por^{2-}, leading to π-electron localization (the symmetry of the system becomes D_{2h}). Such localization of charge prompts a significant decrease of aromaticity in the inner N16 and pyrrole N5 rings. These results are in agreement with previous ring current studies of Steiner and coworkers that assigned diatropic (aromatic) ring current to $PorH_2$ and paratropic (antiaromatic) ring current to Por [76].

Next, we studied how the aromaticity of the porphyrin ring changes upon complexation of the metal cation ($M^{2+} + Por^{2-} \rightarrow MPor$). First, we focused our attention on the aromatic character of inner N16 and outer N20 circuits in metalloporphyrins compared to the free-based porphyrin. As can be seen in Table 11, the aromaticity of N16 keeps practically unchanged upon complexation of the metal center; HOMA values are close to 1 and FLU is close to 0 in all cases. Therefore, the inner N16 ring retains the aromatic character of the Por^{2-}. On the other hand, results on Table 11 show that the presence of a metal center induces a significant enhancement of aromaticity with respect to the outer N20 ring in Por^{2-}. As we explained before, the outer N20 ring of Por^{2-} is nonaromatic because the C–C bonds of the pyrrole rings are strongly localized. The presence of the metal in the center of the porphyrin triggers an increase of aromaticity in the outer ring, reducing the alternation pattern observed in the free-based porphyrin (Por^{2-}). This observation is attributed to two main factors: polarization of the π-cloud and charge transfer. First, the divalent positive charge of the metal moves the more

Table 11 HOMA, FLU, and MCI values for Por^{-2} and MPor. MCI values in a.u.

Metal	HOMA (N16)	HOMA (N20)	HOMA (N5)	FLU (N16)	FLU (N20)	FLU (N5)	MCI (N5)
Por^{-2}	0.897	0.330	0.399	0.005	0.027	0.024	0.029
^{2}Ag	0.938	0.598	0.605	0.004	0.020	0.020	0.025
^{1}Ca	0.906	0.574	0.591	0.003	0.020	0.018	0.026
^{1}Cd	0.879	0.525	0.579	0.004	0.020	0.019	0.026
^{2}Co	0.918	0.726	0.672	0.004	0.018	0.018	0.024
^{4}Co	0.944	0.683	0.644	0.004	0.018	0.016	0.025
^{1}Cr	0.892	0.798	0.711	0.005	0.016	0.017	0.023
^{3}Cr	0.906	0.803	0.750	0.006	0.021	0.027	0.024
^{5}Cr	0.934	0.710	0.663	0.004	0.018	0.018	0.025
^{2}Cu	0.951	0.695	0.656	0.004	0.019	0.018	0.026
^{1}Fe	0.909	0.745	0.680	0.004	0.017	0.018	0.024
^{3}Fe	0.912	0.752	0.687	0.005	0.016	0.015	0.025
^{5}Fe	0.933	0.696	0.656	0.004	0.018	0.019	0.024
^{1} Mg	0.934	0.658	0.637	0.003	0.019	0.017	0.026
^{2}Mn	0.911	0.734	0.666	0.005	0.016	0.020	0.022
^{4}Mn	0.904	0.782	0.706	0.006	0.016	0.020	0.022
^{6}Mn	0.926	0.641	0.630	0.004	0.019	0.018	0.026
^{1}Ni	0.917	0.724	0.677	0.005	0.018	0.019	0.024
^{3}Ni	0.948	0.689	0.650	0.004	0.019	0.018	0.026
^{1}Pd	0.954	0.692	0.656	0.005	0.019	0.019	0.025
^{3}Pd	0.930	0.794	0.770	0.007	0.016	0.019	0.025
^{1}Ru	0.911	0.754	0.684	0.004	0.017	0.018	0.024
^{3}Ru	0.923	0.734	0.676	0.004	0.017	0.015	0.024
^{5}Ru	0.871	0.883	0.783	0.009	0.010	0.009	0.027
^{2}Sc	0.867	0.775	0.717	0.006	0.013	0.011	0.027
^{1}Ti	0.869	0.797	0.699	0.007	0.014	0.022	0.018
^{3}Ti	0.890	0.749	0.677	0.005	0.015	0.017	0.023
^{2} V	0.869	0.797	0.699	0.007	0.014	0.022	0.018
^{4} V	0.890	0.749	0.677	0.005	0.015	0.017	0.023
^{1}Zn	0.944	0.671	0.642	0.004	0.019	0.018	0.026

polarizable π-electrons near to the N atom, increasing the aromaticity of the pyrrole ring, as HOMA and FLU values show. This trend was also observed by Steiner et al. for MgPor, where they reported a bifurcation of the ring current that makes both inner and outer rings aromatic [76]. Second, there are also significant two-electron two-orbital interactions between the metal cation and the porphyrin ring. These interactions are basically σ-donation from the N lone pair to the empty d-orbitals of the metal and then π-retrodonation from the filled d-orbitals of the metal atom to the π-electron system of the porphyrin. Thus, these charge transfer processes increase the π-electron populations of the 5-MR pyrrole rings, leading to an increase of the aromaticity of the outer N20 ring. In almost all the systems, the

aromatic character of the inner N16 ring is higher than the one observed for the outer N20 ring. The case of Ru^5 is particularly interesting because, as we said before, the quintet state of the ruthenium-porphyrin is obtained from the coupling of Ru(III) quartet with $porphyrin^{3-}$ doublet. In this molecule the electron delocalization patterns show a particular behavior with respect to all the other metalloporphyrins studied on this work. The outer N20 ring shows similar aromaticity as the inner N16 ring (see Table 11). In summary, the aromatic character of outer (N20) and 5-MR pyrrole (N5) rings is significantly enhanced by the presence of the metal center, while the inner ring (N16) is hardly affected. We also showed that both HOMA and FLU indices perform really well to assess the aromaticity of large heterocyclic macrocycles such as porphyrins, metalloporphyrins, and related systems. Other works, however, pointed out that the HOMA index has to be applied with care to the quantification of aromaticity in hexaphyrins since both antiaromatic and aromatic porphyrins have similar values [77, 78].

5 Aromaticity in Heterocyclic Inorganic Compounds

5.1 A Test to Evaluate the Performance of Aromaticity Descriptors in All-Metal and Semimetal Clusters

In this section we have chosen the 4-MR series of valence isoelectronic species that have a predictable trend of aromaticity ($[X_nY_{4-n}]^{q\pm}$; X,Y = Al, Ga, Si and Ge; n = 0–4) [29]. The most aromatic species are those with only one kind of atoms ($X_4^{q\pm}$ or $Y_4^{q\pm}$), while the least aromatic are the $X_2Y_2^{q\pm}$ species. For instance, in the series X = Al and Y = Ge, one can predict a steep decrease of aromaticity when going from Al_4^{2-} to heteroaromatic $GeAl_3^-$ due to the reduction of symmetry and the more electronegative character of Ge atom. The transition from $GeAl_3^-$ to Ge_2Al_2 carries a smooth reduction of aromaticity. The two isomers Ge_2Al_2 should have similar aromaticity, and when passing to Ge_3Al^+ there should be a smooth increase of aromaticity. Finally, passing from the latter species to Ge_4^{2+} shall occur with an important increase of the aromaticity. When X and Y belong to the same group, the changes are subtler because their electronegativity and geometry are more similar. In general, we expect an aromaticity trend that goes like this:

$$X_4^{q\pm} > X_3Y^{q\pm} \geq X_2Y_2^{q\pm} \leq XY_3^{q\pm} < Y_4^{q\pm}$$

Let us first discuss the series $[Al_nGe_{4-n}]^{2-n}$ (n = 0–4) in detail, and afterwards the results for the rest of the $[X_nY_{4-n}]^{q\pm}$ series are briefly analyzed. The number of valence electrons is the same for all members of the series. Figure 1 depicts the trends observed along the series Al_4^{2-} to Ge_4^{2+} for the MCI and NICS indices. Both total MCI and MCI_π curves have a clear concave ∪ shape providing the expected order of aromaticity. We have recently shown that the cross-contributions, the

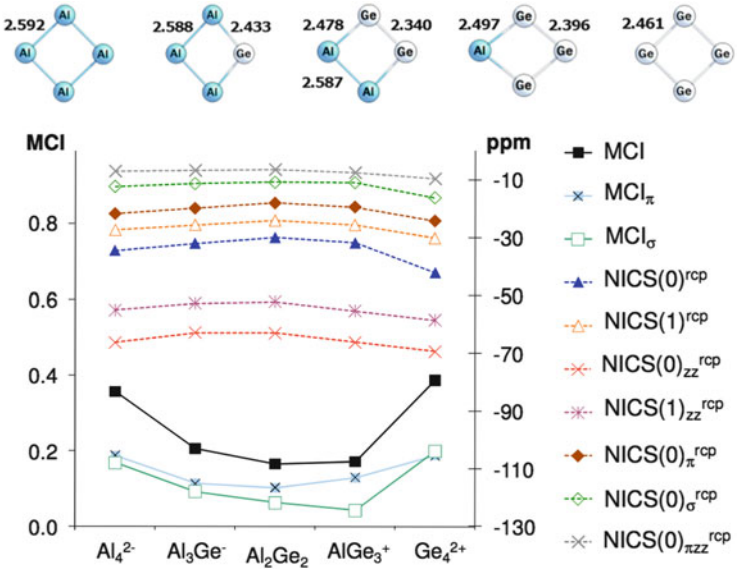
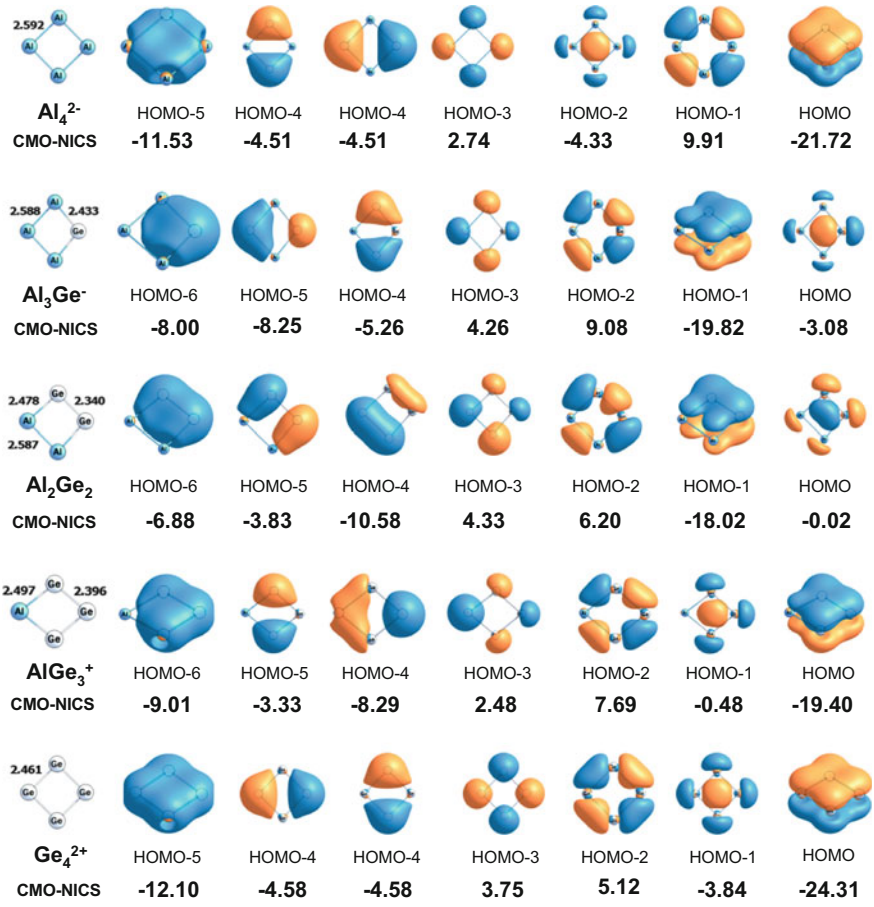


Fig. 1 Several MCI and NICS values (ppm) at the B3LYP/6-311 + G(d) level for the $[X_nY_{4-n}]^{q\pm}$ (X,Y = Al, Ga, Si, and Ge; n = 0–4) series

electron sharing indices (ESI) values of nonbonded atoms in 4-MRs ($\delta^{1,3}$ and $\delta^{1,3}{}_{\pi}$), also reproduce this finding. On the other hand, NICS(0)^{rcp} [calculated at the ring critical point (RCP)] yields the anticipated order of aromaticity, but NICS(0) [calculated at the geometrical ring center] shows an incorrect steady increase of aromaticity when going from Al_4^{2-} to Ge_4^{2+} species, thus indicating the importance of the point selected to compute the NICS value also in inorganic clusters. Actually, the distance between the geometrical ring center and the RCP is 0.19 Å, 0.28 Å, and 0.39 Å in Al_3Ge^-, Al_2Ge_2, and $AlGe_3^+$, respectively, resulting in significant differences between the NICS(0) and NICS(0)^{rcp}. In order to avoid the strong influence of the atoms nearby, it is thus recommended to calculate the NICS value at the point of lowest density in the ring plane, i.e., the RCP. The NICS out-of-plane (zz) component behaves somehow better, and it is less affected by the position where the NICS is calculated. In both cases, NICS(0)$_{zz}$ and NICS(0)$_{zz}{}^{rcp}$, the trend along the series is correct except for Al_3Ge^- that is found to be slightly less aromatic than Al_2Ge_2. In addition, the NICS values computed at 1 Å above the ring plane (NICS(1), NICS(1)^{rcp}, NICS(1)$_{zz}$, and NICS(1)$_{zz}{}^{rcp}$) show the expected ∪ behavior along the series. In order to shed more light into NICS(0) failure, we dissected NICS values for the different molecular orbitals in order to study the separate π and σ contributions (see Scheme 9). Interestingly, radial σ-orbital (HOMO-2 in Al_4^{2-}) and the π-orbital (HOMO in Al_4^{2-}) reproduce the predicted concave shape, while the tangential σ-orbital (HOMO-1) shows the wrong trend. Indeed, summing up orbital contributions, we see that NICS(0)$_{\pi}$ gives the correct



Scheme 9 Canonical molecular orbital contribution to NICS(0)^{rcp} (ppm) for the series Al_4^{2-}, Al_3Ge^-, Al_2Ge_2, $AlGe_3^+$, and Ge_4^{2+}

trend, and we can thus conclude that the σ contribution is responsible for the NICS(0) failure. It is interesting to see that NICS(0)$_\sigma$ gets smaller as more aluminum atoms are replaced by germanium atoms, indicating that the size of the atom is particularly relevant in this case.

The summary of the results obtained for the series analyzed in this work are collected in Table 12. In this table, we write "yes" for those indices that follow the expected trend for a given series, "no" otherwise, and "NC" (non-clear) when the failure of the index is minor (for instance, the index falls short only for the ordering of one species in a given series). The most complicated series for the aromaticity indices are those involving phosphor atoms. The P/S series is discussed with greater detail in the Sect. 5.3. Results in Table 12 show that the multicenter indices perform generally better than NICS, and they reinforce the superior behavior of NICS(0)$_\pi$ as compared to NICS(0), NICS(1), and their corresponding out-of-plane components.

Table 12 Summary of the results obtained at the B3LYP/6-311 + G(d) level for the six series studied with seven descriptors of aromaticity analyzed

Series	MCI	MCI^{π}	$NICS(0)^{rcp}$	$NICS(1)^{rcp}$	$NICS(0)_{zz}{}^{rcp}$	$NICS(1)_{zz}{}^{rcp}$	$NICS(0)_{\pi}{}^{rcp}$
Al/Ge	Yes	Yes	Yes	Yes	NC[a]	Yes	Yes
Al/Si	NC[a]	Yes	NC[a]	Yes	Yes	Yes	Yes
Ga/Si	Yes	Yes	NC[a]	NC[a]	Yes	Yes	Yes
Ga/Ge	Yes	Yes	Yes	Yes	Yes	Yes	Yes
P/S	Yes	Yes	No	No	No	No	Yes
P/Se	Yes	Yes	No	No	No	No	NC[a]

[a]Fails only in ordering one molecule

The series analyzed constitute a new set of tests to evaluate the performance of aromaticity descriptors in all-metal clusters, the first to the best of our knowledge.

5.2 *The Aromaticity of $M_2N_2{}^{2-}$ (M and N = B, Al, and Ga with M ≠ N) Clusters*

In this work we analyzed the molecular structure, relative stability, and aromaticity of the *cis* (C_{2v}) and *trans* (D_{2h}) heteroaromatic isomers of $M_2N_2{}^{2-}$ (M and N = B, Al, and Ga with M ≠ N) clusters [79]. We found that the C_{2v} structures with boron atoms, $Al_2B_2{}^{2-}$ and $Ga_2B_2{}^{2-}$, which present a short B–B bond, are more stable than the D_{2h} alternate systems. On the other hand, for $Ga_2Al_2{}^{2-}$ the D_{2h} isomer is marginally more stable than the C_{2v} structure (see Fig. 2). The same energy ordering is found for same rings incorporating a Li^+ counterion, i.e., the $LiM_2N_2{}^-$ (M and N = B, Al, and Ga with M ≠ N) species. To better understand the origin of the different relative stability of these isomers and the nature of the σ- and π-bonding in these systems, energy decomposition analysis (EDA) calculations were carried out. The $M_2N_2{}^{2-}$ clusters were constructed from $MN^{-\bullet\bullet\bullet}$ fragments in their quartet open-shell valence configuration (three spins up in the bonding orbitals σ_{py}, π_{px}, and π_{pz} of one of the diatomic fragments and three spins down in the same orbitals of the other) to build the σ_t, σ_r, and π molecular orbitals of $M_2N_2{}^{2-}$. Therefore, in the formation of the $M_2N_2{}^{2-}$ cluster by combination of two $MN^{-\bullet\bullet\bullet}$ fragments, two σ- and one π-bonds are formed. We analyzed the bonding situation in D_{2h} isomers obtained from D_{2h} fragments and C_{2v} isomers generated by means of fragments derived from the C_{2v} isomer. The results are collected in Table 13. $Al_2B_2{}^{2-}$ and $Ga_2B_2{}^{2-}$ have similar trends. The C_{2v} structure is the global minimum due to larger attractive orbital interactions despite the higher Pauli repulsion as compared to the D_{2h} counterpart. Both the $\Delta E_{oi\sigma}$ and $\Delta E_{oi\pi}$ terms are more stabilizing for the C_{2v} symmetry, but the main difference stems from the σ-component of the ΔE_{oi} term as a result of the strong B–B σ-bond formed in the C_{2v} symmetry. The $Al_2B_2{}^{2-}$ and $Ga_2B_2{}^{2-}$ molecules energetically prefer to keep a

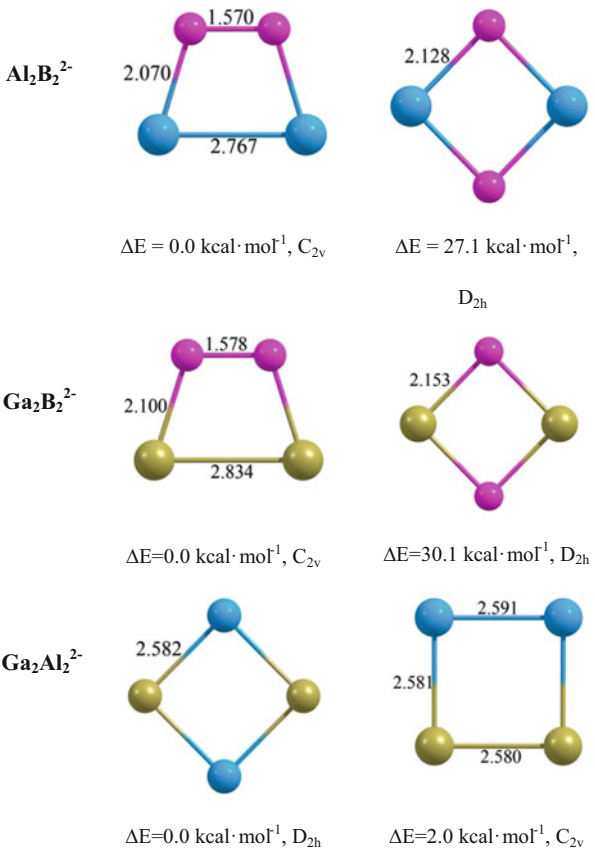


Fig. 2 Optimized structures of C_{2v} and D_{2h} $M_2N_2^{2-}$ species (where M and N = B, Al, Ga with M ≠ N). *Pink*, *blue*, and *yellow* spheres represent boron, aluminum, and gallium atoms, respectively. The relative energies for the isomers were computed at CCSD(T)/6-311G* level theory for the MP2/6-311G* optimized structures. The bond distances are in Å

Table 13 BP86/TZ2P energy decomposition analysis for the cyclic molecules

	$Al_2B_2^{2-}$		$Ga_2Al_2^{2-}$		$Ga_2B_2^{2-}$	
	C_{2v}[a]	D_{2h}[b]	C_{2v}[c]	D_{2h}[d]	C_{2v}[e]	D_{2h}[f]
ΔE_{Pauli}	294.3	228.1	134.2	117.0	296.5	260.8
ΔV_{elstat}	−98.9	−82.5	−52.4	−47.8	−93.3	−97.2
ΔE_{oi}	−244.7	−162.5	−79.8	−68.1	−244.4	−165.8
$\Delta E_{oi\sigma}$	*−210.0*	*−139.6*	*−68.0*	*−58.1*	*−209.9*	*−142.8*
$\Delta E_{oi\pi}$	*−34.7*	*−22.9*	*−11.8*	*−10.0*	*−34.5*	*−23.0*
ΔE_{int}	−49.3	−16.9	2.0	1.1	−41.2	−2.2
ΔE_{def}	0.3	0.1	0.0	0.0	0.1	0.2
ΔE	**−49.0**	**−16.8**	**2.0**	**1.1**	**−41.1**	**−2.0**

Energies are in kcal · mol^{-1}

[a] r(Al–B) = 2.064 Å, r(B–B) = 1.559 Å, r(Al–Al) = 2.871 Å

[b] r(Al–B) = 2.125 Å

[c] r(Ga–Al) = 2.612 Å, r(Al–Al) = 2.617 Å, r(Ga–Ga) = 2.625 Å

[d] r(Ga–Al) = 2.608 Å

[e] r(Ga–B) = 2.124 Å, r(B–B) = 1.561 Å, r(Ga–Ga) = 3.260 Å

[f] r(Ga–B) = 2.197 Å

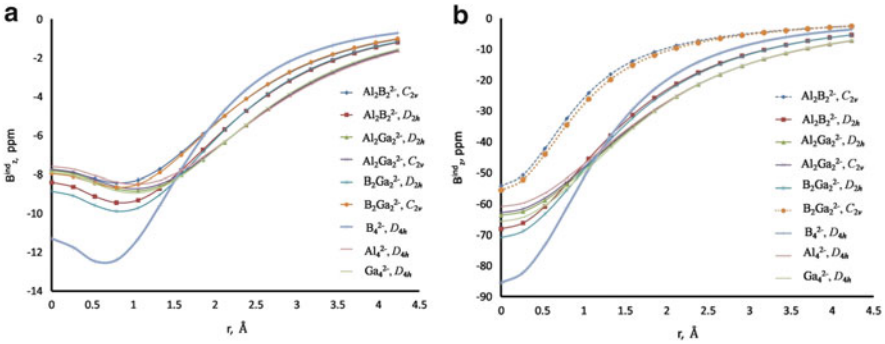


Fig. 3 The (**a**) $B^{ind}_{z\pi}$ and (**b**) $B^{ind}_{z\sigma}$ components of B^{ind}_{z} plotted along the principal axis perpendicular to the molecular plane of the C_{2v} and D_{2h} isomers of the $Ga_2Al_2^{2-}$, $Al_2B_2^{2-}$, and $Ga_2B_2^{2-}$ and of the D_{4h} B_4^{2-}, Al_4^{2-}, and Ga_4^{2-} clusters analyzed by this methodology. The external field is $|\mathbf{B}^{ext}| = 1.0$ T applied in the z (out-of-plane) direction

short B–B bond and a long M–M (M = Al, Ga) one, and this leads to a partial breaking of the M–M bond (d(Al–Al) = 2.767 Å in $Al_2B_2^{2-}$ and d(Ga–Ga) = 2.834 Å in $Ga_2B_2^{2-}$). On the other hand, for $Ga_2Al_2^{2-}$ the two isomers show almost the same values, and the difference in their respective ΔE_{int} is less than 1 kcal · mol^{-1}. The $\Delta E_{oi\pi}$ stabilization remains almost constant for the two forms, and the slightly larger stability of the D_{2h} cluster comes from smaller Pauli repulsion, which is mainly attributed to the existence of the Ga–Ga direct bond in the C_{2v} cluster (Ga has more electrons than Al).

The aromaticity of *cis* (C_{2v}) and *trans* (D_{2h}) isomers of $M_2N_2^{2-}$ (M and N = B, Al, and Ga with M ≠ N) clusters was investigated using the induced magnetic field (B^{ind}_{z}) and the MCI. $B^{ind}_{z\pi}$ is the same as $NICS_{zz\pi}$. A profile of the π contribution of the B^{ind}_{z} ($B^{ind}_{z\pi}$) that includes M_4^{2-} (M = B, Al, and Ga) clusters is plotted in Fig. 3. All $M_2N_2^{2-}$ (M and N = B, Al, and Ga) isomers present a similar trend as expected due the presence of the π-MO similar to the presence of Al_4^{2-}. Only the B_4^{2-} cluster shows a more diatropic character from the ring center up to 1.5 Å over the ring, and this is likely due to the well-known ring size dependence of NICS values. The profile of the σ contribution of the B^{ind}_{z} ($B^{ind}_{z\sigma}$) plotted in Fig. 3 shows a different behavior. In the C_{2v} isomers of $Al_2B_2^{2-}$ and $Ga_2B_2^{2-}$, the radial molecular orbital present in these molecules does not cover all the atoms because of the large Al–Al and Ga–Ga distances, and this affects the σ-aromaticity that is smaller than for the rest of the species. As can be seen in Fig. 3, the diatropic character decreases about 10 ppm along the z-axis if it is compared with their respective D_{2h} isomers.

We also calculated the MCI and its σ and π components (MCI, MCI_π, and MCI_σ) to discuss aromaticity from an electronic point of view. The results listed in Table 14 show that the two cyclical $Ga_2Al_2^{2-}$, both C_{2v} and D_{2h} isomers, are more aromatic than $Ga_2B_2^{2-}$ and $Al_2B_2^{2-}$. Basically, the two isomers of the $Ga_2Al_2^{2-}$ present almost the same total MCI values, and, therefore, the reason for the larger stability of the D_{2h} species in $Ga_2Al_2^{2-}$ cannot be attributed to higher aromaticity. The MCI_π is the same in both isomers. Interestingly, MCI results show

Table 14 The B3LYP/6-311 + G* MCI, MCI_π, and MCI_σ values (in a.u.) of the $M_2N_2^{2-}$ clusters (where M and N = B, Al, and Ga with M ≠ N)

	$Al_2B_2^{2-}$	$Al_2B_2^{2-}$	$Al_2Ga_2^{2-}$	$Al_2Ga_2^{2-}$	$Ga_2B_2^{2-}$	$Ga_2B_2^{2-}$
	(C_{2v})	(D_{2h})	(C_{2v})	(D_{2h})	(C_{2v})	(D_{2h})
MCI	0.202	0.240	0.367	0.359	0.273	0.321
MCI_π	0.129	0.155	0.187	0.185	0.149	0.187
MCI_σ	0.073	0.085	0.180	0.174	0.124	0.134

that the more stable C_{2v} isomer of $Ga_2B_2^{2-}$ and $Al_2B_2^{2-}$ is less σ- and π-aromatic than the D_{2h} form. While the reduction in σ-aromaticity when going from D_{2h} to C_{2v} is in agreement with $B^{ind}_{z\sigma}$ trends given in Fig. 3, the reduction in the π-aromaticity is less evident from the profiles of $B^{ind}_{z\pi}$. On the other hand, whereas the MCI descriptor gives similar weights to the σ- and π-aromaticities of these clusters, B^{ind}_{z} results indicate that the σ-aromaticity is much more important than the π-one in the systems analyzed.

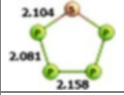
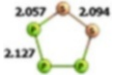
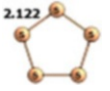
In summary, induced magnetic field and MCI results confirm the double aromatic behavior of all these species. Interestingly, for $Al_2B_2^{2-}$ and $Ga_2B_2^{2-}$, the less aromatic isomer (C_{2v}) is the global minimum in the respective potential energy surface. The reason is the short and strong B–B bond formed in these C_{2v} structures.

5.3 *The Aromaticity of $P_{5-n}S_n^{(n-1)}$ (n = 1–5) Clusters*

In this section we measure the aromaticity of a series of isoelectronic mixed phosphor and sulfur clusters [$P_{5-n}S_n^{(n-1)}$ (n = 1–5)] by different methods. Since these clusters contain unusual P–P, S–S, and P–S bonds, we cannot use any method based on reference molecules such as FLU or HOMA. We employ multicenter indices, NICS, and a model based on classical electrodynamics to measure the intensity of the ring currents (I) and characteristics of ring current loops (ring current radii, R, and height of the current loops above/below the ring plane, r) [29, 80]. These systems with six π-electrons are π-aromatic systems [81, 82]. They do not have σ-(anti)aromaticity. The values of I_π, R, and r are obtained using equations derived from classical electrodynamics and the Biot–Savart law by fitting $NICS_{\pi zz}$ values calculated at several perpendicular distances from the center of the ring located in the molecular xy plane. In the process, it is assumed that two electronic currents are flowing in infinitely narrow superconductor wires of ring current radii R and at a distance r above and below the ring plane.

From the NICS results in Table 15, only $NICS(0)_\pi$ and its zz component ($NICS(0)_{\pi zz}$) indicate the expected change in aromaticity along the $P_{5-n}S_n^{(n-1)}$ (n = 1–5) series of clusters, i.e., $P_5^- > P_4S > P_3S_2^+ \simeq P_2S_3^{2+} < PS_4^{3+} < S_5^{4+}$. The π-ring current intensity for P_5^- is the largest and decreases with the substitution of P by S^+ until the PS_4^{3+} species is reached, in a similar way as $NICS(0)_{\pi zz}$ values. Ring current loops radii are larger than the distance r of the current loops above/below

Table 15 Values of the MCI, MCI_{π}, and MCI_{σ} (in electrons), NICS (in ppm) calculated at the ring critical point (RCP), I_{π} (π-ring current intensity)[a], *R* (ring current radius in Å), *r* (height of ring current loop above and below the ring plane in Å) for the series P_5^-, P_4S, $P_3S_2^+$, $P_2S_3^{2+}$, PS_4^{3+}, and S_5^{4+} at the B3LYP/6-311 + G(d) level of theory

	P_5^-	P_4S	$P_3S_2^+$	$P_2S_3^{2+}$	PS_4^{3+}	S_5^{4+}
						
Symmetry	D_{5h}	C_{2v}	C_{2v}	C_{2v}	C_{2v}	D_{5h}
MCI	0.096	0.061	0.049	0.039	0.088	0.122
MCI_{π}	0.095	0.061	0.047	0.037	0.087	0.121
$NICS(0)^{rcp}$	−16.66	−15.59	−15.26	−14.14	−12.67	−10.25
$NICS(1)^{rcp}$	−15.83	−14.63	−14.19	−13.08	−11.77	−10.02
$NICS(0)_{zz}^{rcp}$	−31.80	−25.51	−23.34	−18.99	−14.30	−9.58
$NICS(1)_{zz}^{rcp}$	−39.24	−34.06	−31.93	−28.69	−25.03	−20.73
$NICS(0)_{\pi}^{rcp}$	−18.17	−17.64	−17.56	−17.72	−17.88	−18.20
$NICS(0)_{\sigma}^{rcp}$	2.81	3.26	3.37	4.53	5.98	8.41
$NICS(0)_{\pi,zz}^{rcp}$	−24.18	−20.85	−19.68	−19.96	−20.00	−20.86
I_{π}	32.62	27.23	24.93	24.53	24.02	24.63
R	1.984	1.975	1.942	1.897	1.882	1.873
r	0.937	0.891	0.855	0.822	0.789	0.762

[a]For calculation of I values, it is assumed that $-\mu_0/2B$ equals to 1 so I_{π} values are relative intensities and will change in different magnetic fields. However, variation in the relative intensities of currents will remain the same for different molecules in every arbitrarily chosen magnetic field

the ring plane. Interestingly, both the R and r values decrease with the increase of the total positive charge.

6 Conclusions

In this chapter we discuss the molecular structure, chemical bond, and aromaticity of a series of heteroaromatic molecules studied in our research group. The examples analyzed include both organic and inorganic species. As to the organic heterocycles, we study the series of six-membered heterocycles that result from the single to triple substitution of a CH group in benzene by an N atom, the effect of substitution in five-membered heterocycles C_4H_4X and substituted pyrazoles and imidazoles, as well as the 1,2- and 1-3-diazasubstituted four-, five-, and six-membered rings. Changes on the heteroaromaticity of organic compounds as a result of the interaction with metals are investigated in the guanine-cytosine base pair and in porphyrin rings. In the last decade aromaticity has moved well beyond the realm of purely aromatic and heteroaromatic classic molecules to encompass all-metal and semimetal aromatic compounds. The heteroaromaticity in these clusters is less known due to two main reasons: first, all-metal and semimetal heteroaromaticity is experimentally still a relatively unexplored field of research, and, second, there is a lack of reliable measures of aromaticity for all-metal and semimetal clusters. We discuss here the results for the series $[X_nY_{4-n}]^{q\pm}$; X,Y = Al, Ga, Si, and Ge (n = 0–4) and $P_{5-n}S_n^{(n-1)}$ (n = 1–5) both having an expected trend in aromaticity along the series and the heteroaromatic isomers of $M_2N_2^{2-}$ (M and N = B, Al, and Ga with $M \neq N$) clusters. Our results point out the need to use several indicators of aromaticity to get reliable orderings of aromaticity. The use of a single indicator of aromaticity is discouraged because we found many cases in which some indicators do not provide the correct aromaticity ordering for the systems under study. Our results also indicate that the electronic multicenter indices are among the most reliable indicators of aromaticity, and they are clearly recommended in studies involving heteroaromatic compounds.

Acknowledgments The following organizations are thanked for financial support: the Ministerio de Ciencia e Innovación (MICINN, project numbers CTQ2011-23156/BQU and CTQ2011-25086/BQU), the Generalitat de Catalunya (project numbers 2009SGR637 and 2014SGR931 and Xarxa de Referència en Química Teòrica i Computacional), and the FEDER fund (European Fund for Regional Development) for the grant UNGI08-4E-003. Excellent service by the Centre de Serveis Científics i Acadèmics de Catalunya (CESCA) is gratefully acknowledged. Support for the research of M. Solà was received through the ICREA Academia 2009 prize for excellence in research funded by the DIUE of the Generalitat de Catalunya. E.M. acknowledges financial support of the EU under the Marie Curie Career Integration grant (PCI09-GA-2011-294240) and the Beatriu de Pinós program from AGAUR for the postdoctoral grant (BP_B_00236). F.F acknowledges financial support from AGAUR for the Beatriu de Pinós postdoctoral Grant (BP_A_00339).

References

1. Hofmann AW (1856) Proc R Soc Lond 8:1–3
2. Kekulé A (1865) Bull Soc Chim Fr (Paris) 3:98–110
3. Anderson T (1868) Trans R Soc Edinb 25:205–216
4. Balaban AT, Oniciu DC, Katritzky AR (2004) Chem Rev 104:2777–2812
5. Stock A, Pohland E (1926) Ber. Dtsch. Chem. Ges. (A and B Series) 59:2215–2223
6. Marwitz AJV, Matus MH, Zakharov LN, Dixon DA, Liu S-Y (2009) Angew Chem Int Ed 48:973–977
7. Elliott GP, Roper WR, Waters JM (1982) J Chem Soc Chem Commun 1982:811–813
8. Li X, Kuznetsov AE, Zhang H-F, Boldyrev A, Wang L-S (2001) Science 291:859–861
9. Boldyrev AI, Wang L-S (2005) Chem Rev 105:3716–3757
10. Tsipis CA (2005) Coord Chem Rev 249:2740–2762
11. Zubarev DY, Averkiev BB, Zhai H-J, Wang L-S, Boldyrev AI (2008) Phys Chem Chem Phys 10:257–267
12. Feixas F, Matito E, Poater J, Solà M (2013) WIREs Comput Mol Sci 3:105–122
13. Li X, Zhang H-F, Wang L-S, Kuznetsov AE, Cannon NA, Boldyrev AI (2001) Angew Chem Int Ed 40:1867–1870
14. Huang X, Zhai H-J, Kiran B, Wang L-S (2005) Angew Chem Int Ed 44:7251–7254
15. Zhai H-J, Averkiev BB, Zubarev DY, Wang L-S, Boldyrev AI (2007) Angew Chem Int Ed 46:4277–4280
16. Ugrinov A, Sen A, Reber AC, Qian M, Khanna SN (2008) J Am Chem Soc 130:782–783
17. Tsipis AC, Kefalidis CE, Tsipis CA (2008) J Am Chem Soc 130:9144–9155
18. Li X, Wang L-S, Boldyrev AI, Simons J (1999) J Am Chem Soc 121:6033–6038
19. Castro AC, Audiffred M, Mercero JM, Ugalde JM, Méndez-Rojas MA, Merino G (2012) Chem Phys Lett 519–520:29–33
20. Katritzky AR, Jug K, Oniciu DC (2001) Chem Rev 101:1421–1449
21. Cyrański MK (2005) Chem Rev 105:3773–3811
22. Matito E, Duran M, Solà M (2005) J Chem Phys 122:014109
23. Fradera X, Austen MA, Bader RFW (1999) J Phys Chem A 103:304–314
24. Poater J, Fradera X, Duran M, Solà M (2003) Chem Eur J 9:400–406
25. Giambiagi M, de Giambiagi MS, dos Santos CD, de Figueiredo AP (2000) Phys Chem Chem Phys 2:3381–3392
26. Giambiagi M, de Giambiagi MS, Mundim KC (1990) Struct Chem 1:423–427
27. Bultinck P, Ponec R, Van Damme S (2005) J Phys Org Chem 18:706–718
28. Boldyrev AI, Kuznetsov AE (2002) Inorg Chem 41:532–537
29. Feixas F, Jiménez-Halla JOC, Matito E, Poater J, Solà M (2010) J Chem Theory Comput 6:1118–1130
30. Feixas F, Matito E, Poater J, Solà M (2008) J Comput Chem 29:1543–1554
31. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, S. Dapprich, Daniels AD, Farkas Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09. Gaussian 09, Revision A.02 edn. Gaussian, Pittsburgh
32. Baerends EJ, Autschbach J, Bérces A, Bickelhaupt FM, Bo C, de Boeij PL, Boerrigter PM, Cavallo L, Chong DP, Deng L, Dickson RM, Ellis DE, Fan L, Fischer TH, Fonseca Guerra C, van Gisbergen SJA, Groeneveld JA, Gritsenko OV, Grüning M, Harris FE, van den Hoek P,

Jacob CR, Jacobsen H, Jensen L, van Kessel G, Kootstra F, van Lenthe E, McCormack DA, Michalak A, Neugebauer J, Osinga VP, Patchkovskii S, Philipsen PHT, Post D, Pye CC, Ravenek W, Ros P, Schipper PRT, Schreckenbach G, Snijders JG, Solà M, Swart M, Swerhone D, te Velde G, Vernooijs P, Versluis L, Visscher L, Visser O, Wang F, Wesolowski TA, van Wezenbeek EM, Wiesenekker G, Wolff SK, Woo TK, Yakovlev AL, Ziegler T (2007) ADF2007.01. SCM, Amsterdam
33. te Velde G, Bickelhaupt FM, Baerends EJ, Fonseca Guerra C, van Gisbergen SJA, Snijders JG, Ziegler T (2001) J Comput Chem 22:931–967
34. Cheeseman JR, Trucks GW, Keith TA, Frisch MJ (1996) J Chem Phys 104:5497–5509
35. Wolinski K, Hinton JF, Pulay P (1990) J Am Chem Soc 112:8251–8260
36. Bader RFW (1990) Atoms in molecules: a quantum theory. Clarendon, Oxford
37. Salvador P, Ramos-Cordoba E (2011) APOST-3D. Institut de Química Computacional i Catàlisi. University of Girona, Girona
38. Mayer I, Salvador P (2004) Chem Phys Lett 383:368–375
39. Matito E (2006) ESI-3D: Electron Sharing Indices Program for 3D molecular space partition. Institut de Química Computacional i Catàlisi, University of Girona, Girona
40. Matito E, Solà M, Salvador P, Duran M (2007) Faraday Discuss 135:325–345
41. Morokuma K (1977) Acc Chem Res 10:294–300
42. Ziegler T, Rauk A (1977) Theor Chim Acta 46:1–10
43. Ziegler T, Rauk A (1979) Inorg Chem 18:1558–1565
44. Mandado M, Otero N, Mosquera RA (2007) Tetrahedron 62:12204–12210
45. Cioslowski J, Matito E, Solà M (2007) J Phys Chem A 111:6521–6525
46. Cyrański MK, Krygowski TM, Katritzky AR, Schleyer PR (2002) J Org Chem 67:1333–1338
47. Alonso M, Herradón B (2010) J Comput Chem 31:917–928
48. Curutchet C, Poater J, Solà M, Elguero J (2011) J Phys Chem A 115:8571–8577
49. Wang Y, Wu JI-C, Li Q, Schleyer PR (2010) Org Lett 12:4824–4827
50. Feixas F, Matito E, Poater J, Solà M (2007) J Phys Chem A 111:4513–4521
51. Krygowski TM, Ejsmont K, Stepien BT, Cyranski MK, Poater J, Solà M (2004) J Org Chem 69:6634–6640
52. Alonso M, Herradón B (2010) Phys Chem Chem Phys 12:1305–1317
53. Fernández I, Dyker CA, DeHope A, Donnadieu B, Frenking G, Bertrand G (2009) J Am Chem Soc 131:11875–11881
54. Poater J, Visser R, Solà M, Bickelhaupt FM (2007) J Org Chem 72:1134–1142
55. El-Hamdi M, Tiznado W, Poater J, Solà M (2011) J Org Chem 76:8913–8921
56. Bird CW (1996) Tetrahedron 52:9945–9952
57. Bird CW (1997) Tetrahedron 53:13111–13118
58. Poater J, Sodupe M, Bertran J, Solà M (2005) Mol Phys 103:163–173
59. Cyrański MK, Gilski M, Jaskolski M, Krygowski TM (2003) J Org Chem 68:8607–8613
60. Huertas O, Poater J, Fuentes-Cabrera M, Orozco M, Solà M, Luque FJ (2006) J Phys Chem A 110:12249–12258
61. Hückel E (1937) Z Elektrochemie 43:752–788, 827–849
62. Noguera M, Bertran J, Sodupe M (2003) J Phys Chem A 108:333–341
63. Poater J, García-Cruz I, Illas F, Solà M (2004) Phys Chem Chem Phys 6:314–318
64. Spaulding LD, Chang CC, Yu N-T, Felton RH (1975) J Am Chem Soc 97:2517–2525
65. Fuhrhop J-H, Kadish KM, Davis DG (1973) J Am Chem Soc 95:5140–5147
66. Irikura KK, Beauchamp JL (1991) J Am Chem Soc 113:2767–2768
67. Jones DH, Hinman AS, Ziegler T (1993) Inorg Chem 32:2092–2095
68. Scheidt WR, Reed CA (1981) Chem Rev 81:543–555
69. Feixas F, Swart M, Solà M (2009) Can J Chem 87:1063–1073
70. Bonomo L, Lehaire M-L, Solari E, Scopelliti R, Floriani C (2001) Angew Chem Int Ed 40:771–774
71. Feixas F, Matito E, Solà M, Poater J (2010) Phys Chem Chem Phys 12:7126–7137

72. Swart M (2008) J Chem Theory Comput 4:2057–2066
73. Cyrański MK, Krygowski TM, Wisiorowski M, Van Eikema Hommes NJR (1998) Schleyer PvR. Angew Chem Int Ed 37:177–180
74. Jusélius J, Sundholm D (2000) Phys Chem Chem Phys 2:2145–2151
75. Steiner E, Fowler P (2002) ChemPhysChem 3:114–116
76. Steiner E, Soncini A, Fowler P (2005) Org Biomol Chem 3:4053–4059
77. Fliegl H, Sundholm D, Taubert S, Pichierri F (2010) J Phys Chem A 114:7153–7161
78. Alonso M, Geerlings P, de Proft F (2012) Chem Eur J 18:10916–10928
79. Islas R, Poater J, Matito E, Solà M (2012) Phys Chem Chem Phys 14:14850–14859
80. Foroutan-Nejad C, Shahbazian S, Feixas F, Rashidi-Ranjbar P, Solà M (2011) J Comput Chem 32:2422–2431
81. De Proft F, Fowler PW, Havenith RWA, Schleyer PR, van Lier G, Geerlings P (2004) Chem Eur J 10:940–950
82. Kraus F, Korber N (2005) Chem Eur J 11:5945–5959

Chemical Bonding and Aromaticity in Poly-heterocyclic Compounds

Truong Ba Tai, Vu Thi Thu Huong, and Minh Tho Nguyen

Contents

Abstract This chapter is devoted to a theoretical analysis of the nature of chemical bonding, and in particular the aromaticity, of some intriguing poly-heterocyclic compounds which have recently been synthesized including the sulflowers and derivatives. The novel concept of *disk aromaticity* proposed as a measure of the aromatic character of planar double-ring boron clusters is applied in this context for poly-heterocyclic compounds.

Keywords Poly-heterocyclic compounds • Electron localization function (ELF) • Chemical bonding • Aromaticity • Disk aromaticity

T.B. Tai • V.T.T. Huong • M.T. Nguyen (✉)
Department of Chemistry, University of Leuven, Celestijnenlaan 200 F, 3001 Leuven, Belgium
e-mail: minh.nguyen@chem.kuleuven.be

F. De Proft, P. Geerlings (eds.), *Structure, Bonding and Reactivity of Heterocyclic Compounds*, Topics in Heterocyclic Chemistry 38, DOI 10.1007/978-3-642-45149-2_6

1 A Brief Overview

Aromaticity is one of the most popular and widely used concepts in chemistry. Its enduring influence on the chemical thinking is witnessed by, among others, the special issues of Chemical Reviews in 2001 and 2005 [1, 2] and a recent book [3]. Initially introduced to interpret the high thermodynamic stability of planar and π-conjugated monocyclic hydrocarbons [4], the concept has nowadays extensively been applied to other types of compounds such as heterocyclic [5–8] and polycyclic organic [9–11] compounds, inorganic complexes, clusters of elements [12–19], and three-dimensional and spherical structures [20–32]. However, despite its universal interest, aromaticity is rather a qualitative concept, and there are no direct measures to quantitatively evaluate it. Alternatively, aromaticity is usually associated with some characteristics of compounds such as closed-shell electronic configurations, highly symmetrical geometries, high thermodynamic stability, large resonance energy, and diatropic magnetic property [1, 2]. As a consequence, various measures have been proposed on the basis of these characteristics, such as the Hückel rule of $(4N+2)$ electrons [33], the nucleus-independent chemical shift (NICS) [34–36], ring-current theory [37, 38], resonance energies [39–42], etc. The theoretical and methodological backgrounds of these measures have been well described in the literature [1, 2], thus, we only summary in this context some outlines of the most popular approaches.

1.1 The Hückel and Möbius Aromaticity

In 1825, benzene was for the first time isolated by Faraday as a stable organic compound, despite its unsaturated electronic configuration [43]. Forty years later, the cyclohexatriene formula of benzene was proposed by Kekulé and was ever since considered as a structural basis of aromaticity [4]. In 1931, Hückel marked a great breakthrough when he proposed a rule for electron count to evaluate the aromatic character of monocyclic organic compounds [33]. Accordingly, a cyclic compound will be aromatic if it contains $(4N+2)$ π-electrons and antiaromatic if it contains $4N$ π-electrons. Probably Hückel did not expect that his proposition would become one of the most important structural indexes in chemistry, as the Hückel rule can extensively be applied for many different types of chemical compounds. It is thus extended to σ-electron systems [44, 45], and some new concepts were subsequently proposed, such as doubly (π and σ) and conflicting (π and σ) [12, 46], triple [47], and three-dimensional [48] aromaticity.

A compound whose both π- and σ-electron systems satisfy the Hückel rule of $(4N+2)$ will be doubly (π and σ) aromatic. If one of two π- and σ-electron systems of a compound contains $(4N+2)$ electrons, the other one only has $4N$ electrons; it will be conflicting aromatic. Subsequently, a compound will be doubly (π and σ) antiaromatic if both π- and σ-electron systems contain each $4N$ electrons.

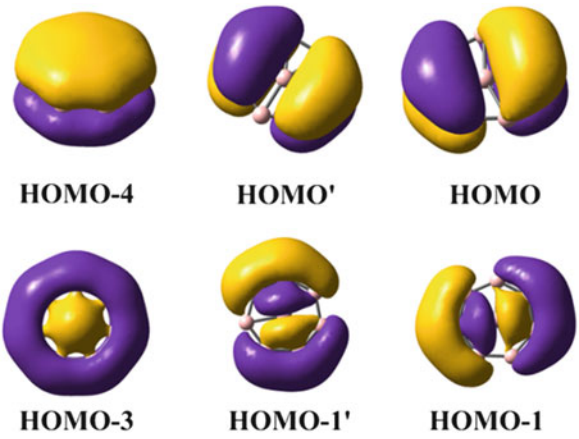


Fig. 1 Shapes of selected molecular orbitals of the boron cluster dianion B_8^{2-}

For instance, the charged boron (B) cluster B_8^{2-} is a typical doubly (π and σ) aromatic species [49] (Fig. 1).

The B_8^{2-} cluster dianion contains in total 26 valence electrons which can be divided into three different subsets. The first subset contains 16 valence σ-electrons and is responsible for 8 two-electron-two-center bonds between boron atoms. The second subset has 6 valence π-electrons (HOMO, HOMO′, and HOMO-4) and the third subset has 6 valence σ-electrons (HOMO-1, HOMO-1′, and HOMO-3) (Fig. 1). The valence electrons of both latter subsets are delocalized over the whole B_8^{2-} structure and subsequently make the B_8^{2-} doubly (π and σ) aromatic with $N = 1$. Similar observations can be found for an abundant number of atomic clusters such as B_3^-, B_4^{2-}, Al_4^{2-}, etc. in the literature [50].

On the other hand, the Möbius systems emerge as an interesting different type of aromaticity, called the Möbius aromaticity that was first proposed by Heilbronner [51]. This term was named to the memory of the German mathematician August Ferdinand Möbius (1790–1868) since he and Johann Bebedict Listing published for the first time their work on projective planes and one-sided surfaces, being now known as the Möbius strip (Fig. 2a) [52, 53]. Detailed descriptions about this concept and its applications were abundantly discussed in recent reviews [54, 55]. Generally, compounds that display a Möbius aromaticity should have rigid frameworks that enforce a smoothly twisted conjugation. A Möbius structure can be considered as a combination of a "normal" aromatic structure and a belt-shaped aromatic structure as shown in Fig. 2b [56]. Based on simple the Hückel molecular orbital theory, it was shown that the Möbius systems containing $4N$ π-electrons will be stable as they possess a closed-shell electronic configuration. This is in contrast to the Hückel systems where systems containing $4N$ π-electrons often have open-shell electronic configurations. As a consequence, the Möbius compounds are aromatic when they have $4N$ π-electrons and antiaromatic when they contain $4N + 2$π-electrons.

A typical example for a Möbius system is the cyclononatetraenyl cation $(CH)_9^+$. Scheleyer and coworkers [57] found that the cyclononatetraenyl cation possesses a

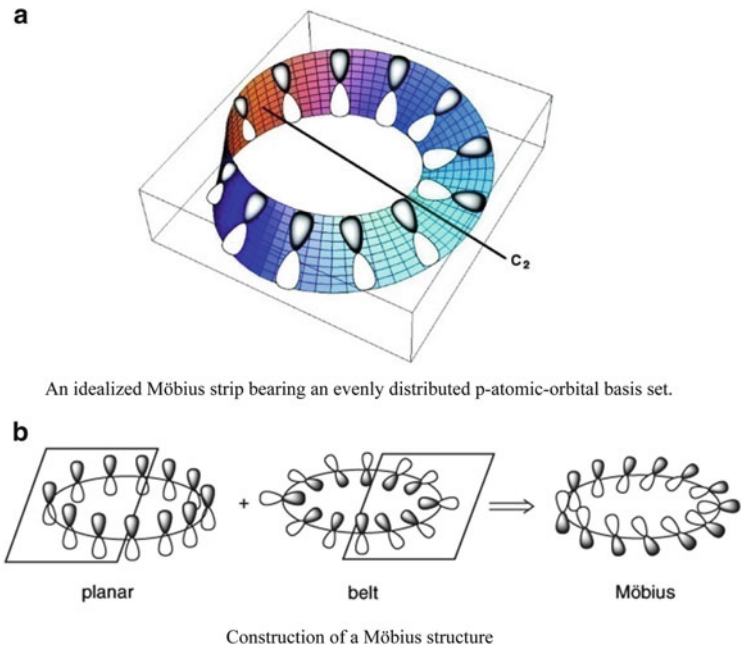


Fig. 2 Shape of an idealized Möbius strip and schematic construction of a Möbius structure (Reproduced with permission from Refs. [55] and [56])

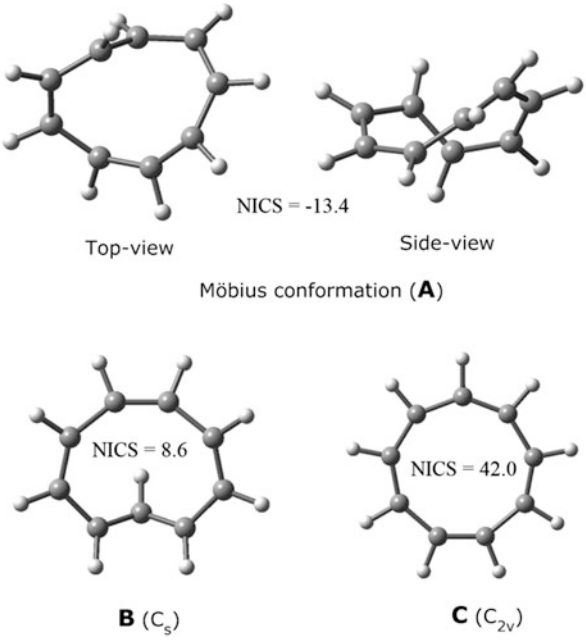


Fig. 3 Shapes of conformations of $(CH)_9^+$. NICS values were calculated at the B3LYP/6-311 + G(d) level (Ref. [57])

Möbius conformation **A** which is 22 and 26 kcal/mol more stable than the second isomer **B** and the planar isomer **C**, respectively. Interestingly, they indicated that the Möbius structure **A** exhibits an aromatic feature with the value of NICS = −13.4, whereas both conformers **B** and **C** are antiaromatic, with positive values NICS(**B**) = 8.6 and NICS(**C**) = 42.0 (Fig. 3).

1.2 Magnetic Criteria and Nucleus-Independent Chemical Shift (NICS)

Magnetic criteria-based indexes are also used to evaluate aromaticity [1, 2]. Among the indexes proposed such as the exalted magnetic susceptibilities (Λ) [58, 59] and the proton and Li^+ NMR chemical shifts [60], the nucleus-independent chemical shift (NICS) is more popularly used, owing to its friendly evaluation and efficiency. The NICS was proposed as a better alternative for indexes of proton and Li^+ NMR chemical shifts by Schleyer and coworkers in 1996 [34]. By using a virtual nucleus (called as ghost atom) placed at a center of structure, the "absolute" chemical shielding obtained avoids effects which are caused by perturbation of wave functions of H and Li nuclei and systems considered. As a consequence, NICS provides more realistic values to quantify aromaticity. However, subsequent studies found that chemical shifts of chemical compounds do not only depend on the π-electron systems of compounds, but they are also affected by other magnetic shielding contributions from the σ-electrons, lone pairs, and also core electrons. Thus, NICS value calculated at points 1 Å above the ring center (being labeled as NICS (1)) was suggested as an improved measure of π-electron delocalization [35, 36]. In order to gain more insight into distributions of σ- and π-electron systems, some dissected NICS techniques were introduced, including localized molecular orbital (LMO)-NICS [61] and canonical molecular orbital (CMO)-NICS methods [36]. In the LMO-NICS method based on the IGLO formalism, chemical shielding of molecules is only calculated on the localized π-electron systems which are separated from the σ-electron systems using the Pipek-Mezey localization procedure. The LMO-NICS is effective to evaluate π-aromatic feature of planar molecules, but its applicability to nonplanar molecules is less convenient. In other ways, the CMO-NICS based on GIAO method provides us with distribution of chemical shielding of all canonical MOs, and the σ- and π-aromatic features are subsequently evaluated. Additionally, it can conveniently be calculated, since the GIAO method is available in most quantum chemical codes such as the Gaussian, ADF, Gamess, etc. In spite of its shortcomings, the CMO-NICS technique becomes nowadays one of the most widely used (magnetic) indexes to quantitatively determine the aromatic character of different chemical compounds.

The ring-current model is another magnetic criterion which is effectively applied for evaluation of aromaticity. The first proposition of a "ring current" was initiated by London [38] and Pauling [62] in the 1930s and then further developed by Pople

[63] and McWeeny [64] in the 1950s. However, the ring-current model has only widely been used to evaluate aromaticity of compounds since Lazzeretti and Zanasi and coworkers had put considerable efforts to improve the rigorous ab initio methods to calculate molecular magnetic properties [65, 66]. These authors used a coupled Hartree-Fock method with large basis sets to study electron current density of several aromatic molecules. Following works by Lazzeretti [67], Fowler [68] and Sundholm [69], etc. gave not only more insights into this magnetic criterion but also presented visualization of the ring currents of molecules. We would refer to the comprehensive reviews by Lazzeretti [67] and Gomes and Mallion [70] that described in detail the theory and applications. In general, in an external magnetic field oriented along the main symmetry axis of the molecule, molecules will sustain induced ring currents whose strengths can be used as a measure of aromaticity. On a map of current densities, a diamagnetic circulation presented by counterclockwise arrows corresponds to an aromatic character. Oppositely, a paramagnetic circulation which is presented by clockwise arrows corresponds to an antiaromatic feature.

2 Method of Electron Localization Function (ELF)

The electron localization function (ELF) has become an effective tool for describing the nature of chemical bond and thereby aromaticity of a wide range of compounds, from novel organic compounds to elemental clusters, to inorganic and organometallic complexes [6, 11, 15, 16, 18, 71]. To get a better understanding about this intriguing function and its applications, let us briefly describe the theoretical backgrounds on which the concept was developed.

The ELF was originally introduced by Becke and Edgecome in 1990 [72] as a local function which measures the probability of finding an electron in the neighborhood of another electron with the same spin. Thus, the ELF is basically a measure of the Pauli repulsion at a given point of a molecular space. Some years later, the ELF was re-interpreted by Silvi and Savin [73] using arguments based on the excess of kinetic energy density due to Pauli repulsion. Accordingly, the ELF is defined by Eq. 1:

$$\mathrm{ELF} = \left\{ 1 + \left[\frac{t_P(r)}{t_{P,h}(\Omega(r))} \right]^2 \right\}^{-1} \tag{1}$$

where $t_P(r)$ presents the excess of local kinetic energy due to Pauli restriction and $t_{P,h}(r)$ stands for the $t_P(r)$ value for a homogeneous electron gas with the same density. The ELF values thus lie in the range of $0 \leq \mathrm{ELF} \leq 1$ in which the upper limit $\mathrm{ELF} = 1$ corresponds to the perfect localization and the $\mathrm{ELF} = 0.5$ to an electron gas-like pair probability. Theoretically, the above formula (1) only applies to closed-shell single-determinant wave functions. However, it was practically

showed that the ELF also gives meaningful results for open-shell systems in both Hartree-Fock MO and density functional theory-based approximations [74].

Silvi and Savin [73] proposed a topological analysis of the molecular space using the gradient field of ELF. On the one hand, when electrons either are lone pair or form anti-spin pair such as localized bonds, the effect of Pauli principle on their behavior and the excess local kinetic energy are very little, and the ELF value is thus close to 1. On the other hand, at the boundaries of such regions, the probability of finding electrons with the same spin is rather high. Consequently, the excess local kinetic energy is higher and the ELF value is smaller. The topological analysis is performed in order to partition the molecular space into basins Ω_A of attractors. Each basin is characterized by an attractor which is a local maximum of ELF. The synaptic order of each basin is defined as the number of core basins included. For instance, lone pairs are associated to monosynaptic basins, whereas localized two-electron-two-center bonds can be described by disynaptic basins. Integration of the electron density of a basin Ω_A will give us with its electron population. To have an efficient visualization tool, the authors [73] introduced an *f*-localization domain which is defined as formal bodies bounded by a given isosurface $\eta(r) = f$ and enclosing points for which ELF $> f$. These domains are considered to be reducible when they contain more than one contractor and irreducible when they contain only one contractor. At low ELF value, only one reducible *f*-localization domain is presented for molecular space and covers all its populations. This reducible *f*-domain will be partitioned into many smaller irreducible and reducible *f*-localization domains when the ELF value is increased.

The topological analysis of the ELF of benzene shown in Fig. 4 is a typical example. At low value of ELF = 0.5, all populations are included into only one reducible domain. At ELF = 0.65, this domain is separated into one "aromatic" domain containing populations of all carbon atoms and six disynaptic $V(C_i,H_i)$ domains. When the ELF value is increased to 0.70, the "aromatic" domain is split into six V(C,C) domains, each of which corresponds to a single C–C bond of benzene. We would refer to a previous report for a detailed analysis on the ELF of benzene and its derivatives [75].

The analysis of the separation of *f*-localization domains along the increase of ELF value gives interesting information about the location of electrons and thereby the nature of chemical bond in a molecule. Additionally, it is the basis on which quantitative aromaticity of a molecule can be evaluated. On the basis of topological analysis of ELF, Santos and coworkers [76] proposed a proper method to evaluate aromatic character. They performed an analysis over the separate ELF_π and ELF_σ. Accordingly, a molecule is π-/σ-aromatic when its bifurcation value of ELF_π/ELF_σ is high. Oppositely, it is antiaromatic when the ELF value is low.

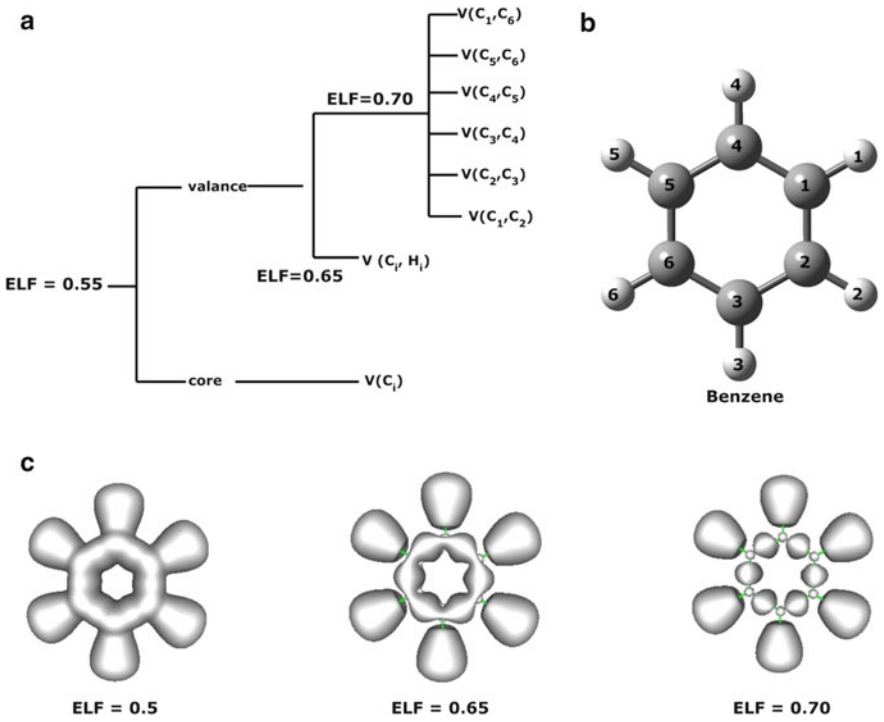


Fig. 4 Separation of *f*-localization domains (**a**), shape of molecule (**b**), and plot of total ELF (**c**) of benzene

3 Aromaticity of Thiophene and Its Derivatives

Thiophene and derivatives are important materials that form the basic units in different types of polymers and π-conjugated organic materials. Characteristics of aromaticity and chemical bonding of these cyclic molecules were reported in recent literature [36]. According to total NICS calculations, thiophene and all heteroatomic derivatives which are constructed by either substituting n CH-group by n atoms of other hetero-elements such as nitrogen (N) and phosphorus (P) or replacing S by oxygen (O) and selenium (Se) are aromatic. Total NICS values obtained at the centers of rings and at positions of 1 Å above these centers are highly negative and are comparable to the typical value of NICS(1) = −10.2 of benzene (Fig. 5).

Concerning the D_{nh} [n]-annulenes and hetero-derivatives, detailed description about their aromatic features is quite limited. It is well known that the total NICS of a molecule depends on many factors. Analysis of aromaticity based on dissected NICS techniques is thus necessary to gain more insight into distributions of σ- and π-electron subsystems. For instance, based on CMO-NICS analysis for [n]-annulenes, Heine et al. [77] showed that the largest NICS distributions come

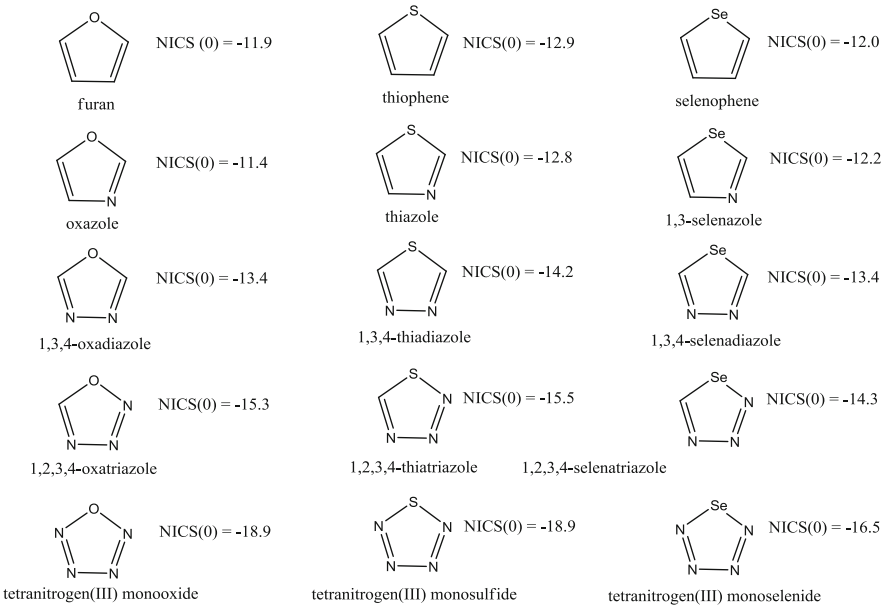


Fig. 5 Aromatic five-membered heterocyclic compounds (NICS(0) values obtained at B3LYP/6-311 + G(d,p) are taken from Ref. [36])

from the lowest-energy σ- and π-orbitals. Accordingly, the diatropic character will be decreased as the number of nodes around the ring is increased. These observations disagree with those of Fowler et al. [78], who previously found that frontier orbitals of [*n*]-annulenes are mainly responsible for the ring-current density. However, subsequent studies recognized that the distributions of frontier orbitals are dominating by the *zz*-component of the shielding tensor [79]. The latter authors concluded that the $NICS_{zz}$ values describe the aromatic character better than the total NICS. A legitimate question thus arises as to whether similar characters can be found for thiophenes.

We performed the dissected CMO-NICS analysis on thiophene and some derivatives using the GIAO method at the B3LYP/6-31 + G(d)//B3LYP/6-31 + G(d,p) level. The shapes of MOs and CMO-NICS distributions of thiophene are depicted in Fig. 6.

There is a large difference between $NICS_{zz}(0)$ and $NICS_{zz}(\pi)(0)$ values obtained at the center of thiophene (Fig. 6). The $NICS_{zz}(\pi)(0)$ is equal to −26.8 that is much more negative than the $NICS_{zz}(0)$ of −7.6. Interestingly, at position of 1 Å above the center where distribution of σ-electron system is small, this gap is remarkably decreased for $NICS_{zz}(1)$ and $NICS_{zz}(\pi)(1)$, namely, $NICS_{zz}(1) = -18.9$ and $NICS_{zz}(1)(\pi) = -26.0$. Two $NICS_{zz}(\pi)(0)$ and $NICS_{zz}(\pi)(1)$ values are close to each other (Fig. 6). These computed results clearly show that thiophene is a highly π-aromatic molecule. Additionally, Fig. 6 also reveals that the lowest-lying π-orbital (MO-19) gives larger contribution to π-aromatic character of thiophene

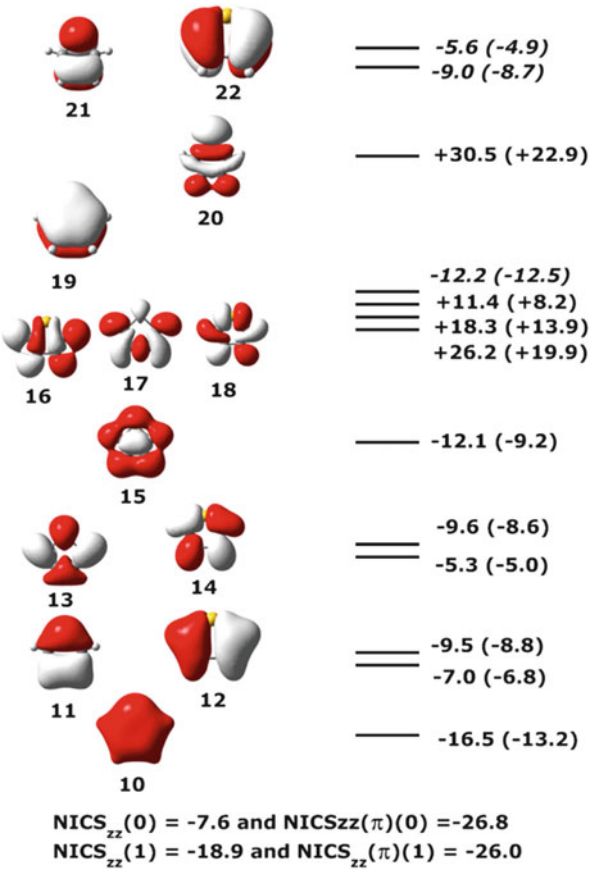


Fig. 6 Shapes of valence MOs and CMO-NICS distributions of *zz*-component of shielding tensor of thiophene in ppm. $NICS_{zz}$ values are obtained at the center of ring ($NICS_{zz}(0)$) and at position of 1 Å above the center ($NICS_{zz}(1)$, in parentheses)

than the higher-energy π-orbitals (MO-21 and MO-22). Similar observation is also found for σ-electron system that the lowest-energy σ-MOs (from MO-10 to MO-15) have diatropic shielding, whereas distributions of the higher σ-MOs are paratropic. Our predictions are similar to those of Heine et al. [77] derived for [*n*]-annulenes. However, it is worthy to note that the *zz*-component of the shielding tensor ($NICS_{zz}$) and total NICS of thiophene give the same trend.

Our dissected CMO-NICS analysis for furan and selenophene shows similar predictions. The lowest-energy π-MOs (MO-12 for furan and MO-28 for selenophene) have the largest diatropic shielding. Similar computed results are found for σ-electron systems where the lowest-lying MOs give diatropic CMO-NICS, whereas higher-energy σ-MOs induce paratropic shielding. Table 1 also reveals the large difference between $NICS_{zz}(0)$ and $NICS_{zz}(0)(\pi)$ obtained at the ring centers, whereas their $NICS_{zz}(1)$ and $NICS_{zz}(1)(\pi)$ values at 1 Å positions are close to each other. Consequently, both furan and selenophene are, as expected, strong π-aromatic species. Although earlier reports showed that furan exhibits aromaticity somewhat weaker than thiophene and selenophene [80], such a trend

Table 1 Individual CMO-NICS contributions of furan and selenophene (in ppm unit) obtained at the GIAO-B3LYP/6-31 + G(d)//B3LYP/6-31 + G(d,p) level. Bold and italic values correspond to contributions of π-MOs

Furan				Selenophene			
MO		$NICS_{zz}(0)$	$NICS_{zz}(1)$	MO		$NICS_{zz}(0)$	$NICS_{zz}(1)$
6	σ	−11.8	−5.3	19	σ	−16.8	−7.9
7	σ	−12.0	−6.4	20	σ	−9.6	−6.7
8	σ	−8.6	−5.7	21	σ	−10.6	−7.2
9	σ	−9.5	−7.1	22	σ	−4.6	−4.0
10	σ	−6.1	−4.2	23	σ	−9.6	−5.8
11	σ	−11.5	−4.9	24	σ	−10.2	−4.5
12	***π***	**−11.8**	**−9.2**	25	σ	26.6	11.5
13	σ	17.7	6.6	26	σ	20.1	9.1
14	σ	23.4	9.3	27	σ	5.2	0.4
15	σ	28.1	9.1	***28***	***π***	**−13.1**	**−11.3**
16	σ	18.6	6.3	29	σ	33.3	13.5
17	***π***	**−8.5**	**−7.2**	***30***	***π***	**−9.0**	**−7.1**
18	***π***	**−7.1**	**−4.4**	***31***	***π***	**−3.0**	**−2.4**
NICS(π)		−27.4	−20.8			−25.0	−20.8
Total NICS		−7.5	−26.9			−4.5	−26.6

is not clearly found in this context. We should note that previous reports evaluated aromaticity of five-membered heterocycles using NICS(0) and NICS(1) values. As stated above, these NICS values are less accurate to describe aromaticity of molecules due to the effects of components along the x- and y-axes of shielding tensors.

In order to gain more insights into individual CMO-NICS contributions in heterocyclic compounds, we performed a CMO-NICS analysis for some derivatives such as oxazole, thiazole, and selenazole. Table 2 showed that the lowest-energy π-MOs of all these molecules (MO-12 for oxazole, MO-18 for thiazole, and MO-27 for selenazole) give the largest diatropic CMO-NICS values, which are similar to the predictions observed above for furan, thiophene, and selenophene. Similarly, the low-lying energy σ-MOs of all these molecules also reveal diatropic characters, whereas their higher-energy σ-MOs have paratropic CMO-NICS. These molecules are accordingly aromatic systems with high negative $NICS_{zz}$ values (Table 2).

Generally, based on dissected CMO-NICS analysis, we can conclude that all the five-membered heterocycles considered are π-aromatic. The diatropic characters tend to be decreased from the lowest- to the highest-energy MOs. The same trend is holds for both π- and σ-electron systems. Such an observation is in variance with the case of [*n*]-annulenes, where their highest-energy π-MOs give larger contribution on $NICS_{zz}(\pi)$ values.

As for a better understanding about the chemical bonding and aromaticity of hetero-monocycles considered, we now performed a topological analysis of their ELF picture. The ELF_{π} plots for these molecules and their bifurcation values shown

Table 2 Individual CMO-NICS contributions in oxazole, thiazole, and selenazole (in ppm unit) obtained at the GIAO-B3LYP/6-31 + G(d)//B3LYP/6-31 + G (d,p) level. Bold and italic values correspond to contributions of π-MOs

Oxazole			Thiazole			Selenazole		
MO	$NICS_{zz}(0)$	$NICS_{zz}(1)$	MO	$NICS_{zz}(0)$	$NICS_{zz}(1)$	MO	$NICS_{zz}(0)$	$NICS_{zz}(1)$
6 (σ)	−13.2	−5.4	10 (σ)	−17.4	−7.6	19 (σ)	−20.5	−9.1
7 (σ)	−10.3	−5.3	11 (σ)	−7.5	−5.5	20 (σ)	−5.7	−4.9
8 (σ)	−9.6	−5.9	12 (σ)	−9.3	−6.0	21 (σ)	−7.2	−5.3
9 (σ)	−9.4	−7.1	13 (σ)	−6.1	−5.2	22 (σ)	−5.4	−5.1
10 (σ)	−6.2	−4.9	14 (σ)	−8.5	−5.6	23 (σ)	−8.1	−4.9
11 (σ)	−9.9	−4.3	15 (σ)	−8.2	−3.9	24 (σ)	−7.6	−3.5
12 (π)	**−12.2**	**−9.5**	16 (σ)	17.7	7.3	25 (σ)	19.4	9.2
13 (σ)	18.7	6.8	17 (σ)	14.4	5.1	26 (σ)	6.3	0.9
14 (σ)	23.5	9.2	***18 (π)***	**−13.1**	**−11.0**	***27 (π)***	**−14.5**	**−12.0**
15 (σ)	20.3	6.5	19 (σ)	20.2	8.0	28 (σ)	18.8	7.3
16 (π)	**−7.8**	**−6.0**	***20 (π)***	**−7.2**	**−5.6**	***29 (π)***	**−4.7**	**−3.9**
17 (σ)	22.6	6.4	21 (σ)	26.6	9.7	30 (σ)	29.5	11.5
18 (π)	**−6.6**	**−4.0**	***22 (π)***	**−5.8**	**−4.3**	***31 (π)***	**−5.2**	**−4.0**
NICS(π)	−26.7	−19.6	NICS(π)	−26.1	−20.8	NICS(π)	−24.4	−19.9
NICS	−9.4	−27.8	NICS	−11.5	−29.0	NICS	−8.9	−28.1

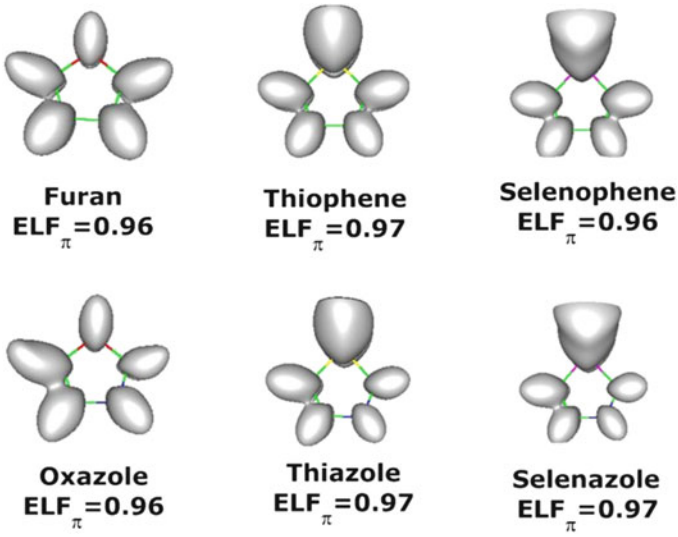


Fig. 7 The ELF_π plots and their bifurcation values for five-membered heterocycles

in Fig. 7 are in good agreement with our above discussion about their aromatic features. All molecules reveal high bifurcation values of 0.96–0.97 that are comparable to the $ELF_\pi = 0.91$ of benzene. Consequently, these molecules are aromatic.

There are similar characters about total ELF plots of the cycles considered and also their electron distributions. Figure 8 reveals monosynaptic basins of oxygen (O), sulfur (S), and selenium (Se) with electron populations of ~4.2–4.6 electrons, which correspond to their lone pairs. For furan, thiophene, and selenophene, two different types of disynaptic V(C,C) basins are equally present. The electron populations of the first type corresponding to V1(C,C) basins contain each ~3.5 electrons, which are responsible for double C=C bonds. The electron populations of the second type that corresponds to V2(C,C) basins only vary in a range of 2.1–2.5 electrons. These basins are responsible for the formation of single C–C ring bonds. In addition, each of these molecules contains two disynaptic V(C,X) basins (with X = O, S, and Se). Each of these basins is populated by ~1.8 electrons, which are responsible for single C–X bonds. Similar observations are found for heterocycles containing nitrogen. A considerable difference concerns the presence of monosynaptic V(N) basins in these cycles.

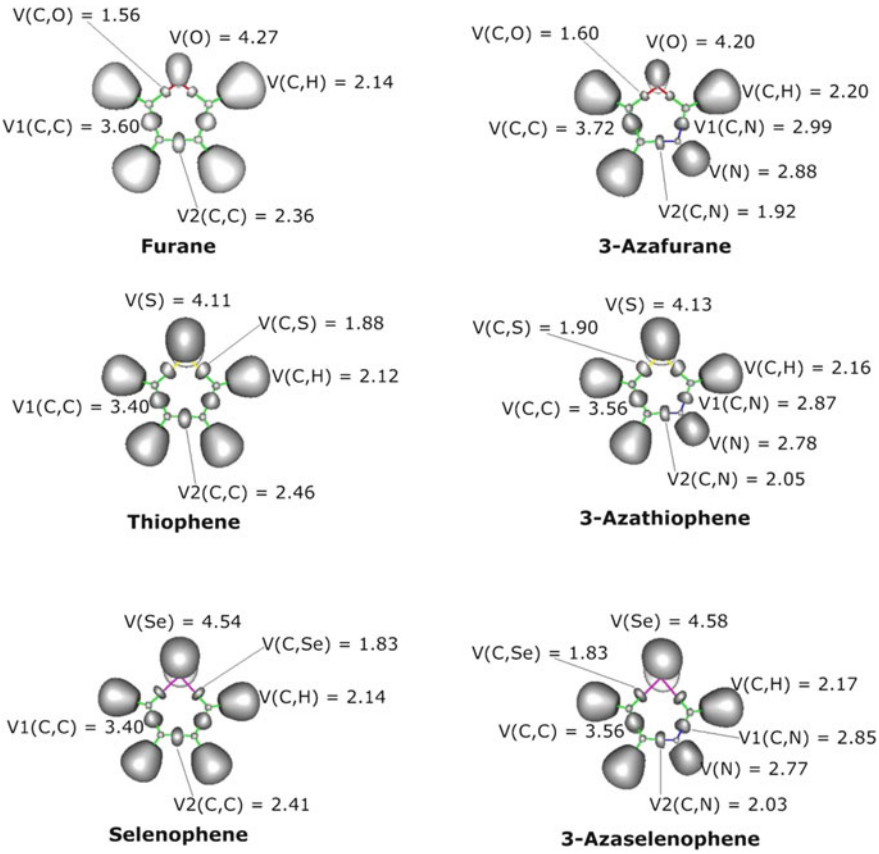


Fig. 8 Total ELF plots of molecules at bifurcation values of 0.80–0.83

4 Aromaticity of Poly-heterocycles Containing Sulfur and Their Derivatives

Oligomers and polymers containing heterocyclic units of chalcogen elements have greatly been attractive as promising units for π-conjugated organic semiconductors, which are at the heart of the fabrication of high-performance organic devices such as organic light-emitting diodes (OLEDs), organic field-effect transistors (OFETs), and organic photovoltaic cells (OPVs) [81–83]. We have shown in the previous section that possessing six delocalized π-electrons similar to that of benzene, the five-membered heterocyclic units such as furan, thiophene, selenophene, etc. exhibit highly aromatic features. Additionally, materials containing thiophene fused rings usually are more stable against oxidation reagents as compared with those containing the corresponding acenes [84]. Moreover, the compounds containing furan or thiophene are found to have high-packing crystallized structures in solid state, which is a very important property to gain effective charge transport

[85, 86]. During the past decades, numerous compounds having such features have extensively been synthesized in the laboratory, and their electronic and optical properties were carefully investigated [87–89]. In this context, we give hereafter some details about their chemical bonding.

Let us first consider the family of [*n*]-circulenes which are composed of a central *n*-membered carbon ring and a full saturation of fused benzene rings on their edges. This class emerges as a special type of polyaromatic hydrocarbons (PAHs). The [*n*]-circulenes containing 5-, 6-, and 7-membered central rings were experimentally synthesized, and their electronic structure, aromaticity, and chemical bonding features were extensively studied by theory [90–94]. Both the [5]-circulene (known as bowl-shaped corannulene) and [*6*]-circulene (known as planar coronene) are found to be hub-antiaromatic and rim-aromatic systems [92]. Interestingly, the [*7*]-circulene has a saddle-shaped structure of C_2 symmetry in the solid state and a time-averaged planar structure of D_{7h} symmetry in solution. Both structures of this molecule were found to exhibit aromatic characters similar to each other [93]. Using the ipsocentric CTOCD-DZ/6-31G(d,p)//RHF/6-31G(d,p) method, Jenneskens et al. [93] found that the structures sustain rim-diatropic and hub-paratropic currents, which are similar to those of corannulene and coronene. From the viewpoint of heterocyclic chemistry, analogous compounds of [*n*]-circulenes are of interest. In fact, many heterocyclic circulenes in which *n* fused benzene rings are replaced by *n* fused thiophene rings were experimentally synthesized [95, 96]. In what follows we analyze a few typical systems.

4.1 Aromaticity of the Sulflower $C_{16}S_8$

Ever since the *sulflower* $C_{16}S_8$ was experimentally prepared in 2006 by Chernichenko et al. [96], the annulated oligothiophenes and their derivatives have received much attention. $C_{16}S_8$ was found to be a good p-type semiconducting material for application in organic electronic devices [97]. Subsequently, its derivatives $C_{16}X_8$ with X = O, Se, PH, PF, and NH were extensively studied [98]. We should note that while [*n*]-circulenes with $n = 5$–7 were clearly established by experiment [90, 91], the [*8*]-circulene was theoretically probed to be unstable molecule with rim and hub antiaromaticity [99]. Thus, the existence of $C_{16}S_8$ has marked a great breakthrough in PHAs chemistry. $C_{16}S_8$ contains eight π-electrons which are populated at inner C_8 ring and 24 π-electrons populated at the outer hetero ring. In the view of "annulene within an annulene" model, it is the inner and outer antiaromatic molecule. Based on NICS calculations at the centers of the inner and outer rings, the $C_{16}S_8$ is however considered as a hub-nonaromatic and rim-aromatic molecule [98]. To probe the distribution of π- and σ-electron systems, we performed a CMO-NICS analysis using the GIAO-B3LYP/6-31G(d)//B3LYP/6-31 + G(d,p) method.

The computed values listed in Table 3 reveal that larger differences between the $NICS_{zz}(\pi)(0)$ and $NICS_{zz}(0)$ values occur at the ring centers. Similar to the cases of

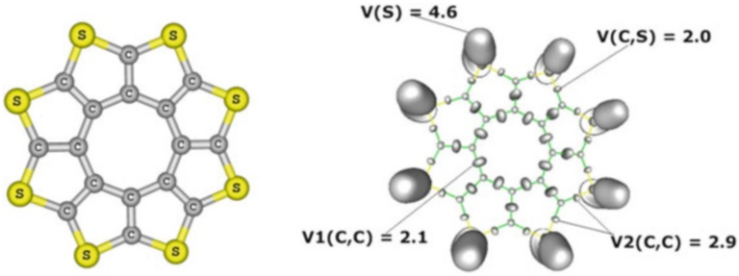


Fig. 9 Shape and total ELF plot of the sulflower $C_{16}S_8$

mono-heterocycles such as furan, thiophene, and selenophene, these gaps are considerably decreased between the $NICS_{zz}(\pi)(1)$ and $NICS_{zz}(1)$ values obtained at positions of 1 Å above centers. Additionally, all $NICS_{zz}(\pi)$ values obtained at the outer five-membered rings are consistently negative. However, while $NICS_{zz}(\pi)$ values of inner eight-membered ring are highly positive ($NICS_{zz}(\pi)(0) = +13.1$ and $NICS_{zz}(\pi)(1) = +9.9$), its $NICS(\pi)$ values are close to zero. Since $NICS_{zz}$ is considered as more realistic than total NICS, our computed results suggest that the sulflower $C_{16}S_8$ also exhibits hub-antiaromatic and rim-aromatic character. Thus, it appears that the "annulene within an annulene" model is again failed in evaluating the aromatic character of poly-heterocyclic compounds.

A more interesting information from CMO-NICS analysis is perhaps the contributions of orbitals to the global aromaticity. Table 3 reveals that NICS values tend to be decreased with increase of energy levels of π-orbitals, especially for NICS values obtained at the center of inner ring. The low-lying π-orbitals (MO-82, MO-87, MO-88, MO-94, MO-95, and MO-104) give larger diatropic distributions than the others. Importantly, the highest-energy π-orbitals (from MO-108 to MO-112) give remarkably paratropic contributions to the antiaromatic feature of the inner ring.

For a better understanding about aromaticity and chemical bonding of $C_{16}S_8$, we also performed a topological analysis of its ELF (Fig. 9). Firstly, a good electron delocalization is effective over the entire structure of $C_{16}S_8$. Eight monosynaptic basins V(S) are localized on total ELF plot where each of these basins contains 4.6 electrons. Additionally, 16 disynaptic basins V(C,S) are present, each of them contains 2.0 valence electrons which are responsible for two-electron-two-center bonds between sulfur and carbon atoms. The C–C bonds between carbon atoms of the inner eight-membered ring are presented by eight disynaptic basins V1(C,C), each being populated by 2.1 valence electrons. Finally, the ELF plot shows eight disynaptic basins V2(C,C) with each containing 2.9 electrons. These electrons together with contributions from lone pairs of sulfur atoms are responsible for the aromaticity of $C_{16}S_8$.

More interestingly perhaps, our analysis reveals that both NICS calculations and topological analysis of ELF agree with each other in the evaluation of aromaticity. The first bifurcation value of ELF_π is observed at $ELF_\pi = 0.57$ that corresponds to a

Table 3 Individual CMO-NICS contributions of valence molecular orbitals of $C_{16}S_8$. Values are calculated at the B3LYP/6-31G(d)//B3LYP/6-31 + G(d,p) level

		Inner C8 ring				Outer rings			
		NICS(0)		NICS(1)		NICS(0)		NICS(1)	
		$NICS_{zz}$	NICS	$NICS_{zz}$	NICS	$NICS_{zz}$	NICS	$NICS_{zz}$	NICS
MO-57	σ	−7.8	−3.2	−6.3	−2.3	−3.0	−2.0	−2.1	−0.6
MO-58	σ	−8.3	−3.3	−6.6	−2.4	−2.7	−1.5	−2.3	−0.8
MO-59	σ	−8.3	−3.3	−6.6	−2.4	−5.5	−3.5	−3.3	−0.8
MO-60	σ	−5.8	−2.3	−5.3	−1.9	−6.3	−3.1	−3.7	−1.1
MO-61	σ	−5.8	−2.3	−5.3	−1.9	−6.3	−3.1	−3.7	−1.1
MO-62	σ	−5.5	−2.1	−5.0	−1.8	−5.9	−2.5	−3.5	−1.2
MO-63	σ	−5.5	−2.1	−5.0	−1.8	−13.0	−5.5	−5.9	−1.6
MO-64	σ	−7.9	−3.0	−5.8	−1.6	−2.8	−1.3	−2.6	−1.0
MO-65	σ	−6.4	−2.2	−5.6	−1.9	−11.0	−4.5	−5.0	−1.4
MO-66	σ	−8.8	−2.9	−6.1	−1.9	−1.1	−0.6	−1.4	−0.6
MO-67	σ	−8.8	−2.9	−6.1	−1.9	−1.6	−0.9	−2.5	−1.0
MO-68	σ	−5.4	−1.9	−4.2	−1.5	0.4	0.1	−1.0	−0.5
MO-69	σ	−5.4	−1.9	−4.2	−1.5	0.4	0.1	−1.0	−0.5
MO-70	σ	0.8	0.2	0.4	0.1	−5.4	−1.9	−3.0	−1.1
MO-71	σ	−0.3	0.0	−0.3	−0.1	−0.2	0.0	0.2	−0.1
MO-72	σ	−0.3	0.0	−0.3	−0.1	−4.0	−1.0	−2.6	−1.0
MO-73	σ	−8.1	−3.9	−5.0	−1.7	0.8	0.1	−0.4	−0.3
MO-74	σ	−6.0	−2.2	−5.5	−2.1	0.4	−0.4	−2.4	−0.9
MO-75	σ	−6.0	−2.2	−5.5	−2.1	−2.6	−1.4	−3.9	−1.5
MO-76	σ	−0.2	0.5	−0.2	0.1	1.0	0.3	0.3	0.0
MO-77	σ	−0.2	0.5	−0.2	0.1	−1.4	−0.7	−0.7	−0.5
MO-78	σ	0.8	0.0	0.4	−0.1	−1.3	−0.7	−1.3	−0.7
MO-79	σ	0.8	0.0	0.4	−0.1	−1.3	−0.7	−1.3	−0.7
MO-80	σ	−6.0	−1.1	−5.7	−1.6	5.3	1.5	0.3	0.0
MO-81	σ	−1.4	−0.1	−1.6	−0.3	−4.4	−1.0	−3.8	−1.3
MO-82	**π**	**−8.4**	**−4.6**	**−7.3**	**−3.2**	**−2.3**	**−3.1**	**−2.2**	**−1.5**
MO-83	σ	9.0	3.5	7.6	2.7	8.4	3.0	4.8	1.6
MO-84	σ	9.0	3.5	7.6	2.7	22.6	8.3	9.8	3.3
MO-85	σ	5.0	2.9	2.9	1.4	−4.9	−1.5	−1.7	−0.6
MO-86	σ	5.0	2.9	2.9	1.4	−4.9	−1.5	−1.7	−0.6
MO-87	**π**	**−7.2**	**−3.2**	**−6.3**	**−2.5**	**−2.9**	**−1.9**	**−2.7**	**−1.2**
MO-88	**π**	**−7.2**	**−3.2**	**−6.3**	**−2.5**	**−4.0**	**−4.9**	**−3.9**	**−2.3**
MO-89	σ	8.0	2.8	4.7	1.6	−0.6	0.2	−0.3	0.0
MO-90	σ	8.0	2.8	4.7	1.6	−3.8	−0.3	−1.9	−0.7
MO-91	σ	27.1	9.5	24.5	8.5	48.1	16.8	25.7	8.6
MO-92	σ	6.2	3.2	4.6	2.0	0.3	0.1	1.4	0.4
MO-93	σ	6.2	3.2	4.6	2.0	−5.9	−1.7	−0.4	−0.2
MO-94	**π**	**−5.3**	**−2.0**	**−4.8**	**−1.7**	**−4.8**	**−3.4**	**−4.5**	**−2.0**
MO-95	**π**	**−5.3**	**−2.0**	**−4.8**	**−1.7**	**−4.8**	**−3.4**	**−4.5**	**−2.0**

(continued)

Table 3 (continued)

		Inner C8 ring				Outer rings			
		NICS(0)		NICS(1)		NICS(0)		NICS(1)	
		$NICS_{zz}$	NICS	$NICS_{zz}$	NICS	$NICS_{zz}$	NICS	$NICS_{zz}$	NICS
MO-96	σ	0.0	−0.6	0.5	−0.2	7.9	3.9	2.3	1.0
MO-97	σ	8.1	3.6	6.2	2.4	−0.1	0.0	2.1	0.6
MO-98	σ	9.3	2.7	7.5	2.2	9.3	3.7	4.9	1.7
MO-99	σ	9.3	2.7	7.5	2.2	15.3	6.8	7.5	2.7
MO-100	σ	20.0	6.4	13.9	4.4	3.0	1.9	3.2	1.1
MO-101	σ	20.0	6.4	13.9	4.4	3.0	1.9	3.2	1.1
MO-102	**π**	**−3.4**	**−1.1**	**−3.1**	**−1.0**	**−2.7**	**−1.5**	**−2.5**	**−0.9**
MO-103	**π**	**−3.4**	**−1.1**	**−3.1**	**−1.0**	**−5.3**	**−3.3**	**−5.0**	**−2.0**
MO-104	**π**	**−6.9**	**−4.4**	**−6.0**	**−3.0**	**−2.2**	**−1.0**	**−2.2**	**−0.6**
MO-105	**π**	**−4.0**	**−2.3**	**−3.3**	**−1.5**	**−1.4**	**−0.6**	**−1.2**	**−0.3**
MO-106	**π**	**−4.0**	**−2.3**	**−3.3**	**−1.5**	**1.1**	**−0.1**	**1.0**	**0.6**
MO-107	**π**	**−1.2**	**−0.4**	**−1.1**	**−0.3**	**−1.6**	**−1.2**	**−1.4**	**−0.5**
MO-108	**π**	**3.9**	**1.5**	**3.2**	**1.3**	**−1.7**	**−0.7**	**−1.4**	**−0.2**
MO-109	**π**	**18.1**	**6.0**	**15.4**	**5.2**	**5.2**	**1.6**	**5.4**	**2.1**
MO-110	**π**	**18.1**	**6.0**	**15.4**	**5.2**	**5.2**	**1.6**	**5.4**	**2.1**
MO-111	**π**	**14.8**	**5.1**	**12.6**	**4.3**	**12.0**	**4.1**	**11.1**	**4.1**
MO-112	**π**	**14.8**	**5.1**	**12.6**	**4.3**	**−5.3**	**−1.6**	**−4.6**	**−1.1**
NICS(π)		13.1	−3.0	9.9	0.4	−15.4	−19.3	−13.3	−5.7
NICS		31.7	5.2	15.9	2.6	5.9	−13.9	−14.9	−7.3

separation of basins of inner ring. This value of ELF_π is quite low and indicates that the polycycle $C_{16}S_8$ is a hub-antiaromatic system. The second bifurcation value of ELF_π, which corresponds to a separation of the basins of outer five-membered rings, amounts to 0.89. This value is very high and is comparable to the $ELF_\pi = 0.91$ of benzene. Consequently, $C_{16}S_8$ can be considered as rim-aromatic system.

4.2 Annulated Heterocycles Containing an Inner Eight-Membered C_8 Ring

Derivatives of $C_{16}S_8$ have also recently drawn much attention in theory and experiment. Previous studies showed that seleno-[8]-circulene and P-flowers $C_{16}(PH)_8$ and $C_{16}(PF)_8$ in which all S atoms of $C_{16}S_8$ are replaced by Se and –PH and –PF groups, respectively, are the most promising candidates [7, 98]. The seleno-[8]-circulene $C_{16}Se_8$ was also synthesized in experiment by Chernichenko et al. [95] using tetraselenophene as a precursor. This procedure was also used effectively to synthesize sulfur-selenium [8]-circulene $C_{16}S_4Se_4$ (Fig. 10).

Our theoretical predictions agree well with earlier studies that all above molecules are planar similarly to $C_{16}S_8$. Because N-C and O-C bond lengths are quite

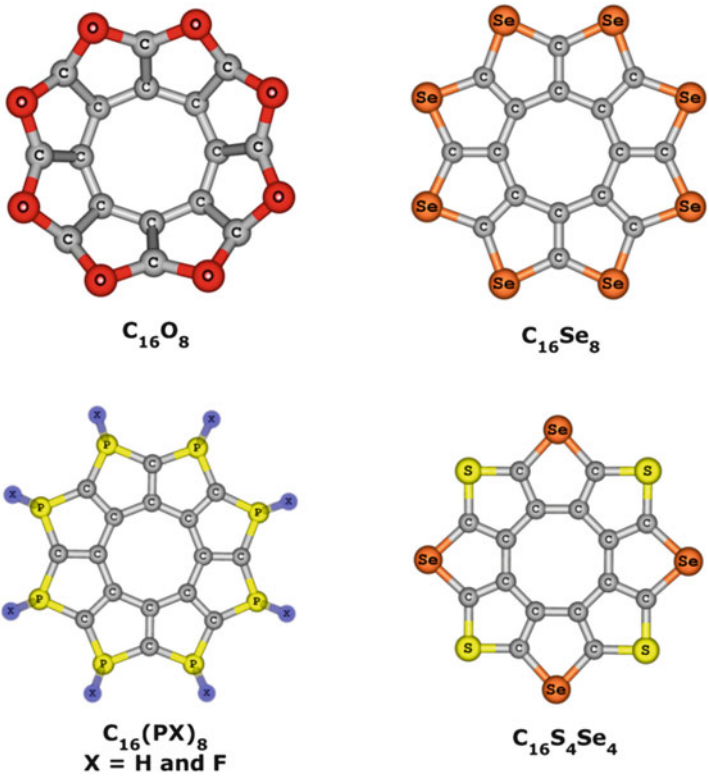


Fig. 10 Shapes of some eight-membered poly-heterocycles

shorter than C–S distance, the molecules such as $C_{16}O_8$ and $C_{16}(NH)_8$ were found to have a bowl shape rather than a planar structure. We also performed computations on the oxygen derivative $C_{16}S_4O_4$ (sulfur-oxygen [8]-circulene). However, the optimized geometry of $C_{16}S_4O_4$ is again bowl shaped.

All planar molecules $C_{16}Se_8$, $C_{16}(PH)_8$, $C_{16}(PF)_9$, and $C_{16}Se_4S_4$ turn out to have aromatic features similar to that of sulflower $C_{16}S_8$ (Table 4). They are found to be in the mean time hub-antiaromatic and rim-aromatic systems. The compounds containing sulfur and selenium have similar aromatic features. In fact, NICS values obtained at the center of the inner C_8 rings are similar, whereas NICS values obtained at the center of fused selenophene rings are somewhat less negative than those obtained at the center of fused thiophenes. Interestingly, the molecules $C_{16}(PH)_8$ and $C_{16}(PF)_8$ showed smaller magnetic characters than those of compounds containing sulfur and selenium. While NICS values obtained at the center of their C_8 rings are smaller than those of $C_{16}S_8$, the five-membered rings containing phosphorus have weaker aromatic character than the five-membered rings of sulfur and selenium.

Table 4 NICS and $NICS_{zz}$ values of poly-heterocycles obtained at the centers of rings and position of 1 Å above centers. The values are calculated at the B3LYP/6-31G(d)//B3LYP/6-31 + G(d,p) level

	Inner C8 ring				Outer rings			
	NICS(0)		NICS(1)		NICS(0)		NICS(1)	
	$NICS_{zz}$	NICS	$NICS_{zz}$	NICS	$NICS_{zz}$	NICS	$NICS_{zz}$	NICS
$C_{16}S_8$	31.7	5.2	15.9	2.6	5.9	−13.9	−14.9	−7.3
$C1_6Se_8$	30.9	5.2	15.7	2.5	14.2	−10.7	−11.3	−6.8
$C_{16}S_4Se_4$	31.2	5.1	15.8	2.5	8.8[a]	−11.4[a]	−14.0[a]	−7.0[a]
					14.0[b]	−10.8[b]	−14.6[b]	−6.8[b]
$C_{16}(PH)_8$	24.7	4.6	11.6	1.8	11.8	−8.6	−6.9	−4.2
$C_{16}(PF)_8$	17.6	1.9	4.6	0.1	12.6	−7.6	−3.9	−1.5

[a]NICS values obtained at center of outer thiophene rings
[b]NICS values obtained at center of outer selenophene rings

5 Disk Aromaticity of Poly-heterocycles

As discussed above, the classical Hückel rule of $(4N + 2)$ electrons is beyond any doubt the most applied model to probe aromaticity. Let us state again that a planar system containing $(4N + 2)$ delocalized π-valence electrons exhibits aromaticity, whereas a planar system that contains $4N$ delocalized π-valence electrons has an antiaromatic feature [33]. This rule has consistently been proven to be effective in examining the aromaticity of planar monocycles, including organic compounds, inorganic complexes, and metal clusters. However, this rule seems to be failed when applied to polycyclic compounds. For instance, Wang et al. [100] reported a combined theoretical and experimental investigation on the anionic boron cluster B_{19}^-. Their findings showed that B_{19}^- exhibits a beautiful planar structure that is composed of two boron rings with one central B atom. Although this anionic cluster has an aromatic feature on the basis of negative $NICS_{zz}$ value, its molecular orbital (MO) pattern does not follow the Hückel rule of $(4N + 2)$ at all. The anion contains in fact 12 delocalized π-electrons. In order to rationalize this feature, the authors [100] proposed that the anion has a concentric doubly π-aromaticity corresponding to two concentric delocalized π-electron systems, with six electrons each. In a recent communication, we found that the anionic B_{20}^- and B_{20}^{2-} clusters also exhibit polycyclic planar structures, being composed of 1 outer 13-membered ring, 1 inner 7-membered ring, and one central B atom [14]. The B_{20}^{2-} dianion is also found to be an aromatic system whose MO pattern is similar to that of the B_{19}^- anion. More fundamentally, we pointed out that the energy levels and shapes of MOs of B_{20}^{2-}, and also of B_{19}^-, match well with those obtained from the simple model of a particle in a circular box. The solutions of the Schrödinger equation for the model are constructed using the Bessel functions as basic wave functions. These anions are proposed to have a *disk aromaticity*, and concretely, *each disk-aromatic system bears 12 valence π-electrons*. A legitimate question arises here as to whether this represents only a special case for systems containing 12 delocalized valence

π-electrons or it is associated with a more general feature of polycyclic compounds. Let us first give a short introduction to the disk aromaticity model.

5.1 Theoretical Backgrounds

The "particle in a box" is the simplest example for the solution of the Schrödinger equation and widely used in textbooks of elementary quantum chemistry to show the ways of constructing wave functions, thereby of qualitatively explaining many basic phenomena related to electronic structure. The "particle in a circular box" is another simple but more basic quantum chemical model. This model can be considered as a two-dimensional problem in which the particle is moving in a potential well with the perimeters having infinite height. The particle is restricted to be within $r = \mathrm{R}$, where R is a constant. In polar coordinates, the Schrödinger equation for this problem is written as follows:

$$\frac{-\hbar^2}{2\mu}\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}\right)\psi(\theta,r)=E\psi(\theta,r) \tag{2}$$

where $\hbar$ is the Plank constant and μ is the mass of particle.

Because this is a circular symmetry and polar coordinate problem, the wave function ψ (θ, r) can be written as $R(r).\Theta(\theta)$. The solution for this equation will be $\psi = R(r)\ \mathrm{e}^{i\mu\theta}$. Substitution of the latter into the Schrödinger equation gives us

$$\frac{\partial^2 R(r)}{\partial r^2}+\frac{1}{r}\frac{\partial R(r)}{\partial r}+\frac{1}{r^2}\left(k^2-\frac{m^2}{r^2}\right)R(r)=0 \quad \text{with} \quad k^2=\frac{E2\mu}{\hbar^2} \tag{3}$$

This equation is known as the Bessel differential equation [101], and its solutions are the Bessel functions $J_l(x)$.

The Bessel functions $J_l(x)$ for integer l have the form

$$J_l(x)=\left(\frac{x}{2}\right)^l\sum\nolimits_{m=0}^{\infty}\frac{(-1)^m}{m!(l+m)!}\left(\frac{x}{2}\right)^{2m} \tag{4}$$

This is the Bessel function of the first kind of order l. This function is finite and converges for all values of x. The values of x for which $J_l(x) = 0$ is called the *zeros* of the functions.

The energy of a particle in this circular box can thus be obtained as

Table 5 Radial (n) and rotational (l) quantum number, energies (E, in unit of $\hbar^2/2ma^2$) of the lowest eigenstates, and corresponding number of π-electrons

n	l	E	Number of π-electrons
1	0	5.783	**2**
1	±1	14.682	**6**
1	±2	26.374	**10**
2	0	30.472	**12**
1	±3	40.707	**16**
2	±1	49.219	**20**
1	±4	57.582	**24**
2	±2	70.849	**28**
3	0	74.887	**30**
2	±3	95.277	**34**
3	±2	103.502	**38**

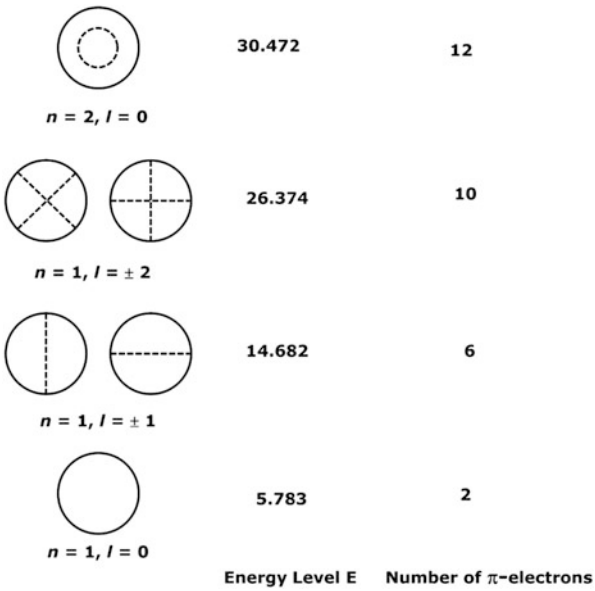


Fig. 11 Maps of node characteristics of the wave functions for the lowest eigenstates of a circular box

$$E_{n,l} = \alpha_{n,l}^2 \frac{\hbar^2}{2m} = x_{n,l}^2 \frac{\hbar^2}{2ma^2} \quad n = 1, 2, 3 \ldots \text{and } l = 0, \pm 1, \pm 2 \ldots \qquad (5)$$

where $x_{n,l}$ is the *zeros* of the Bessel functions $J_l(x)$, a is the radius of box, and n and l are the radial and rotational quantum numbers, respectively.

By replacing the zeros of the Bessel functions into Eq. 5, the energies (in unit of $\hbar^2/2ma^2$) of the lowest-lying states of a particle can be determined and are given in Table 5. The maps of the signs and nodes of the wave functions determined from the use of the Bessel functions for some lowest-lying states are sketched and represented in Fig. 11.

It turns out that this simple model of a particle in a circular box offers an amazingly accurate description of the π-states in the cyclic boron structures. The eigenstates in the box model are characterized by a radial ($n = 1, 2, \ldots$) and a rotational ($l = \sigma, \pi, \delta, \phi, \gamma, \ldots$ or $l = 0, \pm 1, \pm 2, \pm 3, \pm 4, \ldots$) quantum number, and the energies are determined by the zeros of the cylindrical Bessel functions. The lowest eigenstates in ascending order are 1σ, 1π, 1δ, and 2σ. The nodal characteristics of these cylindrical waves exactly match the sequence of the occupied delocalized π-orbitals of boron-based clusters. Accordingly, the molecules containing the number of 2, 6, 10, 12, 16... delocalized valence π-electrons that fully occupy the circular shells of the model exhibit an enhanced stability that we proposed as to arise from a *disk aromaticity*. Consequently, the molecules containing the numbers of 4, 8, 14, 18... delocalized valence π-electrons exhibit a *disk-antiaromatic* character.

5.2 Some Applications

As for a crucial test, we re-examine the cases of boron-carbon wheels $C_5B_{11}^+$ and $C_6B_{12}^{2-}$ [102, 103] whose structures and shapes have well been established. They were in fact reported as stable structures with interesting fluxional features and beautiful symmetrical geometries.

While the cation $C_5B_{11}^+$ was found to be aromatic with ten delocalized valence π-electrons, the dianion $C_6B_{12}^{2-}$ is antiaromatic despite the fact that $C_6B_{12}^{2-}$ contains 14 delocalized valence π-electrons [103]. From this observation, the authors [98] claimed that the Hückel rule of $(4N + 2)$ is inefficient for polycyclic wheels, but a further explanation was not given.

Let us reconsider these structures on the basis of the model of a particle in a circular box. For the purpose of comparison, the species $C_5B_{11}^-$ and C_6B_{12} are also examined. The shapes of the valence π-MOs of these structures are given in Fig. 12 together with the node characteristics of the model. It is interesting to see that both charged species $C_5B_{11}^+$ and $C_5B_{11}^-$ are disk-aromatic systems that satisfy the model of a particle in a circular box. Their pictures of π-MOs agree well with the node characteristics of the model (see above). The cation $C_5B_{11}^+$ contains ten delocalized π-MOs that doubly occupy the eigenstates of (1σ1π1δ). Adding two excess electrons to π-HOMO of the cation $C_5B_{11}^+$, the anion $C_5B_{11}^-$ now has 12 delocalized π-electrons that occupy the eigenstates of (1σ1π1δ2σ). Consequently, they both exhibit disk-aromatic features.

More interestingly, while the neutral C_6B_{12} is aromatic, its dianion $C_6B_{12}^{2-}$ is antiaromatic. It can be seen in Fig. 10 that the dianion $C_6B_{12}^{2-}$ only possesses 14 valence π-electrons that thereby cannot fulfill the eigenstates of the model ($n = 1$ and $l = \pm 3$). Consequently, it exhibits an anti-disk-aromatic character. Following detachment of two electrons from HOMO of the $C_6B_{12}^{2-}$, the picture of MOs of C_6B_{12} is similar to that of the $B_5C_{11}^-$ and B_{20}^{2-} and the whole feature makes it disk aromatic.

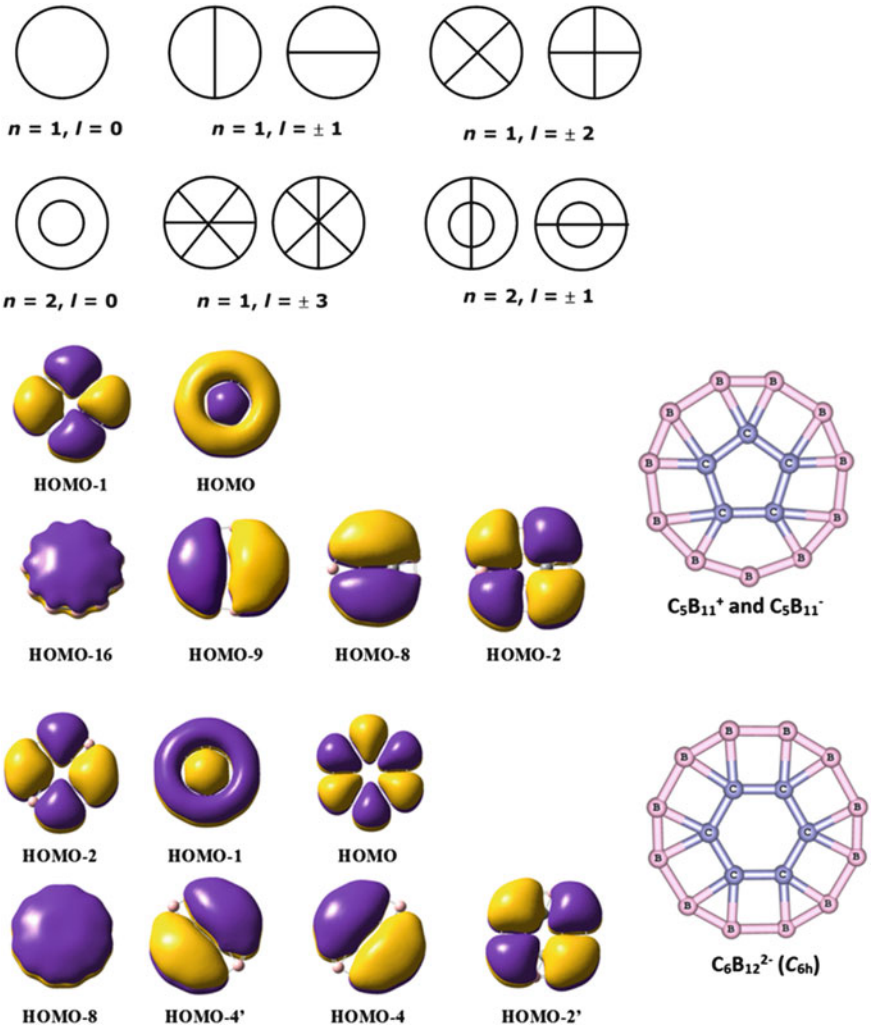


Fig. 12 Shapes of the valence π-orbitals of $C_5B_{11}^-$ and $C_6B_{12}^{2-}$ and the lowest-lying wave functions for a particle in a circular box

It is worthy to note that there is a concurrence in the number of valence electrons between the Hückel rule of $(4N+2)$ electrons and the disk aromaticity model at small systems ($N \leq 2$). While the Hückel rule is not realistic for polycyclic compounds, the planar monocyclic compounds whose number of valence π-electrons is larger than 10 is rather rare. Thus, the proposed concept of disk aromaticity is also expected as an effective measure of aromaticity for both planar monocyclic and polycyclic compounds.

6 Conclusion and Perspective

In this chapter, we reviewed some of the characteristics of aromaticity and chemical bonding of a few known mono- and poly-heterocyclic compounds containing sulfur and their hetero-derivatives. The inherent aromaticity was carefully considered using the topological analysis of the dissected CMO-NICS and electron localization function (ELF). The novel concept of *disk aromaticity* has been applied to some poly-heterocycles. We would hope that the theoretical approaches described in this chapter will be helpful for the heterocyclic chemists.

Acknowledgments The authors are indebted to the KU Leuven Research Council for continuing support (GOA, IDO, and PDM programs). We thank FWO Vlaanderen for a research grant on non-classical aromaticity and a postdoctoral fellowship (TBT).

References

1. Schleyer PR (2001) Chem Rev 101:1115–1118
2. Schleyer PR (2005) Chem Rev 105:3433–3435
3. Chattaraj PK (2012) Aromaticity and metal clusters. CRC Press, London
4. Kekulé A (1865) Bull Soc Chim Paris 3:98–110
5. Radenkovic S, Gutman I, Bultinck P (2012) J Phys Chem A 116:9421–9430
6. Krygowski TM, Cyranski MK (2009) Aromaticity in heterocyclic compounds in topic hetero chemistry, vol 19. Springer, Berlin
7. Huong VTT, Tai TB, Nguyen MT (2012) Phys Chem Chem Phys 14:14832–14841
8. Huong VTT, Nguyen MT, Tai TB, Nguyen MT (2013) J Phys Chem C. doi:10.1021/jp401191a
9. Watson MD, Fechtenkötter A, Müllen K (2001) Chem Rev 101:1267–1300
10. Randic M (2003) Chem Rev 103:3449–3605
11. Tai TB, Nguyen MT (2012) Angew Chem Int Ed 52:4554–4557
12. Boldyrev AI, Wang LS (2005) Chem Rev 105:3716–3757
13. Tai TB, Nguyen MT, Nguyen MT (2012) Theor Chem Acc 131:1241
14. Tai TB, Ceulemans A, Nguyen MT (2012) Chem Eur J 18:4510–4512
15. Tai TB, Nguyen MT, Nguyen MT (2012) Chem Phys Lett 530:71–76
16. Tai TB, Kadłubanski P, Roszak S, Majumdar D, Leszczynski J, Nguyen MT (2011) Chem Phys Chem 12:2948–2958
17. Tai TB, Nguyen MT (2010) Chem Phys 375:35–45
18. Tai TB, Grant DJ, Nguyen MT, Dixon DA (2010) J Phys Chem A 114:994–1007
19. Tai TB, Nguyen MT (2009) Chem Phys Lett 483:35–42
20. Hirsch A, Chen Z, Jiao H (2000) Angew Chem Int Ed 39:3915–3917
21. Chen Z, King RB (2005) Chem Rev 105:3613–3642
22. Bühl M, Hirsch A (2001) Chem Rev 101:1153–1183
23. Tai TB, Nguyen MT (2013) Chem Commun 49:913–915
24. Tam NM, Tai TB, Nguyen MT (2012) J Phys Chem C 116:20086–20098
25. Tai TB, Nguyen MT (2012) J Comput Chem 33:800–809
26. Tai TB, Tam NM, Nguyen MT (2011) Chem Phys 388:1–8
27. Tai TB, Nguyen MT (2011) J Phys Chem A 115:9993–9999
28. Tai TB, Nguyen MT (2011) J Chem Theory Comput 7:1119–1130
29. Tai TB, Hue NTM, Nguyen MT (2011) Chem Phys Lett 502:187–193

30. Tai TB, Nhat PV, Nguyen MT (2010) Phys Chem Chem Phys 12:11477–11486
31. Tai TB, Nguyen MT (2010) Chem Phys Lett 489:75–80
32. Tai TB, Nguyen MT (2010) Chem Phys Lett 492:290–296
33. Hückel E (1931) Z Phys 70:204–286
34. Schleyer PR, Maerker C, Dransfeld A, Jiao H, Hommes NJRE (1996) J Am Chem Soc 118:6317–6318
35. Schleyer PR, Jiao H, Hommes HJRE, Malkin VG, Malkina O (1997) J Am Chem Soc 119:12669–12670
36. Chen Z, Wannere CS, Corminboeur C, Puchta R, Schleyer PR (2005) Chem Rev 105:3842–3888
37. Pauling L (1936) J Chem Phys 4:673–677
38. London F (1937) J Phys Radium 8:397–409
39. Pauling L, Sherman J (1933) J Chem Phys 1:606–617
40. Dewar MJS, Gleicher GJ (1965) J Am Chem Soc 87:685–692
41. Schaad LJ, Hess BA (2001) Chem Rev 101:1465–1476
42. Aihara J (1976) J Am Chem Soc 98:2750–2758
43. Faraday M (1825) Philos Trans R Soc Lond 115:440–466.
44. Dewar MJS, McKee ML (1980) Pure Appl Chem 52:1431–1441
45. Dewar MJS (1984) J Am Chem Soc 106:669–682
46. Chandrasekhar J, Jemmis ED, Schleyer PR (1979) Tetrahedron Lett 20:3707–3710
47. Baird NC (1972) J Am Chem Soc 94:4941–4948
48. Aihara J (1978) J Am Chem Soc 100:3339–3342
49. Zhai HJ, Alexandrova AN, Birch KA, Boldyrev AI, Wang LS (2003) Angew Chem Int Ed 42:6004–6008
50. Alexandrova AN, Boldyrev AI, Zhai HJ, Wang LS (2006) Coord Chem Rev 250:2811–2866
51. Heilbronner E (1964) Tetrahedron Lett 1923–1926
52. Möbius AF (1865) über die Bestimmung des Inhaltes eines Polyëders. 17: 31–68
53. Listing JB (1861) Abhandlungen der Mathematischen Classe der Königlichen Gesellschaft der Wissenschaften zu Göttingen 10:97–182
54. Herges R (2006) Chem Rev 106:4820–4842
55. Rzepa HS (2005) Chem Rev 105:3697–3715
56. Kawasw T, Oda M (2004) Angew Chem Int Ed 43:4396–4398
57. Mauksch M, Gogonea V, Jiao H, Schleyer PR (1998) Angew Chem Int Ed 37:2395–2397
58. Pascal P (1910) Ann Chim Phys 19:5–70
59. Dauben HJ Jr, Wilson JD, Laity JL (1968) J Am Chem Soc 90:811–813
60. Paquette LA, Bauer W, Sivik MR, Bühl M, Feigel M, Schleyer PR (1990) J Am Chem Soc 112:8776–8789
61. Bohmann JA, Weinhold F, Farrar TC (1997) J Chem Phys 107:1173–1184
62. Pauling LJ (1936) Chem Phys 4:673–677
63. Pople JA (1958) Mol Phys 1:175–180
64. McWeeny R (1958) Mol Phys 1:311–321
65. Lazzeretti P, Zanasi R (1981) J Phys Chem 75:5019–5027
66. Lazzeretti P, Zanasi R (1981) Chem Phys Lett 80:533–536
67. Lazzeretti P (2000) Ring currents, vol 36, Progress in nuclear magnetic resonance spectroscopy. Elsevier, Amsterdam, pp 1–88
68. Steiner E, Fowler P (1996) Int J Quantum Chem 60:609–616
69. Juselius J, Sundholm D (1999) Phys Chem Chem Phys 1:3429–3435
70. Gomes JANF, Mallion RB (2001) Chem Rev 101:1349–1383
71. Poater J, Duran M, Sol M, Silvi B (2005) Chem Rev 105:3911–3947
72. Becke AD, Edgecombe KE (1990) J Chem Phys 92:5397
73. Silvi B, Savin A (1994) Nature 371:683–686
74. Melin J, Fuentealba P (2003) Int J Quantum Chem 92:381–390

75. Nguyen MT, Kryachko ES, Vanquickenborne LG (2003) In: Rappoport Z (ed) The chemistry of phenols, Patai series, The chemistry of functional groups, part 1. Wiley, Hoboken, pp 1–198
76. Santos JC, Tiznado W, Contreras R, Fuentealba F (2004) J Chem Phys 120:1670–1673
77. Heine T, Schleyer PR, Corminboeuf C, Seifert G, Reviakine R, Weber J (2003) J Phys Chem A 107:6470–6475
78. Steiner E, Fowler PW (2001) J Phys Chem A105:9553–9562
79. Corminboeuf C, Heine T, Seifert G, Schleyer PR, Weber J (2004) Phys Chem Chem Phys 6:273–276
80. Nyulaszi L, Schleyer PR (1999) J Am Chem Soc 1121:6872–6875
81. Dimitrakopoulos CD, Malenfant PRL (2002) Adv Mater 14:99–117, and references therein
82. Nelson SF, Lin YY, Gundlach DJ, Jackson TN (1998) Appl Phys Lett 72:1854/01-03
83. Cicoira F, Santato C (2002) Adv Funct Mater 17:3421–3434
84. Laquindanum JG, Katz HE, Lovinger AJ (1998) J Am Chem Soc 120:664–672
85. Garnier F, Yassar A, Hajlaoui R, Horowitz G, Deloffre F, Servet B, Ries S, Alnot P (1993) J Am Chem Soc 115:8716–8721
86. Wang C, Dong H, Hu W, Liu Y, Zhu D (2012) Chem Rev 112:2208–2267
87. Anthony JE (2006) Chem Rev 106:5028–5048
88. Murphy AR, Frechet JMJ (2007) Chem Rev 107:1066–1096
89. Mishra A, Ma CQ, Bauerle P (2009) Chem Rev 109:1141–1276
90. Barth WE, Lawton RG (1966) J Am Chem Soc 88:380–381
91. Robertson JM, White JG (1945) J Chem Soc: 607–617
92. Steiner E, Fowler PW, Jenneskens L (2001) Angew Chem Int Ed 40:362–365
93. Acocella A, Havenith RWA, Steiner E, Fowler PW, Jenneskens LW (2002) Chem Phys Lett 363:64–72
94. Yamamoto K, Harada T, Nakazaki M, Nakao T, Kay Y, Harada S, Kasai K (1998) J Am Chem Soc 110:3578–3584
95. Chernichenko KY, Balenkova ES, Nenajdenko VG (2008) Mendeleev Commun 18:171–179
96. Chernichenko KY, Sumerin VV, Shpanchenko RV, Balenkova ES, Nenajdenko VG (2006) Angew Chem Int Ed 45:7367–7370
97. Dadvand A, Cicoira F, Chernichenko KY, Balenkova ES, Osuma RM, Rosei F, Nenajdenko VG, Perepichka DF (2008) Chem Commun 5354–5356
98. Gahungu G, Zhang J (2008) Phys Chem Chem Phys 10:1743–1747
99. Salcedo R, Sansores LE, Picazo A, Sanson L (2004) J Mol Struct THEOCHEM 678:211–215
100. Huang W, Sergeeva AP, Zhai HJ, Averkiev BB, Wang LS, Boldyrev AI (2010) Nat Chem 2:202–206
101. Steiner E (2008) The chemistry maths book. Oxford University Press, Oxford, pp 391–413
102. Erhardt S, Frenking G, Chen Z, Schleyer PR (2005) Angew Chem Int Ed 44:1078–1082
103. Wu YB, Yuan CX, Yang P (2006) J Mol Struct THEOCHEM 765:35–38

Index

F. De Proft, P. Geerlings (eds.), *Structure, Bonding and Reactivity of Heterocyclic Compounds*, Topics in Heterocyclic Chemistry 38, DOI 10.1007/978-3-642-45149-2