Juan J. Novoa
Manuel F. Ruiz-López   *Editors*

8th Congress on Electronic Structure: Principles and Applications (ESPA 2012)

A Conference Selection from Theoretical Chemistry Accounts

Springer

Highlights in Theoretical Chemistry

Vol. 5
Series Editors: Ch.J. Cramer • D.G. Truhlar

Juan J. Novoa • Manuel F. Ruiz-López
Volume Editors

8th Congress on Electronic Structure: Principles and Applications (ESPA 2012)

A Conference Selection from Theoretical Chemistry Accounts

With contributions from

Manuel Alcamí • Diego R. Alcoba • Sergey Aldoshin • Muhannad Altarsha
Juan Aragó • Luis Miguel Azofra • V. G. Baonza • Xavier Barril
M. I. Bernal-Uruchurtu • Konstantin Bozhenko • Stefan T. Bromley
Joaquín Calbo • Josep M. Campanera • Rodrigo Casasnovas • Luigi Cavallo
A. Cedillo • Bo Y. Chang • A. Cimas • Veronica Collico • I. Corral
Mercè Deumal • Sergio Díaz-Tendero • Nina Emel'yanova • Volker Engel
Joaquin Espinosa-García • Mirjam Falge • Juan Frau • Hong Fu
Ryusuke Futamura • Ricard Gelabert • José R. B. Gomes
Sławomir J. Grabowski • Tobias Hell • Stefan E. Huber • Francesc Illas
Francesca Ingrosso • Miguel Jorge • Alexander Krivenko • Luis Lain
Oriol Lamiel-Garcia • Al Mokhtar Lamsabhi • José M. Lluch • Xabier Lopez
F. Javier Luque • Roman Manzhos • M. Marqués • Antonio M. Márquez
Fernando Martín • Jon M. Matxain • J. M. Menéndez • Otilia Mó
Manuel Monge-Palacios • M. Merced Montero-Campillo • A. Morales-García
Miquel Moreno • Francisco Muñoz • Marc Nadal-Ferret • Roger Nadler
Juan J. Novoa • Josep M. Oliva • Enrique Ortí • Alexander Ostermann
Mario Piris • José J. Plata • Albert Poater • Michael Probst • Carlos Randino
Cipriano Rangel • J. M. Recio • Maitreyi Robledo • Manuel F. Ruiz-López
Nataliya Sanina • D. Santamaría-Pérez • J. J. Santoyo-Flores • Javier Fdez Sanz
Sebastián Sastre • Ignacio R. Sola • Alicia Torre • Mark M. Turnbull
Jesus M. Ugalde • Sergi Vela • R. Verzeni • Patricia Vindel-Zandbergen
Manuel Yáñez • Minghui Yang

 Springer

Volume Editors
Juan J. Novoa
Departament de Química Física & IQTCUB
Facultat de Química
Universitat de Barcelona
Barcelona, Spain

Manuel F. Ruiz-López
SRSMC, Theoretical Chemistry
 and Biochemistry Group
University of Lorraine, CNRS
Vandoeuvre-les-Nancy, France

Contents

Preface to the ESPA-2012 special issue

Juan J. Novoa · Manuel F. Ruiz-López

This issue of Theoretical Chemistry Accounts contains a recollection of some of the work presented and discussed at the 8th edition of the Electronic Structure: Principles and Applications (in short, ESPA-2012). The ESPA events are biennial international research conferences organized within the activities of the Spanish Theoretical Chemistry groups that co-organize the Interuniversity Doctorate in Theoretical Chemistry and Computational Modeling. The main aim behind all ESPA conferences, shared by the organizers of ESPA-2012, is promoting scientific excellence and exchange of ideas among their Ph. D. students, in a friendly environment. ESPA-2012 follows previous events held in Madrid, San Sebastián, Sevilla, Valladolid, Santiago de Compostela, Palma de Mallorca, and Oviedo.

ESPA-2012 took place in Barcelona from the 26th up to the 29th of June 2012, in a magnificent location: the Auditorium of CosmoCaixa in Barcelona, the Science Museum created and supported by "La Caixa" savings bank in the hills that overlook Barcelona from the North. We all remember the superb auditorium facilities, together with its amazing views to the Science Museum and the city of Barcelona. The conference was organized by Prof. Juan J. Novoa (Chairman), helped by (al alphabetical order) Albert Bruix (Ph. D. student), Prof. Rosa Caballol, Marçal Capdevila (Ph. D. student), Dr. Mercè Deumal, Prof. Javier Luque, Dr. Iberio de P. R Moreira, Dr. Fernando Mota, Dr. Jordi Ribas-Ariño, Prof. Ramón Sayós, Dr. Carmen Sousa, and Sergi Vela (Ph. D. student). A picture of the Organizing Committee is displayed in Fig. 1.

ESPA-2012 was designed guided by three main principles: (1) passion for discovery, (2) scientific excellence, and (3) a friendly environment. For sure, all ESPA-2012 participants shared the same emotions beautifully described by Herman Melville in his "Moby Dick" book: "… but as for me, I am tormented with an everlasting itch for things remote. I love to sail forbidden seas, and land on barbarous coasts." Concerning our passion for Science, for sure, most ESPA-2012 participants went to Barcelona with the aim of reporting their discoveries while "sailing the Theoretical Chemistry and Computational Modeling seas," and also listening at other participant's reports. After all, modern scientific research is a cooperative effort, where it is still valid Isaac Newton's statement: "If I have seen further is by standing on the shoulders of giants."

In relation to excellence, it is sometimes stated that the quality of a conference can be measured, at least partially, by the stature of its invited speakers. Aiming at excellence, in ESPA-2012, we had as invited speakers some of the world leaders in the field of Theoretical Chemistry and Computational Modeling. Each one gave one of the nine Invited Plenary Talks: The Opening Plenary Talk was delivered by Prof. M. A. Robb (Imperial College London; Fellow of the Royal Society of Chemistry) and the Closing Plenary Talk was given by Prof. W. L. Jorgensen (Yale University, CT, USA, Co-editor of Journal of Chemical Theory and Computation). The remaining seven Invited

J. J. Novoa (✉)
Departament de Química Física & IQTCUB, Facultat de Química, Universitat de Barcelona, Av. Diagonal 645, Barcelona 08028, Spain
e-mail: juan.novoa@ub.edu

M. F. Ruiz-Lopez (✉)
SRSMC, Theoretical Chemistry and Biochemistry Group, University of Lorraine, CNRS, 54506 Vandoeuvre-les-Nancy, France
e-mail: Manuel.Ruiz@univ-lorraine.fr

Plenary Talks were presented (in alphabetical order) by Prof. Johan Aqvist (Uppsala University), Prof. Bjork Hammer (Aarhus University), Prof. Pavel Hobza (Institute of Organic Chemistry and Biochemistry), Prof. Frank Neese (Max Planck Institute for Bioinorganic Chemistry), Prof. Matthias Scheffler (Fritz Haber Institute), Prof. Sason S. Shaik (The Hebrew University), and Prof. Manuel Yañez (Universidad Autónoma de Madrid). Each Invited Plenary Talk had an allocated time of 45 min (40 min of presentation, followed by 5 min of questions, that is, $40' + 5'$ talks). Besides them, there were 25 Contributed Talks ($15' + 5'$ each), selected by the Organizing Committee among all propositions, and about 200 posters, also previously evaluated by the Organizing Committee. Two poster sessions were allocated for their presentation by one of their authors (each session lasting 2 hours).

Fig. 1 a Picture of the ESPA-2012 Organizing Committee taken in the Main Entrance to the Chemistry Building of the University of Barcelona. *Lower row* (from *left* to *right*): J. J. Novoa, C. Sousa, R. Sayós, J. Ribas-Ariño, I. de P. R Moreira, M. Deumal; *Upper row* (from *left* to *right*): J. Luque, R. Caballol, A. Bruix, S. Vela, M. Capdevila, F. Mota. b ESPA-2012: A cooperative work, illustrated by a picture of a six-floor Human Castle, a Catalan tradition
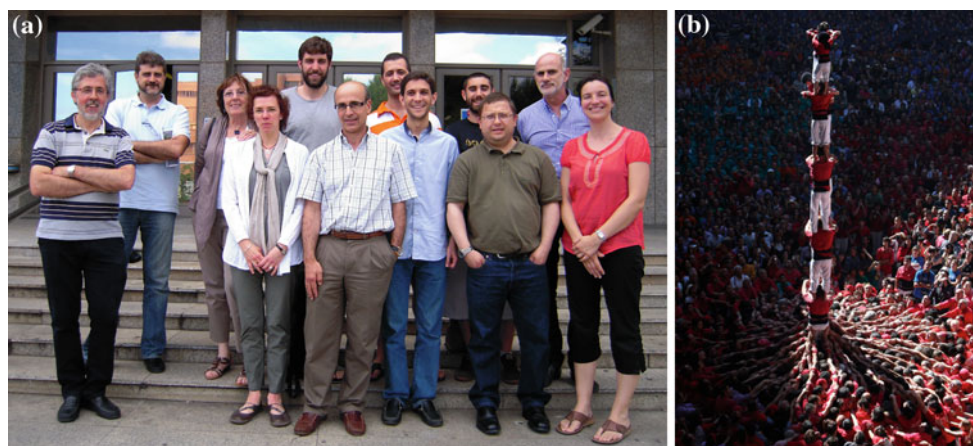
Cosmocaixa			Wed., 27 / 6 / 2012
	9:00 - 9:30		OPENING CEREMONY
			ORAL SESSION M1
Michael A. Robb	9:30 - 10:15	Plenary 1	Non-adiabatic Chemistry at the Conical Intersection Seam Studied With "On the fly" Quantum Dynamics
Ángel Martín Pendás	10:15 - 10:35	Theory, methods & foundations 1	Fluctuation of electron populations and chemical bonding
Coen de Graaf	10:35 - 10:55		The effect of thermal motion on the spectroscopic properties of NiO
Alessandro Erba	10:55 - 11:15		Temperature Effects on the Density Matrix of Crystals: An ab initio Monte Carlo Harmonic Approach
	11:15 - 11:45		COFFEE BREAK
			ORAL SESSION M2
Johan Aqvist	11:45 - 12:30	Plenary 2	Cold-Adaptation of Enzyme Reaction Rates
Alexandra T. P. Carvalho	12:30 - 12:50	Biomolecular modelling 1	Role of the amide groups in the tuning of redox reactions
Carles Curutchet	12:50 - 13:10		Ultrafast light-harvesting dynamics in the PE545 photosynthetic antenna
Dan T. Major	13:10 - 13:30		Electrostatically Guided Enzyme Dynamics – the Root of Fidelity in a Promiscuous Terpene Synthase: The Case of Bornyl Diphosphate Synthase
	13:30 - 15:30		LUNCH
			ORAL SESSION T1
Matthias Scheffler	15:30 - 16:15	Plenary 3	The random-phase approximation for the electron correlation energy and beyond
Carmen J. Calzado	16:15 - 16 35	Materials science 1	Magnetic interactions in polynuclear transition metal systems from wave function based methods
Johannes Gierschner	16:35 - 16:55		Tuning of Optical & Photophysical Properties in Organic Crystals
José R.B. Gomes	16:55 - 17:15		The Early Stages of the Synthesis of Hybrid Mesoporous Organosilicas: DFT and Molecular Dynamics Simulations
	17:15 - 17:45		COFFEE BREAK
	17:45 - 19:00		POSTER SESSION P1
	20:00 - 23:30		CULTURAL ACTIVITY: A NIGHT AT THE OPERA

Fig. 2 Detailed Wednesday 27 Program

Cosmocaixa			Thu., 28 / 6 / 2012
ORAL SESSION M3			
Pavel Hobza	9:00 - 9:45	Plenary 4	Nonempirical, semiempirical and empirical interaction energies
Slawomir J. Grabowski	9:45 - 10:05	Theory, methods & foundations 2	The Description of Non-Covalent Interactions in Terms of Bent's Rule
Enrique Sánchez Marcos	10:05 - 10:25		A multitechnique approach to elucidate axial hydration of square-planar complexes of metal cations
Julia Contreras-García	10:25 - 10:45		NCI: a new quantitative method for revealing non-covalent interactions
Manuel Yáñez	10:45 - 11:30	Plenary 5	Non-covalent interactions: From hydrogen bonds to beryllium bonds or the importance of deformation
	11:30 - 12:00	**COFFEE BREAK**	
ORAL SESSION M4			
Sason Shaik	12:00 - 12:45	Plenary 6	The Valence Bond way in Chemistry
Gregori Ujaque	12:45 - 13:05	Chemical reactivity 1	AIMD approach to Homogeneous Catalysis: The Wacker and Shilov Reactions
Carmen Barrientos	13:05 - 13:25		Searching for the Origin of Biomolecules in Space
Margarita I. Bernal-Uruchurtu	13:25 - 13:45		Bromine Dissociation in aqueous clusters. The catalytic role of water
	13:45 - 16:00	**LUNCH**	
	16:00 - 20:00	**CULTURAL ACTIVITY: EXCURSION**	

Fig. 3 Detailed Thursday 28 Program

For didactic reasons, all presentations were grouped into one of the following four thematic areas that we have drawn on for the presentation of this TCAC Volume: [1] Theory, methods and foundations (TMF), [2] Chemical Reactivity (CR), [3] Biomolecular Modeling (BM), and (4) Materials Science (MS). In order to further facilitate the effort of the audience, they were presented in thematic sessions, constituted by one Plenary Talk and three Contributed Talks, whenever possible. There were two morning sessions, separated by a coffee break, on 3 days with scientific sessions (27, 28, and 29 of June) and two afternoon sessions, separated by another coffee break, on the 27th and also on the 29th. There was a first poster session on the afternoon of the 27th (posters of the thematic areas TMF and MS) and another on the afternoon of the 29th (poster of thematic areas CR and BM).

It was in our stated aim to make the atmosphere of ESPA-2013 as friendly as possible. With this idea in mind, we prepared a rich social program, with events every day and non-overlapping in time with the scientific program. The ESPA-2012 Conference started in the afternoon of the 26th with the Registration and Welcome Party. All participants were asked to register at the Historical Building of University of Barcelona, located downtown Barcelona. Registration was followed by the first social activity: a Welcome Party that took place, in the late afternoon hours, under the shade of the old trees planted in the Historical Building gardens. All participants had a chance to meet old

friends and make new ones, while enjoying live piano music and a snack served with wine or non-alcoholic beverages. The activities of the social program ended with a Conference Dinner on the 29th, in a restaurant overlooking Barcelona and with superb views over the city night-lights. In between these two events, on 27th, there was an "A night at the Opera" event for all participants interested on opera, which took place at the Barcelona Opera House (whose local nickname is "El Liceu"), where we watched and listened the start-up performance of "Pelléas et Mélisande," a Debussy's opera. On 28th, we all had a "paella" at the Barcelona Olympic Harbour, followed by an afternoon visit to the three most impressive Gaudi's architectural masterpieces located in Barcelona: "Parc Guell," "Sagrada Familia," and "La Pedrera." Besides these activities, lunch on the 27th and 29th was arranged by the Organizing Committee for all participants in a high-end restaurant located nearby CosmoCaixa Auditorium (a bus shuttle service was provided by the organization, both directions).

The scientific program of ESPA-2012 started in the morning of June 27, with the Opening Ceremony presided by the Chancellor of the University of Barcelona, Prof. Didac Ramirez. Afterward, we had the six morning sessions, two afternoon sessions, and two poster sessions, and their speakers and titles are shown in Figs. 2, 3, and 4 (for Wednesday 27, Thursday 28, and Friday 29). The scientific part of ESPA-2012 ended on the afternoon of June 29th

Cosmocaixa			Fri., 29 / 6 / 2012
			ORAL SESSION M5
Bjørk Hammer	9:00 - 9:45	Plenary 7	Observation and modeling of chiral transfer on Pt(111)
José A. Gámez	9:45 - 10:05		Co-operating N_2- and Phenyl Rotations in trans-Azobenzene Photo-Isomerisation
Lluís Blancafort	10:05 - 10:25	Chemical reactivity 2	Wave Packet Dynamics at an Extended Seam of Conical Intersection: Mechanism of the Light-induced Wolff Rearrangement
Alejandro Toro-Labbé	10:25 - 10:45		The Mechanism of Hydrogen Activation by (Amino)Carbenes
	10:45 - 11:15		**COFFEE BREAK**
			ORAL SESSION M6
Frank Neese	11:15 - 12:00	Plenary 8	Efficient Coupled Cluster Calculations on Large Molecules Using Pair Natural Orbitals
Andreas Heidenreich	12:00- 12:20	Materials science	Coulomb explosion dynamics and energetics of multiply ionized clusters
Francesca Ingrosso	12:20 - 12:40	Biomolecular Modelling	β-Cyclodextrins in Supercritical CO_2: a Theoretical Chemistry Viewpoint
Irena Efremenko	12:40 - 13:00	Chemical reactivity	Protonation of Polyoxometalates and its Implications for the Initial Steps of Mars-van Krevelen Mechanism
	13:00 - 14:45		**LUNCH**
	14:45 - 16:00		**POSTER SESSION P2**
			ORAL SESSION T2
Jörg Grunenberg	16:00 - 16:20		Complexity in Molecular Recognition
José J. Ruiz-Pernía	16:20 - 16:40	Biomolecular Modelling 2	Dissecting the role of Entropy in Catalysis through simulations Chorismate Mutase revisited
Petra Imhof	16:40 - 17:00		RNA Catalysis: Molecular Simulations of the Diels-Alderase Ribozyme
William Jorgensen	17:00 - 17:45	Plenary 9	Drug Discovery Accelerated by Computational Methods
	17:45 - 18:00		**CLOSING REMARKS**
	21:00 - 23:00		**CULTURAL ACTIVITY: BANQUET**

Fig. 4 Detailed Friday 29 Program

with the Closing Remarks Ceremony. The ceremony started with the presentation of the four poster prizes to their winners, one per each of the four thematic areas in which all posters were grouped (see above). Afterward, the Conference Chairman wished a safe trip back home to all participants and also strength for the difficult economic times still to come. The Closing Remarks Ceremony ended with a special "see you soon, friends," using the first two verses of a beautiful farewell Catalan song: "If you tell me farewell, I wish that it would be in a clear and bright day." Then, following the traditions of this part of Spain, all participants had a chance to say farewell while drinking a cup of Catalan cava, served chilled in the gardens of CosmoCaixa.

There were 261 participants at ESPA-2012, 165 of them with a Ph. D. degree and 96 Ph. D. students. Most

participants were Spanish. All others came from 20 different countries (in alphabetical order, Argentina, Algeria, Austria, Brazil, Bulgaria, Chile, Czech Republic, Denmark, France, Germany, India, Israel, Italy, Mexico, New Zealand, Portugal, Russia, Sweden, United Kingdom, and the United States of America).

As you have seen, we all enjoyed the meeting at ESPA-2012: Lots of good scientific ideas, time to talk with our old and new friends about them, time to enjoy visiting Barcelona and some of its cultural highlights.

We hope to see you again at the next ESPA, ESPA-2014! In the meantime, our best wishes to all with a final quote, attributed to Albert Einstein: "The most beautiful thing we can experience is the mysterious. It is the source of all true Art and all Science."

REGULAR ARTICLE

The one-electron picture in the Piris natural orbital functional 5 (PNOF5)

**Mario Piris · Jon M. Matxain ·
Xabier Lopez · Jesus M. Ugalde**

Abstract The natural orbital functional theory provides two complementary representations of the one-electron picture in molecules, namely, the natural orbital (NO) representation and the canonical orbital (CO) representation. The former arises directly from the optimization process solving the corresponding Euler equations, whereas the latter is attained from the diagonalization of the matrix of Lagrange multipliers obtained in the NO representation. In general, the one-particle reduced-density matrix (1-RDM) and the Lagrangian cannot be simultaneously brought to the diagonal form, except for the special Hartree-Fock case. The 1-RDM is diagonal in the NO representation, but not the Lagrangian, which is only a Hermitian matrix. Conversely, in the CO representation, the Lagrangian is diagonal, but not the 1-RDM. Combining both representations we have the whole picture concerning the occupation numbers and the orbital energies. The Piris natural orbital functional 5 leads generally to the localization of the molecular orbitals in the NO representation. Accordingly, it provides an orbital picture that agrees closely with the empirical valence shell electron pair repulsion theory and the Bent's rule, along with the theoretical valence bond method. On the other hand, the equivalent CO representation can afford delocalized molecular orbitals adapted to the symmetry of the molecule. We show by means of the extended Koopmans' theorem that the one-particle energies associated with the COs can yield reasonable principal ionization potentials when the 1-RDM remains close to the diagonal form. The relationship between NOs and COs is illustrated by several examples, showing that both orbital representations complement each other.

Published as part of the special collection of articles derived from the 8th Congress on Electronic Structure: Principles and Applications (ESPA 2012).

M. Piris (✉) · J. M. Matxain · X. Lopez · J. M. Ugalde
Kimika Fakultatea, Donostia International Physics Center
(DIPC), Euskal Herriko Unibertsitatea (UPV/EHU),
P.K. 1072, 20080 Donostia, Euskadi, Spain
e-mail: mario.piris@ehu.es

J. M. Matxain
e-mail: jonmattin.matxain@ehu.es

X. Lopez
e-mail: xabier.lopez@ehu.es

J. M. Ugalde
e-mail: jesus.ugalde@ehu.es

M. Piris
IKERBASQUE, Basque Foundation for Science,
48011 Bilbao, Euskadi, Spain

1 Introduction

One-electron pictures have long helped to our understanding of chemical bonding. The simplest one-electron model is based on the independent-particle Hartree-Fock (HF) approximation [8, 12]. However, it shows limitations due to the lack of the electron correlation. Many-electron effects can be taken into account with an adequate approximation of the 2-RDM since the molecular energy is determined exactly by the two-particle reduced-density matrix (2-RDM). Correlated wavefunction theory (WFT) approximations provide accurate 2-RDMs, hence Brueckner [2] and Dyson orbitals [23, 35] are reliable methods for determining a set of one-particle functions [36].

Unfortunately, such theories demand significant computational resources as the size of the systems of interest increase.

On the other hand, the density functional theory (DFT) [37] has become very popular in the computational community because electron correlation is treated in an effective one-particle framework. DFT replaces the two-particle problem with a one-particle exchange-correlation potential. In doing so, a calculation comparable to a HF one is possible with a relatively low computational cost, even though practical DFT methods suffer from several errors like those arising from electron self-interaction, the wrong long-range behavior of the Kohn-Sham (KS) [17] potentials, etc. Current implementations of DFT are mainly based on the KS formulation, in which the kinetic energy is not constructed as a functional of the density, but rather from an auxiliary Slater determinant. Since the non-interacting kinetic energy differs from the many-body kinetic energy, there is a contribution from a part of the kinetic energy contained in the correlation potential. The incorrect handling of the correlation kinetic energy is one main source of problems of present-day KS functionals.

The density matrix functional theory (DMFT) has emerged in recent years as an alternative method to conventional WFT and DFT. The idea of a one-particle reduced-density matrix (1-RDM) functional appeared few decades ago [9, 21, 22, 56]. The major advantage of a density matrix formulation is that the kinetic energy is explicitly defined, and it does not require the construction of a functional. The unknown functional only needs to incorporate electron correlation. The 1-RDM functional is called natural orbital functional (NOF) when it is based upon the spectral expansion of the 1-RDM. The natural orbitals (NOs) [25] are orthonormal with fractional occupancies, allowing to unveil the genuine electron correlation effects in terms of one-electron functions. Valuable literature related to the NOF theory (NOFT) can be found in Refs. [42] and [43].

It is important to note that functionals currently in use are only known in the basis where the 1-RDM is diagonal. This implies that they are not functionals explicitly dependent on the 1-RDM and retain some dependence on the 2-RDM. So far, all known NOFs suffer from this problem including the exact NOF for two-electron closed-shell systems [11]. The only exception is the special case of the HF energy that may be viewed as a 1-RDM functional. Accordingly, the NOs obtained from an approximate functional are not the exact NOs corresponding to the exact expression of the energy. In this vein, they are NOs as the orbitals that diagonalize the 1-RDM corresponding to an approximate expression of the energy, like those obtained from an approximate WFT.

One route [51, 52] to the construction of approximate NOF involves the employment of a reconstruction functional based on the cumulant expansion [19, 29] of the 2-RDM. We shall use the reconstruction functional proposed in [41], in which the two-particle cumulant is explicitly reconstructed in terms of two matrices, $\Delta(\mathbf{n})$ and $\Pi(\mathbf{n})$, $\mathbf{n}$ being the set of the occupation numbers. The $\Delta(\mathbf{n})$ and $\Pi(\mathbf{n})$ matrices satisfy known necessary N-representability conditions [30, 32] and sum rules of the 2-RDM, or equivalently, of the functional. Moreover, precise constraints that the two-particle cumulant matrix must fulfill in order to conserve the expectation values of the total spin and its projection have been formulated and implemented for the matrices $\Delta(\mathbf{n})$ and $\Pi(\mathbf{n})$ [46]. Appropriate forms of the matrices $\Delta(\mathbf{n})$ and $\Pi(\mathbf{n})$ led to different implementations of NOF, known in the literature as PNOFi ($i = 1$–5) [41, 44, 45, 47, 48]. A detailed account of these functionals can be found elsewhere [43]. Because PNOF theory is based on both the 1- and the 2-RDMs, it has connections to the parametric 2-RDM methods of Refs. [31, 54]

It has recently been pointed out [28] that PNOF5 [24, 27, 44] can provide a NO picture that agrees closely with the empirical valence shell electron pair repulsion theory (VSEPR) [10] and the Bent's rule [1], along with the popular theoretical valence bond (VB) method [13, 58]. Although PNOF5 can predict additionally three- and four-center two-electron bonds, in general, the solutions of the PNOF5 equations lead to orbital hybridization and to localization of the NOs in two centers, providing a natural language for the chemical bonding theory.

Nevertheless, in some systems the electronic structure is better understood through orbital delocalization. Typical cases are the aromatic systems like benzene molecule. This point of view was introduced by Hund [14] and Mulliken [34] within the framework of the linear combination of atomic orbitals–molecular orbital (LCAO-MO) theory, in which orbitals can extend over the entire molecule. Later, Koopmans [18] demonstrated, using the HF approximation in the framework of the LCAO-MO theory, one of the most important connections between orbitals and the experiment: the HF orbital energies are directly associated with ionization energies. Accordingly, it raises the question of how to achieve a delocalized one-particle orbital representation that complements the NO representation in PNOF5.

In this paper, we introduce an equivalent orbital representation to the NO one, in which the molecular orbitals are delocalized. These orbitals are not obtained arbitrarily, but arise from the diagonalization of the matrix of Lagrange multipliers, or the Lagrangian, obtained in the NO representation, so we will call them canonical orbitals (COs) by analogy to the HF COs. It is important to recall that only for functionals explicitly dependent on the 1-RDM, the

1-RDM and the Lagrangian may be simultaneously brought to the diagonal form by the same unitary transformation [6]. On the contrary, in our case, the functional still depends on the 2-RDM, hence both matrices do not commute. Moreover, we cannot expect that it should be possible to bring the 1-RDM and the Lagrangian simultaneously to diagonal form in the case of finite order of the one-particle set [25]. In summary, only in the HF case, it is possible to find one representation in which both matrices are diagonal. For all the other known NOFs, there are two unique representations that diagonalize separately each matrix.

In the NO representation, the 1-RDM is diagonal, but not the Lagrangian, so the eigenvalues of the former afford the occupation numbers of the NOs corresponding to the proposed approximate functional. On the other hand, in the CO representation, the matrix of Lagrange multipliers is diagonal, but not anymore the 1-RDM. Taking into account the terminology developed by Coulson and Longuet-Higgins [3], we have in the CO representation the charge order of the orbital in the diagonal elements of the 1-RDM and the bond order of two orbitals in the off-diagonal elements. But even here, the charge order may be interpreted as the average number of particles in the orbital under consideration [25].

In contrast to the NO representation, the diagonal elements of the Lagragian in the CO representation can be physically meaningful. We demonstrate below, using the extended Koopmans' theorem (EKT) [4, 5, 33, 55], that the new one-particle energies can describe satisfactorily the principal ionization potentials (IPs), when the 1-RDM is close to the diagonal form in the CO representation. These one-particle energies account for the electron correlation effects, but evidently they neglect relaxation of the orbitals in the $(N - 1)$-state and consequently tend to produce too positive IPs. In the next section, the theory related to PNOF5 COs is presented. The relationship between NOs and COs is examined then by several examples.

2 Theory

The Piris natural orbital functional (PNOF) for singlet states reads as [41]

$$E = 2 \sum_p n_p H_{pp} + \sum_{pq} \Pi_{qp} L_{pq}$$
$$+ \sum_{pq} (n_q n_p - \Delta_{qp})(2J_{pq} - K_{pq}) \tag{1}$$

where p denotes the spatial NO and n_p its occupation number (ON). H_{pp} is the pth matrix element of the kinetic energy and nuclear attraction terms, whereas $J_{pq} = \langle pq|pq \rangle$ and $K_{pq} = \langle pq|qp \rangle$ are the usual Coulomb

and exchange integrals, respectively. $L_{pq} = \langle pp|qq \rangle$ is the exchange and time-inversion integral [40, 50]. Note that if Δ and Π vanish, then our reconstruction yields the HF energy, as expected. Moreover, for real orbitals exchange integrals and exchange and time-inversion integrals coincide, $L_{pq} = K_{pq}$. In PNOF5, we have adopted the following expressions [44]

$$\Delta_{pq} = n_p^2 \delta_{pq} + n_p n_{\tilde{p}} \delta_{\tilde{p}q} \tag{2}$$

$$\Pi_{pq} = n_p \delta_{pq} - \sqrt{n_p n_{\tilde{p}}} \delta_{\tilde{p}q} \tag{3}$$

The $\tilde{p}$-state defines the coupled NO to the orbital p, namely, $\tilde{p} = N - p + 1$, N being the number of particles in the system. Bounds that stem from imposing N-representability-necessary conditions on the 2-RDM imply that the ON of the $\tilde{p}$ level must coincide with that of the hole of its coupled state p, namely,

$$n_{\tilde{p}} = h_p, \qquad n_{\tilde{p}} + n_p = 1, \tag{4}$$

where h_p denotes the hole $1 - n_p$ in the spatial orbital p. In accordance to the Eq. (4), all occupancies vanish for $p > N$. Assuming a real set of NOs, the PNOF5 energy for a singlet state of an N-electron system is cast as [44]:

$$E = \sum_{p=1}^{N} \left[n_p (2H_{pp} + J_{pp}) - \sqrt{n_{\tilde{p}} n_p} K_{p\tilde{p}} \right]$$
$$+ \sum_{p,q=1}^{N} {}'' n_q n_p (2J_{pq} - K_{pq}), \tag{5}$$

The double prime in Eq. (5) indicates that both the $q = p$ term and the coupled one-particle state terms $p = \tilde{p}$ are omitted from the last summation. One must look for the pairs of coupled orbitals $(p, \tilde{p})$ that yield the minimum energy for the functional of Eq. (5). The actual p and $\tilde{p}$ orbitals paired are not constrained to remain fixed along the orbital optimization process. As a consequence, an orbital localization occurs generally, which corresponds to the most favorable orbital interactions [28]. This situation contrasts with our previous approximations PNOFi ($i = 1$–4) [41, 45, 47, 48], in which, the off-diagonal elements Δ_{pq} and Π_{pq} were formulated for all possible (p, q) pairs, leading to delocalized NOs.

The solution is established by optimizing the energy functional (5) with respect to the ONs and to the NOs, separately. PNOF5 allows constraint-free minimization with respect to the ONs, which yields substantial savings of computational time [44]. Therefore, one has to minimize the energy (5) with respect to the real orbitals $\{ \varphi_p(\mathbf{r}) \}$ under the orthonormality constraints. Introducing the matrix of symmetric Lagrange multipliers $\mathbf{\Lambda} = \{ \lambda_{qp} \}$, the functional whose extremum we seek is given by

$$\Omega = E - 2\sum_{pq} \lambda_{qp}\left[<\varphi_p|\varphi_q> - \delta_{pq}\right] \tag{6}$$

The Euler equations for the functions $\{\varphi_p(\mathbf{r})\}$ are,

$$n_p \hat{V}_p|\varphi_p\rangle = \sum_q \lambda_{qp}|\varphi_q\rangle \tag{7}$$

Multiplying Eq. (7) by $\langle\varphi_q|$, the matrix representation of this equation is

$$n_p\langle\varphi_q|\hat{V}_p|\varphi_p\rangle = \lambda_{qp} \tag{8}$$

The one-particle operator $\hat{V}_p$ is given by

$$\begin{aligned}\hat{V}_p(1) = \hat{H}(1) + \hat{J}_p(1) - \sqrt{\frac{h_p}{n_p}}\hat{K}_{\bar{p}}(1) \\ + \sum_{q=1}^{N}{}'' n_q\left[2\hat{J}_q(1) - \hat{K}_q(1)\right]\end{aligned} \tag{9}$$

with

$$\hat{J}_q(1) = \langle\varphi_q|\,r_{12}^{-1}|\varphi_q\rangle, \hat{K}_q(1) = \langle\varphi_q|r_{12}^{-1}\hat{P}_{12}|\varphi_q\rangle$$

The $\hat{P}_{12}$ operator permutes electrons 1 and 2, and the integration is carried out only over the coordinates of 2. Notice that the $\hat{V}_p$ operator is pth orbital dependent, it is not a mean field operator like, for instance, the Fock operator. One consequence of this is that the Lagrangian matrix $\mathbf{\Lambda}$ and the 1-RDM $\mathbf{\Gamma}$ do not commute; $[\mathbf{\Lambda}, \mathbf{\Gamma}] \neq 0$, therefore, they cannot be simultaneously brought to diagonal form by the same unitary transformation $\mathbf{U}$. Thus, Eq. (7)–(8) cannot be reduced to a pseudo-eigenvalue problem by diagonalizing the matrix $\mathbf{\Lambda}$. Actually, apart from the special HF case, where the 1-RDM is idempotent and the energy may be viewed as a 1-RDM functional, none of the currently known NOFs have effective potentials that allow to diagonalize simultaneously both matrices $\mathbf{\Gamma}$ and $\mathbf{\Lambda}$.

In this paper, the efficient self-consistent eigenvalue procedure proposed in [53] is employed to solve Eq. (7). It yields the NOs by iterative diagonalization of a Hermitian matrix $\mathbf{F}$. The off-diagonal elements of the latter are determined from the hermiticity of the matrix of the Lagrange multipliers $\mathbf{\Lambda}$. An expression for diagonal elements is absent, so a generalized Fockian is undefined in the conventional sense; nevertheless, they may be determined from an aufbau principle [53].

Using the expressions for diagonal elements of $\mathbf{\Lambda}$, let us rewrite the energy functional (5) as follows:

$$E = \sum_{p=1}^{N}\left[n_p H_{pp} + \lambda_{pp}\right] \tag{10}$$

Let us recall that the trace of an $N \times N$ square matrix is the sum of its diagonal elements, then the Eq. (10) can be rewritten as

$$E = \text{Tr}(\mathbf{H\Gamma} + \mathbf{\Lambda}) \tag{11}$$

Taking into account that the trace of a matrix is invariant under a unitary transformation $\mathbf{U}$, the energy (11) keeps constant under such transformation of the orbitals, that is,

$$\begin{aligned}\text{Tr}(\mathbf{H\Gamma} + \mathbf{\Lambda}) = \text{Tr}(\mathbf{U}^\dagger\mathbf{HUU}^\dagger\mathbf{\Gamma U} + \mathbf{U}^\dagger\mathbf{\Lambda U}) \\ = \text{Tr}(\mathbf{H}'\mathbf{\Gamma}' + \mathbf{\Lambda}')\end{aligned} \tag{12}$$

Accordingly, it is always possible to find a matrix $\mathbf{U}$ such that the transformation $\mathbf{\Lambda}' = \mathbf{U}^\dagger\mathbf{\Lambda U}$ diagonalizes $\mathbf{\Lambda}$. It is worth to note that the transformed 1-RDM $\mathbf{\Gamma}' = \mathbf{U}^\dagger\mathbf{\Gamma U}$ is not anymore a diagonal matrix. Such unitary transformation exists and is unique. Orbitals $\{\chi_p(\mathbf{r})\}$, for which the matrix of Lagrange multipliers is diagonal, will be called COs by analogy to the HF COs. One should note that the Lagrangian $\mathbf{\Lambda}$ is a symmetric matrix only at the extremum; ergo, this procedure for obtaining the COs can be solely used after the NOs have been obtained. In contrast to PNOF5 NOs, which are localized, more in line with our intuitive feeling for chemical bonds, the PNOF5 COs will generally be delocalized.

Analogously to HF COs, PNOF5 COs may also form a basis for an irreducible representation of the point group of the molecule. Similar to the Fock matrix, the Lagrangian $\mathbf{\Lambda}$ depends itself on the orbitals that have to be determined. It is well known that if there is any symmetry present in the initial guess of the HF COs, then this symmetry will be preserved at the SCF solution [15]. In this vein, if the initial guess for the NOs was adapted to the point group symmetry of the molecule, although optimal NOs are mostly not adapted to the symmetry, the Lagrangian contains all symmetry information, so the latter can be transferred to the COs after diagonalization of $\mathbf{\Lambda}$. It is worth to notice that the matrix of the Lagrange multipliers plays the role of the generalized Fock matrix in the NOFT.

Both $\{\varphi_p(\mathbf{r})\}$ and $\{\chi_p(\mathbf{r})\}$ sets of orbitals are pictures of the same solution; therefore, they complement each other. In the Sect. 3, the obtained results are discussed.

2.1 Orbital energies and ionization potentials

The Eq. (10), which now includes correlation effects, looks exactly the same as the total energy of an independent-particle system, hence λ_{pp} can be considered as a one-particle energy of the spatial orbital p. However, in contrast to the HF one-particle energies, $-\lambda_{pp}$ are not IPs of the molecule via the Koopmans' theorem [18]. The IPs in the NOFT [20, 38, 49] must be obtained from the extended Koopmans' theorem [4, 5, 33, 55]. The equation for the EKT may be derived by expressing the wavefunction of the

Fig. 1 PNOF5 valence natural and canonical orbitals for H$_2$O, along with their corresponding diagonal Lagrange multipliers in Hartrees, and diagonal elements of the 1-RDM, in *parenthesis*

Natural Orbital Representation

Canonical Orbital Representation

−0.0084 (0.0069)
−0.0179 (0.0182)

−0.0186 (0.0184)   −0.017 (0.0184)   −0.0085 (0.0085)   −0.0075 (0.0079)

−0.5057 (1.9915)

−0.5821 (1.9869)

−0.7167 (1.9816)

−0.7238 (1.9931)

−0.8538 (1.9818)

−1.3458 (1.9868)
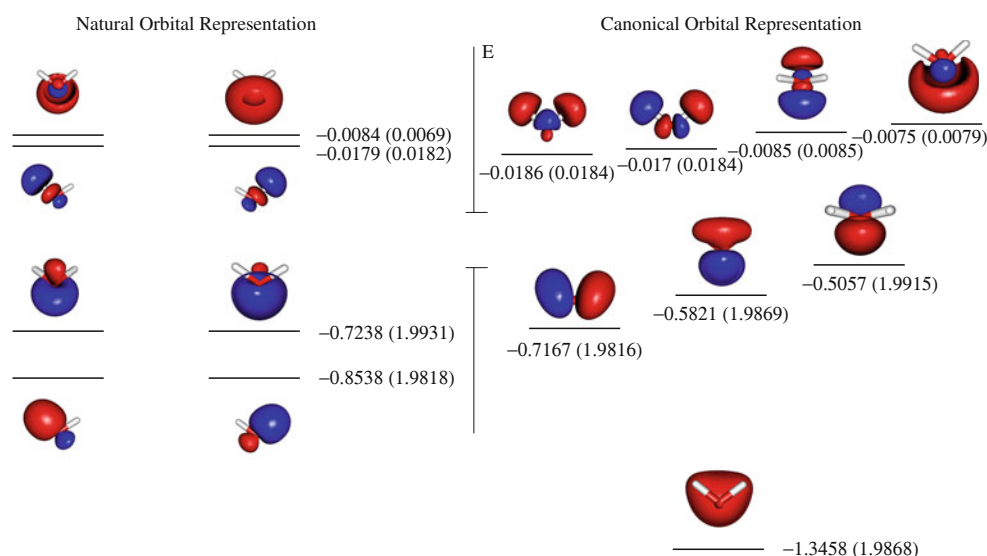
(N − 1)-electron system as the following linear combination

$$\left|\Psi^{N-1}\right\rangle = \sum_i C_i \hat{a}_i \left|\Psi^N\right\rangle \tag{13}$$

In Eq. (13), $\hat{a}_i$ is the annihilation operator for an electron in the spin-orbital $|\phi_i\rangle = |\varphi_p\rangle \otimes |\sigma\rangle (\sigma = \alpha, \beta)$, $|\Psi^N\rangle$ is the wavefunction of the N-electron system, $|\Psi^{N-1}\rangle$ is the wavefunction of the (N − 1)-electron system and $\{C_i\}$ are the set of coefficients to be determined. Optimizing the energy of the state Ψ^{N-1} with respect to the parameters $\{C_i\}$ and subtracting the energy of Ψ^N gives the EKT equations as a generalized eigenvalue problem,

$$\mathbf{FC} = \mathbf{\Gamma C}v \tag{14}$$

where v are the EKT IPs, and the transition matrix elements are given by

$$F_{ji} = \left\langle \Psi^N \left| \hat{a}_j^\dagger [\hat{\mathcal{H}}, \hat{a}_i] \right| \Psi^N \right\rangle \tag{15}$$

Eq. (14) can be transformed by a symmetric orthonormalization using the inverse square root of the 1-RDM. Hence, the diagonalization of the matrix $\mathbf{\Gamma}^{1/2}\mathbf{F\Gamma}^{-1/2}$ yields the IPs as eigenvalues. In a spin-restricted NOFT, it is not difficult to demonstrate that transition matrix elements are given by $-\lambda_{qp}$ [49]. Accordingly, the diagonalization of the matrix v with elements

$$v_{qp} = -\frac{\lambda_{qp}}{\sqrt{n_q n_p}} \tag{16}$$

affords the IPs in the NO representation. If the off-diagonal elements of the Lagrangian may be neglected, then from Eq. (16) follows that $-\lambda_{pp}/n_p$ will be good approximations for the ionization energies. Our calculations have shown that this rarely occurs. On the other hand, $-\mathbf{\Lambda}'$ corresponds

to the transition matrix $\mathbf{F}'$ in the CO representation, hence the diagonalization of the matrix v' with elements

$$v'_{qp} = -\sum_r \left(\Gamma'_{qr}\right)^{-1/2} \lambda'_{rr} \left(\Gamma'_{rp}\right)^{-1/2} \tag{17}$$

provides alternatively the IPs too. In many cases, we have found that the 1-RDM remains close to the diagonal form in the CO representation, so the values $-\lambda'_{pp}/\Gamma'_{pp}$ may be taken as ionization energies. We show below, in Sect. 4, that the one-particle energies associated with the COs may be considered as good estimations of the principal ionization potentials, for several molecules, via this formula.

3 PNOF5 orbitals

In general, PNOF5 yields localized orbitals in the NO representation, whereas it affords delocalized orbitals in the CO representation. The former ones arise directly from the energy minimization process. The COs are obtained from the diagonalization of the matrix of Lagrange multipliers after the NOs have been obtained. Here, both NO and CO representations of PNOF5 valence orbitals are given for a selected set of molecules, namely, H$_2$O, CH$_4$, (BH$_3$)$_2$, BrF$_5$ and C$_6$H$_6$. These molecules have been chosen to show the equivalency between both pictures of the orbitals and how these two pictures are connected. In all figures, the corresponding diagonal elements of the matrix of Lagrange multipliers and 1-RDM have been included. For the latter, twice of its values are reported, for example , the double of the occupancies in the case of the NO representation, and the double of the average number of particles in the orbital under consideration, for the CO representation.

Fig. 2 PNOF5 valence natural and canonical orbitals for CH_4, along with their corresponding diagonal Lagrange multipliers in Hartrees, and diagonal elements of the 1-RDM, in *parenthesis*
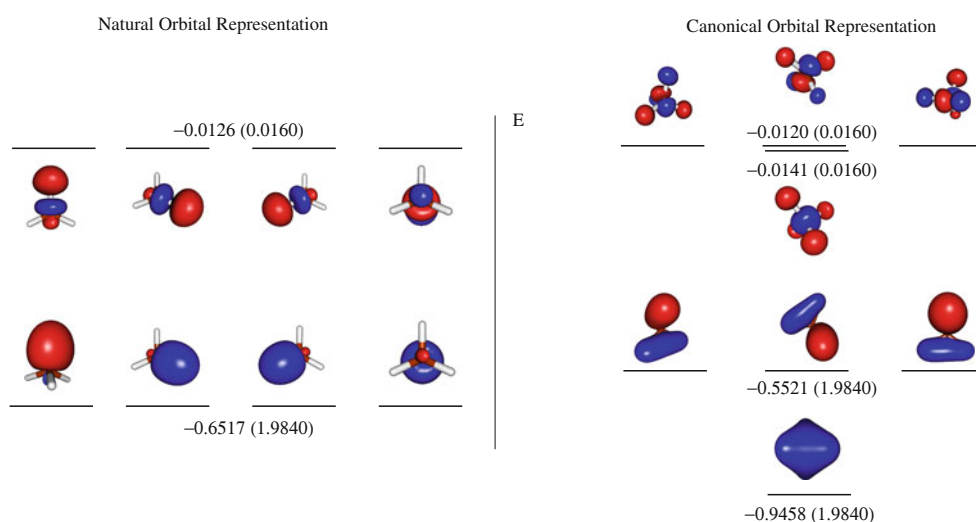
Fig. 3 PNOF5 valence natural and canonical orbitals for $(BH_3)_2$, along with their corresponding diagonal Lagrange multipliers in Hartrees, and diagonal elements of the 1-RDM, in *parenthesis*
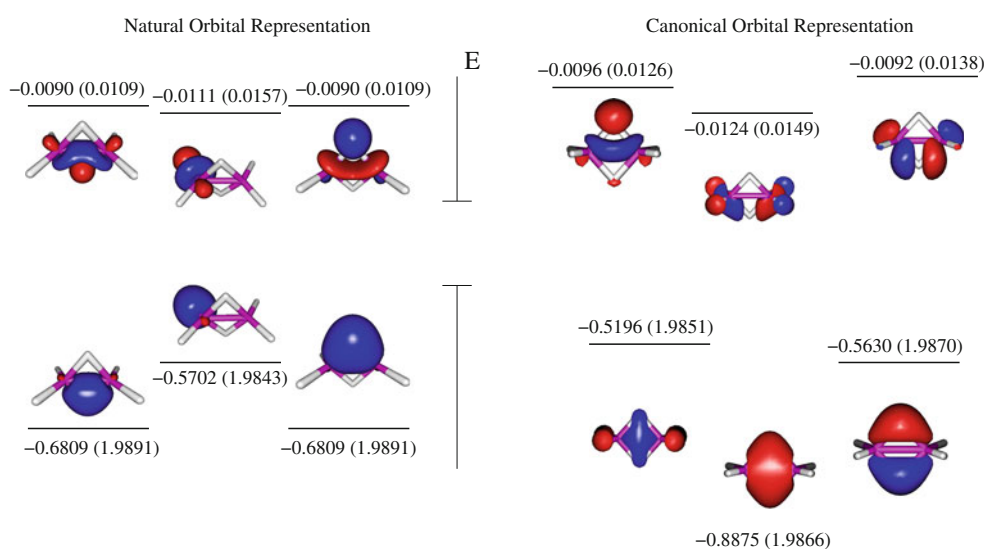
All calculations have been carried out at the experimental geometries [16], using the PNOFID code [39]. The correlation-consistent cc-pVDZ-contracted Gaussian basis sets [7, 57] have been employed. No important differences were observed for orbitals obtained with larger basis sets.

In Fig. 1, the PNOF5 valence natural and canonical orbitals calculated for the water molecule are shown. It is observed that NOs agree closely to the picture emerging from chemical bonding arguments: the O atom has sp^3 hybridization, two of these orbitals are used to bound to H atoms, leading to two degenerate oxygen-hydrogen σ bonds, and the remaining two are degenerated lone pairs. NO representation provides a theoretical basis to the VSEPR model. On the other hand, in the CO representation, the obtained orbitals are symmetry adapted and resemble those obtained by usual molecular orbital theories, for example, HF or DFT.

The valence natural and canonical orbitals of methane are depicted in Fig. 2. The NO representation describes the bonding picture in methane as four equivalent C–H bonds, resembling those that can be obtained with the VB method. Carbon is hybridized to form four sp^3-type orbitals. Each of such orbitals form a covalent bond with the 1s of one of the H atoms. The calculated orbital energies and occupation numbers are the same for these four orbitals.

On the contrary, the COs are symmetry adapted, and one can observe that the fourfold degeneracy is broken into one orbital of a_1 symmetry and threefold degenerate t_2 orbitals.

Focussing on the diagonal elements of the 1-RDM in both representations, we note that these values are close to 1 or 0. Moreover, the off-diagonal elements are exactly zero in the NO representation, and they can be neglected in the CO representation; ergo, the 1-RDM is practically idempotent in both representations. Because the

Fig. 4 PNOF5 valence natural and canonical orbitals for BrF$_5$, along with their corresponding diagonal Lagrange multipliers in Hartrees, and diagonal elements of the 1-RDM, in *parenthesis*

Natural Orbital Representation

Canonical Orbital Representation

−0.0182 (0.0171) −0.0196 (0.0184) −0.0182 (0.0171)

−0.0232 (0.0174) −0.0253 (0.0182) −0.0232 (0.0174)

−0.5452 (1.9852)

−0.6845 (1.9838) −0.6971 (1.9895)

−0.8689 (1.9881)

−0.9155 (1.9829) −1.0075 (1.9816) −0.9155 (1.9829)

−0.8823 (1.9886)

−1.1427 (1.9992)
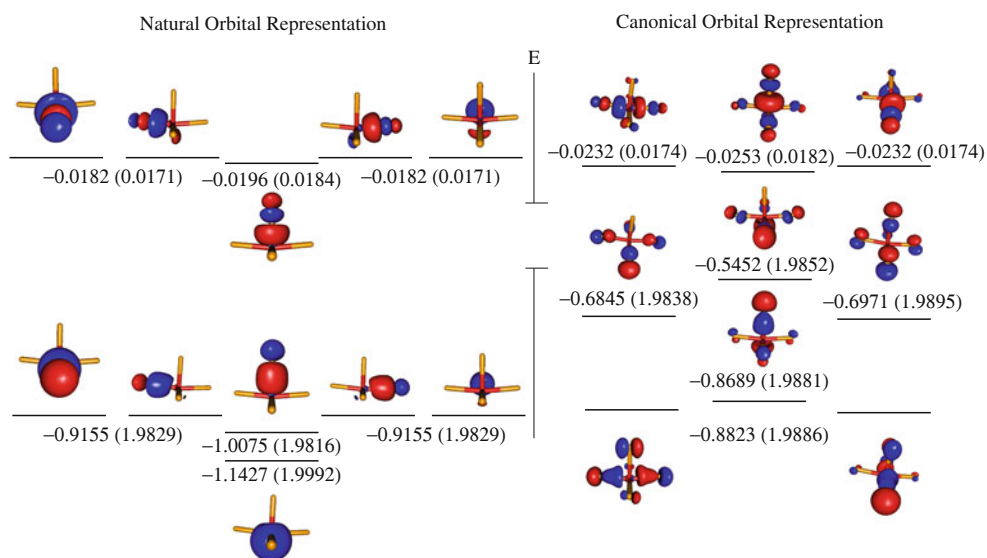
off-diagonal elements of the Lagrangian are not negligible in the NO representation, $-\lambda_{pp}$ cannot approximate the IP, and it is completely wrong to expect one valence ionization energy fourfold degenerate in methane. On the contrary, the obtained $-\lambda'_{pp}$ may be considered as good estimation for the IP. In [26], it was shown that PNOF5 is able to describe the two peaks of the vertical ionization spectra of methane via the EKT. The obtained here IPs for methane, by means of the negative value of the CO energies, are 15.02 and 25.74 eV, which are very close to the PNOF5-EKT values of 15.14 and 25.82 eV, and to the experimental IPs of 14.40 and 23.00 eV, respectively [26]. This is an example of how both one-electron pictures complement each other. It is evident that the NO representation agrees perfectly with the chemical bonding arguments, whereas the CO representation solves the problem raised for the ionization potentials.

Diborane $(BH_3)_2$ is an electron deficient molecule. The PNOF5 valence NO scheme predicts three-center two-electron (3c–2e) bonds, linking together both B atoms through intermediate H atoms, as can be seen on the left side of Fig. 3. Four degenerated B–H σ bonding orbitals are predicted to be coupled with their corresponding anti-bonding orbitals, while two degenerated B–H–B bonding orbitals are coupled with their corresponding antibonding orbitals. The CO scheme, depicted on the right side of Fig. 3, shows a similar delocalized picture as the standard molecular orbital calculations. According to this picture, the 3c–2e bonds are correctly described.

In Fig. 4, the PNOF5 orbitals calculated for BrF$_5$ are given. The NO representation depicts the bonding in this molecule as four degenerated Br–F σ bonds on the equatorial plane, and one Br–F σ bond on the axial axis. Each F atom has sp^3 hybridization, having the three lone pairs (not

Fig. 5 PNOF5 valence natural and canonical orbitals for C$_6$H$_6$, along with their corresponding diagonal Lagrange multipliers in Hartrees, and diagonal elements of the 1-RDM, in *parenthesis*


Natural Orbital Representation

Canonical Orbital Representation

−0.0155 (0.0413)

−0.0122 (0.0413)

−0.0169 (0.0427)

−0.3403 (1.9573)

−0.3938 (1.9587)

−0.5012 (1.9587)

Table 1 First vertical ionization potentials, in eV, obtained at the HF and PNOF5 levels of theory by means of the KT and EKT, respectively, along with the negative values of the corresponding orbital energies in the canonical orbital representation ($-\lambda'_{pp}$)

		HF-KT	PNOF5-EKT	$-\lambda'_{pp}$	EXP
N_2	σ_g	17.05 (1.45)	16.45 (0.85)	16.98 (1.38)	15.60
F_2	π_g	18.03 (2.16)	17.23 (1.36)	18.12 (2.25)	15.87
LiH	σ	8.17 (0.47)	7.53 (−0.17)	8.60 (0.90)	7.70
HF	π	17.12 (0.93)	16.76 (0.77)	17.55 (1.36)	16.19
HCl	π	12.82 (0.09)	12.63 (−0.06)	12.99 (0.22)	12.77
CO	σ	14.93 (0.92)	14.16 (0.15)	14.81 (0.80)	14.01
SiO	σ	11.78 (0.17)	11.82 (0.21)	11.98 (0.37)	11.61
H_2O	b_1	13.42 (0.64)	13.06 (0.28)	13.76 (0.98)	12.78
NH_3	a_1	11.41 (0.61)	11.05 (0.25)	11.67 (0.87)	10.80
H_2CO	b_2	11.84 (0.94)	11.74 (0.84)	12.24 (1.34)	10.90
C_6H_6	e_{1g}	9.05 (−0.20)	9.20 (−0.05)	9.26 (0.01)	9.25
CO_2	π_g	14.59 (0.81)	13.96 (0.18)	14.89 (1.11)	13.78
SO_2	a_1	13.24 (0.74)	13.00 (0.50)	13.47 (0.97)	12.50
BrF_5	a_1	14.46 (1.29)	13.62 (0.45)	14.84 (1.67)	13.17
ClF_3	b_1	14.52 (1.47)	13.53 (0.48)	14.66 (1.61)	13.05
CH_3CClO	a'	10.73 (−0.30)	11.13 (0.10)	12.65 (1.62)	11.03
HCOOH	a'	12.43 (0.93)	12.29 (0.79)	12.98 (1.48)	11.50
CH_3OCH_3	b_1	11.36 (1.26)	11.24 (1.14)	11.62 (1.52)	10.10
HOC–CHO	a_g	11.75 (1.15)	11.58 (0.98)	12.02 (1.42)	10.60
$HCONH_2$	a''	11.20 (1.04)	10.34 (0.18)	11.49 (1.33)	10.16
CH_3SH	a	9.56 (0.12)	9.33 (−0.11)	9.73 (0.29)	9.44

Differences between theoretical values and the experiment, in parenthesis. The cc-pVDZ basis sets have been employed

shown in the figure) as far apart as possible, in accordance to the VSEPR model. In principle, this NO representation of the PNOF5 orbitals provides a picture that could resemble that predicted by the molecular orbital theories. In the axial Br–F bonds, the quasilinear F–Br–F two bonds are constructed mainly by the same bromine p orbital. Consequently, these two bonds are "connected" by the same p orbital in the center.

A better agreement with the molecular orbital pictures is indeed obtained in the complementary CO representation. It may be observed that COs describe perfectly the three-center four-electron (3c–4e) bonds, where four electrons are delocalized along the quasilinear F–Br–F bonds. Furthermore, the obtained COs are symmetry-adapted.

Benzene can be considered as a model molecule for aromatic systems. It is well known that aromaticity can be described by both localized and delocalized orbitals. VB theory describes the π delocalization by combinations of structures containing localized π bonds between adjacent carbon atoms. On the other hand, HF approximation can predict the delocalization effects with only one Slater determinant, in accordance to the Huckel model. The six p orbitals of carbon atoms involved in the π system form six molecular orbitals, one with 0 nodes, 2 degenerate orbitals

with 2 nodes, 2 degenerate orbitals with 4 nodes and finally a totally antibonding orbital with six nodes.

In Fig. 5, the PNOF5 NOs and COs for benzene are shown. For the sake of clarity, only the orbitals involved in the π system are depicted. Focusing on the NO representation, three degenerate π orbitals are obtained, coupled with their corresponding antibonding orbitals. According to this picture, one could infer that the delocalization effects are not fully taken into account for benzene; however, there are significant values for the off-diagonal elements of the matrix of Lagrange multipliers that contain this information. The CO representation corroborates this hyphothesis showing the typical orbital picture. It should be mentioned that, in the NO representation, the remaining σ-type orbitals are localized C–C and C–H bonds, while in the CO representation, these σ-type orbitals are delocalized along the molecule. The obtained COs are symmetry adapted as in above described cases. Accordingly, PNOF5 can also handle aromatic systems.

4 Vertical ionization potentials

We have shown above, in Subsect. 2.1, that if the 1-RDM keeps close to the diagonal form in the CO representation, then the values $-\lambda'_{pp}/\Gamma'_{pp}$ may be considered as good estimations of the principal ionization potentials. Even more, if the 1-RDM is almost idempotent, the negative values of the CO energies may be taken as well. In this section, the calculated vertical ionization energies of an enlarged set of molecules are shown.

Table 1 lists the obtained vertical IPs calculated as $-\lambda'_{pp}$ for a selected set of molecules. For these systems, the 1-RDM is close to the corresponding idempotent matrix, with diagonal elements near to 1 or 0. For comparison, the ionization energies obtained at the HF level of theory, by means of KT, as well as the PNOF5 IPs via the EKT have been included.

We observe that the negative values of the corresponding orbital energies in the CO representation agree well with the experimental data. The better agreement between the HF IPs and the experiment is due to the partial cancellation of the electron correlation and orbital relaxation effects, an issue that has been long recognized in the literature. In our case, $-\lambda'_{pp}$ takes into account the electron correlation, but neglect the orbital relaxation in the $(N - 1)$-state, hence, the corresponding orbital energies in the canonical orbital representation tend to produce too positive IPs.

Table 2 collects a selected set of molecules with vertical IPs, calculated as $-\lambda'_{pp}$, that are smaller than the IPs obtained via EKT. In the CO representation, these molecules have 1-RDMs which could be considered rather

Table 2 First vertical ionization potentials, in eV, obtained at the HF and PNOF5 levels of theory by means of the KT and EKT, respectively, along with the negative values of the corresponding orbital energies in the canonical orbital representation ($-\lambda'_{pp}$ and $-\lambda'_{pp}/\Gamma'_{pp}$)

		HF-KT	PNOF5-EKT	$-\lambda'_{pp}/\Gamma'_{pp}$	$-\lambda'_{pp}$	EXP
Li$_2$	σ	4.91 (−0.23)	5.20 (0.06)	5.21 (0.07)	4.89 (0.25)	5.14
P$_2$	π_u	10.03 (−0.62)	10.57 (−0.08)	10.56 (−0.09)	10.05 (−0.60)	10.65
CH$_4$	t$_2$	14.78 (0.38)	15.14 (0.74)	15.14 (0.74)	15.02 (0.62)	14.40
C$_2$H$_2$	π_u	11.06 (−0.43)	11.61 (0.12)	11.60 (0.11)	11.26 (−0.23)	11.49
C$_2$H$_4$	1b$_3u$	10.14 (−0.37)	10.87 (0.36)	10.87 (0.36)	10.39 (−0.12)	10.51
HCN	π	13.37 (−0.24)	13.96 (0.35)	14.09 (0.48)	13.64 (0.03)	13.61
CH$_3$CN	a$'$	12.43 (−0.03)	12.97 (0.51)	13.13 (0.67)	12.77 (0.31)	12.46

Differences between theoretical values and the experiment, in parenthesis. The cc-pVDZ basis sets have been employed

diagonal instead of idempotent. Certainly, the off-diagonal elements of the CO 1-RDM can be neglected, and in most cases diagonal values Γ'_{pp} differ significantly from 1 and 0. In these particular systems, where both Γ' and Λ' can be considered diagonal matrices, the estimated IPs via EKT reduces to $-\lambda'_{pp}/\Gamma'_{pp}$ according to Eq. (17). We observe an outstanding agreement between the fourth and fifth columns of Table 2.

Exceptions are HCN and CH$_3$CN. These molecules have values different from 1 and 0 in the diagonal of the new 1-RDM Γ, and in addition, the off-diagonal elements cannot be neglected as in previous cases. Accordingly, the orbital energy in the CO representation yields an smaller estimation for the IP than EKT, but the corrected value obtained dividing $-\lambda'_{pp}$ by the corresponding diagonal element of Γ' is larger than it.

5 Conclusions

It has been shown that PNOF5 provides two complementary pictures of the electronic structure of molecules, namely, the NO and the CO representations. In the NO representation, the matrix of Lagrange multipliers is a symmetric non-diagonal matrix, whereas the 1-RDM is diagonal. Conversely, the matrix of Lagrange multipliers is transformed to yield a diagonal matrix in the CO representation, but the 1-RDM becomes a symmetric non-diagonal matrix. This transformation can be done only after solving the problem in the NO representation because Λ is symmetric only at the extremum. Hence, we are forced to obtain firstly the NOs which minimize the energy, and afterward, Λ is transformed from the NO representation, in which it is not diagonal, to the CO representation in which it is diagonal. Unfortunately, both matrices cannot be diagonalized simultaneously; however, NO and CO representations are unique one-particle pictures of the same solution, ergo, complement each other in the description of the electronic structure.

PNOF5 NOs are localized orbitals that nicely agree with the chemical intuition of chemical bonding, VB and VSEPR bonding pictures. On the contrary, PNOF5 COs are symmetry-adapted delocalized orbitals similar to those obtained by molecular orbital theories. The shape of COs obtained for diborane (BH$_3$)$_2$, bromine pentafluoride (BrF$_5$) and benzene (C$_6$H$_6$) supports the idea that PNOF5 is able to describe correctly the electronic structure of molecules containing three-center two-electron (3c–2e) bonds, three-center four-electron (3c–4e) bonds and the delocalization effects related to the aromaticity.

The new COs arises directly from the unitary transformation that diagonalizes the matrix of Lagrange multipliers of the NO representation. The values of this diagonal Lagrangian in the CO representation can be interpreted in the same vein as HF case, but including electron correlation effects, that is, the ionization energies are the negative of the corresponding orbital energies. We have shown theoretically and numerically that this approximation is valid when the obtained CO 1-RDM is close to the idempotent one. In some particular cases, where both the Lagrangian and 1-RDM can be considered diagonal matrices, the corrected value obtained dividing by the corresponding diagonal element of the 1-RDM yields practically same estimations as the EKT.

References

1. Bent HA (1961) An Appraisal of Valence-bond Structures and Hybridization in Compounds of the First-row elements. Chem Rev 61(3):275–311
2. Brueckner KA, Wada W (1955) Nuclear saturation and two-body: self-consistent solutions and the effects of the exclusion principle. Phys Rev 103:1008–1016

3. Coulson CA, Longuet-Higgins HC (1947) The electronic structure of conjugated systems. I. General theory. Proc R Soc Lond A191:39–60

4. Day OW, Smith DW, Garrod C (1974) A generalization of the hartree-fock one-particle potential. Int J Quantum Chem Symp 8(S8):501–509

5. Day OW, Smith DW, Morrison RC (1975) Extension of Koopmans' theorem. II. Accurate ionization energies from correlated wavefunctions for closed-shell atoms. J Chem Phys 62(1):115–119

6. Donnelly RA (1979) On fundamental difference between energy functionals based on first- and second-order density matrices. J Chem Phys 71(7):2874–2879

7. Dunning Jr. TH (1989) Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. J Chem Phys 90(2):1007–1023

8. Fock VA (1930) Näherungsmethode zur Lösung des quantenmechanischen Mehrkörper- problems. Z Phys 61(1–2):126–148

9. Gilbert TL (1975) Hohenberg-Kohn theorem for nonlocal external potentials. Phys Rev B 12(6):2111–2120

10. Gillespie RJ, Nyholm RS (1957) The valence-shell electron-pair repulsion theory. Quart Rev Chem Soc 11:339–80

11. Goedecker S, Umrigar CJ (2000) Natural Orbital Functional Theory. In: Cioslowski J (ed.) Many-electron densities and reduced density matrices, pp 165–181. Kluwer, New York

12. Hartree DR (1928) The wave mechanics of an atom with a non-coulomb central field. Part I. Theory and Methods. Proc. Camb Phil. Soc. 24(1):89–110

13. Heitler W, London F (1927) Wechselwirkung neutraler Atome und homöopolare Bindung nach der Quantenmechanik. Z Phys 44(6–7):455–472

14. Hund F (1928) Zur Deutung der Molekelspektren. IV. Z Phys 51(11–12):759–795

15. Jimenez-Hoyos CA, Henderson TM, Scuseria GE (2011) Generalized Hartree–Fock Description of Molecular Dissociation. J Chem Theory Comput 7:2667–2674

16. Johnson III RD (ed.) (2011) NIST computational chemistry comparison and benchmark database, NIST standard reference database Num 101, Release 15b. http://cccbdb.nist.gov/

17. Kohn W, Sham L (1965) Self-consistent equations including exchange and correlation effects. Phys Rev 140(4A):A1133–A1138

18. Koopmans TA (1934) Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms. Physica 1(1–6):104–113

19. Kutzelnigg W, Mukherjee D (1999) Cumulant expansion of the reduced density matrices. J Chem Phys 110(6):2800–2809

20. Leiva P, Piris M (2006) Calculation of vertical ionization potentials with the Piris natural orbital functional. J Mol Struct Theochem 770(1–3):45–49

21. Levy M (1979) Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. Proc Natl Acad Sci USA 76(12):6062–6065

22. Levy M (1987) Density matrices and density functionals. In: Erdahl R, Smith VHJ (eds.) Density matrices and density functionals. Reidel, Dordrecht, pp 479–498

23. Linderberg J, Öhrn Y (2004) Propagators in quantum chemistry, 2nd edn. Wiley, Hoboken

24. Lopez X, Ruiperez F, Piris M, Matxain JM, Matito E, Ugalde JM (2012) Performance of PNOF5 natural orbital functional for radical formation reactions: hydrogen atom abstraction and C–C and O–O homolytic bond cleavage in selected molecules. J Chem Theory Comput 8:2646–2652

25. Lowdin PO (1955) Quantum theory of many-particle systems. I. Phys Rev 97(6):1490–1508

26. Matxain JM, Piris M, Mercero JM, Lopez X, Ugalde JM (2012) sp3 Hybrid orbitals and ionization energies of methane from PNOF5. Chem Phys Lett 531:272–274

27. Matxain JM, Piris M, Ruipérez F, Lopez X, Ugalde JM (2011) Homolytic molecular dissociation in natural orbital functional theory. Phys Chem Chem Phys 13(45):20129–20135

28. Matxain JM, Piris M, Uranga J, Lopez X, Merino G, Ugalde JM (2012) Nature of chemical bonds by means of NOFT. Chem Phys Chem 13:2297–2303

29. Mazziotti DA (1998) Approximate solution for electron correlation through the use of Schwinger probes. Chem Phys Lett 289:419–427

30. Mazziotti DA (2007) Variational two-electron reduced-density-matrix theory. In: Mazziotti DA (ed.) Reduced-density-matrix mechanics: with applications to many-electron atoms and molecules, 1 edn., Chap 3. Wiley, Hoboken, pp 21–59

31. Mazziotti DA (2008) Parametrization of the two-electron reduced density matrix for its direct calculation without the many-electron wave function: Generalizations and applications. Phys Rev Lett 101:253002

32. Mazziotti DA (2012) Two-electron reduced density matrix as the basic variable in many-electron quantum chemistry and physics. Chem Rev 112(8):244–262

33. Morrell MM, Parr RG, Levy M (1975) Calculation of I.P. from density matrices and natural functions, and the long-range behavior of natural orbitals and electron density. J Chem Phys 62(2):549–554

34. Mulliken RS (1928) The assignment of quantum numbers for electrons in molecules. I. Phys Rev 32:186–222

35. Ortiz JV (1999) Toward an exact one-electron picture of chemical bonding. Adv Quantum Chem 35:33–52

36. Ortiz JV (2004) Brueckner orbitals, Dyson orbitals, and correlation potentials. Int J Quantum Chem 100(6):1131–1135

37. Parr R, Yang W (1989) Density-functional theory of atoms and molecules. Oxford University Press, Oxford

38. Pernal K, Cioslowski J (2005) Ionization potentials from the extended Koopmans' theorem applied to density matrix functional theory. Chem Phys Lett 412(1-3):71–75

39. Piris M PNOFID, http://www.ehu.es/mario.piris/#Software

40. Piris M (1999) A generalized self-consistent-field procedure in the improved BCS theory. J Math Chem 25:47–54

41. Piris M (2006) A new approach for the two-electron cumulant in natural orbital. Int J Quantum Chem 106:1093–1104

42. Piris M (2007) Natural orbital functional theory. In: Mazziotti DA (ed.) Reduced-density-matrix mechanics: with applications to many-electron atoms and molecules, chap 14. Wiley, Hoboken, pp 387–427

43. Piris M (2012) A natural orbital functional based on an explicit approach of the two-electron cumulant. Int J Quantum Chem. doi: 10.1002/qua.24020

44. Piris M, Lopez X, Ruipérez F, Matxain JM, Ugalde JM (2011) A natural orbital functional for multiconfigurational states. J Chem Phys 134(16):164102

45. Piris M, Lopez X, Ugalde JM (2007) Dispersion interactions within the Piris natural orbital functional theory: the helium dimer. J Chem Phys 126(21):214103

46. Piris M, Matxain JM, Lopez X, Ugalde JM (2009) Spin conserving natural orbital functional theory. J Chem Phys 131:021,102

47. Piris M, Matxain JM, Lopez X, Ugalde JM (2010) Communication: The role of the positivity N-representability conditions in natural orbital functional theory. J Chem Phys 133:111101

48. Piris M, Matxain JM, Lopez X, Ugalde JM (2010) Communications: accurate description of atoms and molecules by natural orbital functional theory. J Chem Phys 132:031103

49. Piris M, Matxain JM, Lopez X, Ugalde JM (2012) The extended Koopmans' theorem: vertical ionization potentials from natural orbital functional theory. J Chem Phys 136(17):174116

50. Piris M, Montero LA, Cruz N (1997) The BCS approach to electron correlation in the density matrix formalism. J Chem Phys 107(1):180–187

51. Piris M, Otto P (2003) One-particle density matrix functional for correlation in molecular systems. Int J Quantum Chem 94(6): 317–323

52. Piris M, Otto P (2005) Natural orbital functional for correlation in polymers. Int J Quantum Chem 102(1):90–97

53. Piris M, Ugalde JM (2009) Iterative diagonalization for orbital optimization in natural orbital functional theory. J Comput Chem 30:2078–2086

54. Sand AM, Schwerdtfeger Ca, Mazziotti DA (2012) Strongly correlated barriers to rotation from parametric two-electron reduced-density-matrix methods in application to the isomerization of diazene. J Chem Phys 136(3):034112

55. Smith DW, Day OW (1975) Extension of Koopmans theorem. I. Derivation. J Chem Phys 62(1):113–114

56. Valone SM (1980) Consequences of extending 1 matrix energy functionals pure-state representable to all ensemble representable 1 matrices. J Chem Phys 73(3):1344–1349

57. Woon D, Dunning Jr. TH (1993) Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. J Chem Phys 98(2):1358–1371

58. Wu W, Su P, Shaik S, Hiberty PC (2011) Classical valence bond approach by modern method. Chem Rev 111(11):7557–7593

REGULAR ARTICLE

MS-CASPT2 study of the low-lying electronic excited states of di-thiosubstituted formic acid dimers

R. Verzeni · O. Mó · A. Cimas · I. Corral · M. Yáñez

Abstract The suitability of di-thiosubstituted derivatives of formic acid dimer, both in hydroxyl and carbonyl position, as possible hydrogen-bonded electron transfer linkers in a hypothetical donor–acceptor dyad for photovoltaic cells and artificial photosynthesis reactors has been studied from a theoretical point of view. To this purpose, the valence singlet electronic excited states of the four possible di-thiosubstituted isomers have been characterized through multi-state complete active space second-order perturbation theory (MS-CASPT2). These hydrogen-bonded systems present electronic spectra consisting of $n\pi^*$ and $\pi\pi^*$ excitations, both intra- and intermonomer. The eventual comparison of the calculated spectroscopic characteristics of the isolated hydrogen-bonded linkers with the experimental spectrum of the chromophore in a donor–acceptor dyad could allow establishing whether the linker would compete with the electron donor in the photon absorption process. Additionally, the analysis of the structural changes undergone by these species upon electronic excitation to the S_1 would allow determining whether the population of this state of the linker upon UV–vis light absorption could compromise the formation of the charge transfer complex, key in the performance of photovoltaic devices.

Keywords Hydrogen-bonded linkers · Formic acid dimer · Di-thiosubstituted derivatives · MS-CASPT2 · Solar cells · Charge transfer · Donor–acceptor dyad

R. Verzeni · O. Mó · I. Corral (✉) · M. Yáñez
Departamento de Química, Facultad de Ciencias, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain
e-mail: ines.corral@uam.es

A. Cimas
Centro de Investigação em Química, Department of Chemistry and Biochemistry, Faculty of Science, University of Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

1 Introduction

In the last decades, the increasing demand of new materials and electronic nanodevices for high-performance organic solar cells [1–5] has motivated a growing interest on solar energy convertors based on the same key process, a photoexcitation leading to a charge separation [1, 5]. The simplest version of such photovoltaic devices is a donor–acceptor dyad at least composed by an electron donor chromophore, an electron acceptor and a linker that controls their distance and electronic interactions. In an organic photovoltaic cell, the dyad is connected to two electrodes, which convey the two-formed charges in a circuit, producing electrical current.

The process of charge separation starts when a photon hits the chromophore, generating an exciton. Under normal conditions, the exciton does not travel long distances, and the chromophore remains in the so-called excited state that usually decays rapidly, relaxing either radiatively or thermally. Nevertheless, under certain circumstances, the above relaxation mechanisms compete with other processes such as charge transfer (CT), for instance in those cases where the chromophore is connected to a strong electron acceptor. In these situations, the exciton could be forced to dissociate driving the system into a CT complex,
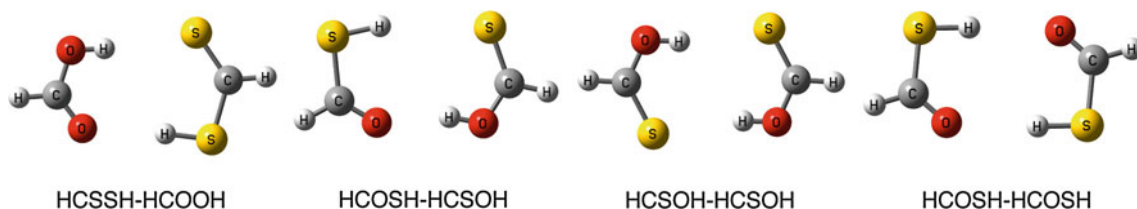
Fig. 1 Formic acid dimer di-thiosubstituted derivatives

where the electron has been transferred to one of the acceptor's lowest lying unoccupied MOs (LUMOs) and the hole still remains on one of the donor highest lying occupied MOs (HOMOs). This hybrid state, lying at the interface between donor and acceptor moieties, governs in solar cells both the voltage-dependent photocurrent as well as the open circuit voltage [5]. The efficiency and the rate of this final step depend on macroscopic values as charge carriers' average mobility, materials' average dielectric constant, and distance below which the CT complex polarons thermally relax, [5, 6] but also on microscopic aspects such as coulombic interactions caused by molecules orientation toward the heterojunction [5, 7].

Inspired by the efficiency of biological photosynthesis, the number of biomimetic studies on the control of electron transfer reactions through a network of hydrogen bonds (HBs) has significantly increased [4, 8]. Indeed, it has been shown that hydrogen-bonded donor–acceptor assemblies ensure more efficient electronic communication than comparable σ- or π-bonding networks [9, 10].

However, in the event that the electronic absorption spectra of the hydrogen-bonded connector and that of the electron donor overlap, the efficiency of the entire device can be seriously compromised, directly, due to the reduction in potentially absorbable photons by the electron donor, preventing the formation of excitons, or indirectly, since geometric changes in the hydrogen-bonded linker, due to its electronic excitation, could interfere with the formation of the CT complex and/or its dissociation, ultimately provoking a collapse of the hydrogen-bonded photovoltaic device. Therefore, the spectroscopic characterization of the connector is of fundamental importance for the successful design of a charge separation reaction center. The aim of this paper is to present such a spectroscopic characterization for the linkers built-up from the substitution of two oxygen atoms in formic acid dimer and leading to the four hydrogen-bonded complexes: HCSSH–HCOOH, HCOSH–HCSOH, HCSOH–HCSOH, and HCOSH–HCOSH, shown in Fig. 1. Similar species based on double HB interactions between a carboxylate anion and an amidinium cation have been used in dyads with photovoltaic activity. The comparison of the results for the di-thiosubstituted dimers with those obtained for formic acid dimer and its mono-thiosubstituted derivatives will allow

Table 1 Summary of formic acid dimer and its monosubstituted derivatives vertical energies and oscillator strengths for the electronic transitions absorbing below 10 eV, analyzed in Ref. [27]

	HCOOH–HCOOH		HCOSH–HCOOH		HCSOH–HCOOH	
	ΔE/eV	f	ΔE/eV	f	ΔE/eV	f
Intra-monomer						
$n_{c=o}\pi^*_{c=o}$	6.13	0.000	6.17	0.001	6.27	0.001
$n_{c=x}\pi^*_{c=x}$(S)	6.21	0.001	4.94	0.000	3.82	0.000
$\pi_{c=o}\pi^*_{c=o}$	8.25	0.655	8.31	0.456	8.33	0.471
$\pi_{c=x}\pi^*_{c=x}$(S)	7.56[a]	0.000	5.98	0.396	5.58	0.477
Inter-monomer						
$n_{c=x}\pi^*_{c=x}$	9.13	0.002	8.71	0.000	8.92	0.001
$n_{c=x}\pi^*_{c=x}$(S)	9.17	0.000	10.63	0.000	8.62	0.002
$\pi_{c=x}\pi^*_{c=x}$	9.80	0.000	9.59	0.004	9.84	0.035
$\pi_{c=x}\pi^*_{c=x}$(S)	9.93	0.005	9.71	0.011	8.73	0.003

(S) only applies to HCOSH–HCOOH and HCSOH–HCOOH dimers and denotes, in the intramonomer section, transitions occurring in the thiosubstituted monomer and, in the intermonomer section, transitions where the electron is promoted from an orbital from the thiosubstituted monomer

[a] Excitation energy underestimated due to the very low reference weight in the CASPT2 calculation. This transition is expected to peak at 8.25 eV [27]

determining the effect that a second sulfur atom has in the UV absorption spectra of these systems and whether their spectroscopic properties could broaden the range of chromophores with which the new linkers can be used.

Finally, the characterization of the structure and bonding of the first electronic excited state in these systems will allow estimating the impact that, in the CT complex of the photovoltaic device, has the change in the structure of the linker in the hypothetical case, these electronic states are populated by UV–vis photons.

To our knowledge, no experimental spectroscopic studies on formic acid dimer di-thiosubstituted derivatives have been reported to date (Table 1).

2 Computational details

The ground state structures of the four studied dimers were optimized using the B3LYP [11, 12] functional in
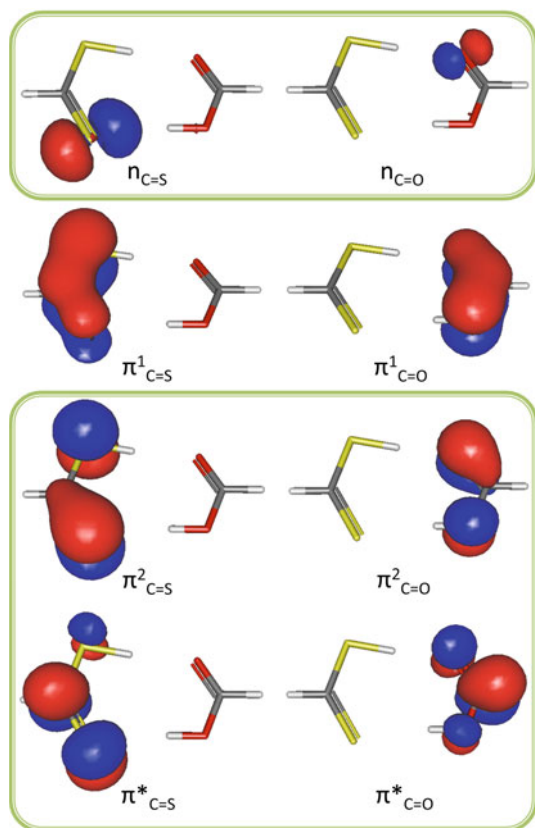
Fig. 2 Exemplary SA-CASSCF valence molecular orbitals used in the calculation of A′ and A″ electronic transitions of the HCSSH–HCOOH dimer. Sulfur and oxygen atoms are represented in *yellow* and *red*, respectively. The n superindex of π^n orbitals denotes the total number of nodes of the MO. Similar orbitals or linear combinations of them were obtained in the excited state calculations of the rest of dithiosubstituted dimers. Framed in *green*, the orbitals included in the active space employed in the geometry optimizations of the S_1 states

Table 2 MS-CASPT2 excitation energies ΔE (eV, nm), configuration interaction coefficients (CI) and oscillator strengths (f), for the valence lower-lying excited states of the HCSSH–HCOOH dimer

State symmetry	MS-CASPT2//CASSCF(12,14)/ANO-L				
	Main configuration	CI coefficient	ΔE/eV	ΔE/nm	f
$1^1A''$ (S_1)[a]	$n_{c=s} \rightarrow \pi^*_{c=s}$	−0.94	2.79	444	0.4901
$2^1A'$ (S_2)[a]	$\pi^2_{c=s} \rightarrow \pi^*_{c=s}$	0.91	4.50	276	0.0864
$2^1A''$ (S_3)[a]	$n_{c=o} \rightarrow \pi^*_{c=o}$	−0.85	5.93	209	0.1802
$3^1A'$ (S_4)[a]	$\pi^1_{c=s} \rightarrow \pi^*_{c=s}$	0.76	6.84	181	0.1704
	DE	−0.50			
$3^1A''$ (S_5)[b]	$n_{c=o} \rightarrow \pi^*_{c=s}$	0.87	7.34	169	0.0211
$4^1A''$ (S_6)[b]	$n_{c=s} \rightarrow \pi^*_{c=o}$	−0.84	7.42	167	0.0006
$4^1A'$ (S_7)[b]	$\pi^2_{c=o} \rightarrow \pi^*_{c=s}$	−0.71	7.86	158	0.0289
	$\pi^2_{c=o} \rightarrow \pi^*_{c=o}$	−0.55			
$5^1A'$ (S_8)[a]	$\pi^2_{c=o} \rightarrow \pi^*_{c=o}$	−0.56	8.36	148	0.0000
	$\pi^2_{c=s} \rightarrow \pi^*_{c=o}$	0.53			
	$\pi^2_{c=o} \rightarrow \pi^*_{c=s}$	0.45			
$6^1A'$ (S_9)[a]	DE	−0.79	8.56	145	0.0009
$7^1A'$ (S_{10})[b]	$\pi^2_{c=s} \rightarrow \pi^*_{c=o}$	−0.70	9.28	134	0.0002
	$\pi^2_{c=o} \rightarrow \pi^*_{c=o}$	−0.43			
$8^1A'$ (S_{11})[a]	DE	−0.69	9.46	131	0.0008
	$\pi^1_{c=s} \rightarrow \pi^*_{c=s}$	0.49			

Ground state total energy: −1023.972452 Eh

DE double excitations

[a] Intramonomer excitations, [b] intermonomer excitations

conjunction with the Pople basis set 6-311++G(3df,2p) [13], recommended in previous works for the optimization of species containing sulfur atoms [14–16]. Tables 2, 3, 4, and 5 collect the valence vertical electronic excitation energies and oscillator strengths, calculated with the CASSCF method [17] along with a triple-zeta contracted set of natural orbitals ANO-L (C,O[4s3p2d]/S[5s4p2d]/ H[3s2p]) [18]. Other excitations (double or Rydberg transitions) above the highest valence states were not included in the tables for simplicity.

Four active spaces of the sizes (12,14), (12,11), (12,14), and (12,10) were employed to model the UV absorption spectra of HCSSH–HCOOH, HCOSH–HCSOH, HCSOH–HCSOH, and HCOSH–HCOSH. All the above active spaces have in common two lone pairs, $n_{c=x}$ (X=O, S), lying at the dimer plane and sitting at the carbonyl/thiocarbonyl position, and 3 pairs of frontier π orbitals, including a bonding $\left(\pi^1_{c=x}\right)$, a non-bonding $\left(\pi^2_{c=x}\right)$ and an antibonding $\left(\pi^*_{c=x}\right)$ orbital, where in the first two cases the

superindex represents the number of nodes of the corresponding MO, (See Fig. 2).

All the remaining orbitals until completing the final sizes correspond to virtual orbitals included to avoid intruder states. All CASSCF calculations were carried out using the state average formalism, under C_s symmetry constraints, entailing the number of roots necessary for describing all valence excited states, that is, 8 roots of both A′ and A″ symmetries for HCSSH–HCOOH, 6 and 3 roots of A′ and A″ symmetries for HCOSH–HCSOH, 8 and 6 roots of A′ and A″ symmetries for HCSOH–HCSOH, and 5 and 3 roots of A′ and A″ symmetries for HCOSH–HCOSH. Dynamic correlation was incorporated via a second-order perturbation theory treatment of the CASSCF wave function through the MS-CASPT2 method [19]. A real level shift [20] parameter of 0.3 was employed in order to remove further problems connected to intruder states.

Excited state geometry optimizations were performed at the CASSCF/aug-cc-pVTZ [21–23] level of theory, using the (8,6) active space defined in Fig. 2. The same protocol was employed for optimizing the ground states in order to analyze the structural changes undergone by these species

Table 3 MS-CASPT2 excitation energies ΔE (eV, nm), configuration interaction coefficients (CI), and oscillator strengths (f), for the valence lower-lying excited states of the HCOSH–HCSOH dimer

| State Symmetry | MS-CASPT2//CASSCF(12,11)/ANO-L | | | | |
	Main configuration	CI coefficient	ΔE/eV	ΔE/nm	f
$1^1A''$ (S$_1$)	$n_{c=s} \rightarrow \pi^*_{c=s}$	−0.94	3.62	343	0.2494
$2^1A''$ (S$_2$)	$n_{c=o} \rightarrow \pi^*_{c=o}$	−0.91	4.95	251	0.6738
$2^1A'$ (S$_3$)	$\pi^2_{c=s} \rightarrow \pi^*_{(c=o+c=s)}$	0.74	5.34	232	0.0051
	$\pi^2_{c=s} \rightarrow \pi^*_{(c=o-c=s)}$	−0.48			
$3^1A'$ (S$_4$)	$\pi^2_{c=o} \rightarrow \pi^*_{(c=o+c=s)}$	−0.71	5.77	215	0.0066
	$\pi^2_{c=o} \rightarrow \pi^*_{(c=o-c=s)}$	−0.47			
$3^1A''$ (S$_5$)	$n_{c=s} \rightarrow \pi^*_{c=o}$	−0.94	6.28	198	0.0000
$4^1A'$ (S$_6$)	$\pi^2_{c=s} \rightarrow \pi^*_{(c=o-c=s)}$	−0.70	7.42	167	0.0000
	$\pi^2_{c=s} \rightarrow \pi^*_{(c=o+c=s)}$	−0.55			
$5^1A'$ (S$_7$)	$\pi^2_{c=o} \rightarrow \pi^*_{(c=o-c=s)}$	−0.70	7.84	158	0.0000
	$\pi^2_{c=o} \rightarrow \pi^*_{(c=o+c=s)}$	0.54			

Ground state total energy: −1023.966949 Eh

(+) and (−) stand for positive and negative linear combinations of the orbitals of the different dimer moieties

upon electronic excitation. Finally, the bonding in the excited states will be investigated in the frame of the AIM theory of Bader [24], through the mapping of the topology of the electron density and localization of bond and ring critical points.

Ground and excited state geometry optimizations were carried out using Gaussian 09 [25], while vertical energies and oscillator strengths were calculated with MOLCAS 7.4 [26] suite of programs.

3 Results and discussion

3.1 UV vertical excitation energies

Before start discussing the calculated UV vertical absorption energies of di-thiosubstituted dimers, we shall briefly revisit the main characteristics of the electronic spectra of formic acid dimer and its monosubstituted derivatives, analyzed in detail in ref. [27]. For the sake of clarity, a summary of the vertical energies and oscillator strengths for the electronic transitions absorbing below 10 eV for these species has been included in Table 1. The calculated UV spectra of these three species are characterized by an imaginary division line which separates intramonomer transitions, taking place in a single moiety, and intermonomer or CT transitions involving the two constituting monomers.

Table 4 MS-CASPT2 excitation energies ΔE (eV, nm), configuration interaction coefficients (CI) and oscillator strengths (f), for the low-lying excited states of the HCSOH–HCSOH dimer

| State symmetry | MS-CASPT2//CASSCF(12,14)/ANO-L | | | | |
	Main configuration	CI coefficient	ΔE/eV	ΔE/nm	f
$1^1A''$ (S$_1$)	$n_{(A-B)} \rightarrow \pi^*_{(A+B)}$	−0.79	3.56	348	0.0119
	$n_{(A+B)} \rightarrow \pi^*_{(A-B)}$	−0.47			
$2^1A''$ (S$_2$)	$n_{(A-B)} \rightarrow \pi^*_{(A-B)}$	0.70	3.67	337	0.6679
	$n_{(A+B)} \rightarrow \pi^*_{(A+B)}$	0.58			
$2^1A'$ (S$_3$)	$\pi^2_{(A)} \rightarrow \pi^*_{(A+B)}$	−0.62	5.03	246	0.0109
	$\pi^2_{(B)} \rightarrow \pi^*_{(A-B)}$	0.35			
$3^1A'$ (S$_4$)	$\pi^2_{(B)} \rightarrow \pi^*_{(A+B)}$	−0.65	5.44	228	0.0009
	$\pi^2_{(A)} \rightarrow \pi^*_{(A-B)}$	0.47			
	$\pi^2_{(B)} \rightarrow \pi^*_{(A-B)}$	0.32			
$3^1A''$ (S$_5$)	DE		6.01	206	0.0000
$4^1A''$ (S$_6$)	DE		6.19	200	0.0000
$5^1A''$ (S$_7$)	$n_{(A-B)} \rightarrow \pi^*_{(A-B)}$	0.59	6.90	180	0.0000
	$n_{(A+B)} \rightarrow \pi^*_{(A+B)}$	−0.69			
$6^1A''$ (S$_8$)	$n_{(A-B)} \rightarrow \pi^*_{(A+B)}$	0.48	7.06	176	0.0000
	$n_{(A+B)} \rightarrow \pi^*_{(A-B)}$	−0.76			
$4^1A'$ (S$_9$)	DE		7.08	175	0.0000
$5^1A'$ (S$_{10}$)	DE		7.09	175	0.0008
$6^1A'$ (S$_{11}$)	$\pi^2_{(B)} \rightarrow \pi^*_{A+B)}$	0.37	7.21	172	0.0000
	$\pi^2_{(B)} \rightarrow \pi^*_{(A-B)}$	0.66			
$7^1A'$ (S$_{12}$)	$\pi^2_{(A)} \rightarrow \pi^*_{(A+B)}$	0.42	7.36	168	0.0000
	$\pi^2_{(B)} \rightarrow \pi^*_{(A+B)}$	−0.37			
	$\pi^2_{(A)} \rightarrow \pi^*_{(A-B)}$	−0.68			

Ground state total energy: −1023.974778 Eh

DE double excitations

(A) and (B) denote the two identical dimer moieties. (+) and (−) stand for positive and negative linear combinations of the orbitals of the different dimer moieties

Within both regions, electronic transitions were found to be arranged according the two established patterns: $n\pi^*$ − $n\pi^*$ − $\pi\pi^*$ − $\pi\pi^*$ characteristic of formic acid dimer spectrum and $n\pi^*$ − $\pi\pi^*$ − $n\pi^*$ − $\pi\pi^*$ governing the absorption spectrum of the mono-thiosubstituted dimers. This change of pattern has its origin in a red-shift of the electronic vertical excitation energies which unevenly affects $n\pi^*$ and $\pi\pi^*$ transitions upon sulfur substitution, which was also found to affect the oscillator strengths of these excitations.

Table 5 MS-CASPT2 excitation energies ΔE (eV, nm), configuration interaction coefficients (CI), and oscillator strengths (f), for the low-lying excited states of the HCOSH–HCOSH dimer

State symmetry	MS-CASPT2//CASSCF(12,10)/ANO-L				
	Main configuration	CI coefficient	ΔE/eV	ΔE/nm	f
$1\,^1A''$ (S$_1$)	$n_{(B)} \rightarrow \pi^*_{(B)}$	−0.95	4.89	253	0.0000
$2\,^1A''$ (S$_2$)	$n_{(A)} \rightarrow \pi^*_{(A)}$	−0.92	4.96	250	0.8380
$2\,^1A'$ (S$_3$)	$\pi^2_{(A-B)} \rightarrow \pi^*_{(A+B)}$	−0.73	5.73	216	0.0190
	$\pi^2_{(A+B)} \rightarrow \pi^*_{(A-B)}$	−0.58			
$3\,^1A'$ (S$_4$)	$\pi^2_{(A-B)} \rightarrow \pi^*_{(A-B)}$	0.64	6.01	206	0.0000
	$\pi^2_{(A+B)} \rightarrow \pi^*_{(A+B)}$	0.71			
$4\,^1A'$ (S$_5$)	$\pi^2_{(A-B)} \rightarrow \pi^*_{(A-B)}$	−0.71	8.06	154	0.0000
	$\pi^2_{(A+B)} \rightarrow \pi^*_{(A+B)}$	0.63			
$5\,^1A'$ (S$_6$)	$\pi^2_{(A-B)} \rightarrow \pi^*_{(A+B)}$	0.59	8.08	153	0.0000
	$\pi^2_{(A+B)} \rightarrow \pi^*_{(A-B)}$	−0.74			
$3\,^1A''$ (S$_7$)	$n_{(B)} \rightarrow \pi^*_{(A)}$	0.94	8.27	150	0.0001

Ground state total energy: −1023.962576 Eh

(A) and (B) denote the two identical dimer moieties. (+) and (−) stand for positive and negative linear combinations of the orbitals of the different dimer moieties

3.1.1 HCSSH–HCOOH

The MS-CASPT2//SA-CASSCF vertical absorption energies for HCSSH–HCOOH dimer are collected in Table 2. According to these results, its UV absorption spectrum would consist of a very intense band peaking at 2.8 eV (S$_1$) followed by three less intense absorptions in the region of 4.5–7 eV (S$_2$–S$_4$). The S$_1$ and S$_3$ states present an $n\pi^*$ character and are localized in the thiosubstituted monomer and in the formic acid moiety, respectively, while the S$_2$ and S$_4$ have a predominant $\pi\pi^*$ nature.

A careful comparison between Tables 1 and 2 reveals that di-thio-substitution within the same monomer breaks the clear division between intra- and intermonomer excitations characteristic of the unsubstituted and monosubstituted dimers. In addition, no clear pattern among $n\pi^*$ and $\pi\pi^*$ transitions is observed in the electronic spectrum summarized in Table 2 in contrast to HCOOH–HCOOH, HCSOH–HCOOH or HCOSH–HCOOH [27].

There are, however, some absorptions common to the electronic spectra of these four species, such as the intramonomer transitions taking place in the common moiety to the four dimers, that is, the HCOOH monomer. Specifically, the $n_{c=o} \rightarrow \pi^*_{c=o}$ and $\pi_{c=o} \rightarrow \pi^*_{c=o}$ excitations, peaking around 6.1 and 8.3 eV for the unsubstituted and monosubstituted dimers [27] (See Table 1), were calculated at 5.93

and 8.36 eV for HCSSH–HCOOH (See Table 2). The small energy differences registered for the above transitions can be either attributed to the use of different basis sets in both works or alternatively to the different chemical environments surrounding the fragment where the excitations occur. Interestingly, the oscillator strength of these transitions is also affected by the introduction of a second sulfur atom in the molecule. Thus, the intensity of the $n\pi^*$ transition becomes 100 times stronger in the HCSSH–HCOOH dimer, while the $\pi\pi^*$ transition goes from being one of the most intense absorptions in HCOOH–HCOOH and its mono-thiosubstituted derivatives to a dark state in HCSSH–HCOOH.

In general, double thio-substitution within the same monomer affects in a larger extent the rest of the transitions. Specifically, intramonomer excitations occurring in the sulfur-substituted monomer and CT transitions shift to lower energies as compared to HCSOH–HCOOH or HCOSH–HCOOH, being this shift slightly larger when the second sulfur substitution takes place in the carbonyl position.

Finally, it is interesting to mention that simultaneous substitution of the hydroxyl and carbonyl oxygens by sulfur in the same monomer greatly stabilize the $\pi^1_{c=s} \rightarrow \pi^*_{c=s}$ transition. As a consequence, while this transition is observed at 6.84 eV for the HCSSH–HCOOH complex, it was not found among the transitions absorbing below 10 eV in formic acid dimer or its mono-thiosubstituted derivatives.

3.1.2 HCOSH–HCSOH

Table 3 summarizes the MS-CASPT2-calculated valence vertical transition energies and relative intensities for formic acid dimer di-thiosubstituted in the carbonyl and the hydroxyl positions of different monomers. The $n_{c=o} \rightarrow \pi^*_{c=s}$ transition is not reported because it was omitted from the calculation due to the huge number of double excitations which precede it and that complicated the perturbation treatment.

Similarly to HCSSH–HCOOH, the electronic absorption spectrum of HCOSH–HCSOH is governed by $n\pi^*$ intramonomer excitations, which carry the largest oscillator strengths, see Table 3. A closer look to this table reveals that intra- and intermonomer excitations localize into two different regions of the spectrum, reminding unsubstituted and monosubstituted dimers spectra, although in this case, intra- and intermonomer $\pi\pi^*$ excitations are difficult to distinguish since the π^* orbitals involved in these transitions spread over the whole molecule.

Also for this species, we find some absorptions peaking at the same wavelengths as in the spectra of monosubstituted

derivatives. In particular, this applies to $n_{c=s} \rightarrow \pi^*_{c=s}$ and $n_{c=o} \rightarrow \pi^*_{c=o}$ transitions peaking at 3.62 and 4.95 eV, also present in HCSOH–HCOOH and HCOSH–HCOOH dimers, respectively. This also applies to $\pi\pi^*$ S_3 and S_4 states, which appear red shifted by 0.2 eV, most probably due to the π^* orbital mixing mentioned above. Red-shifts between 1.3 and 2.8 eV were also registered for the transitions S_5–S_7 taking the thiohydroxyl and thiocarbonyl-substituted derivatives as a reference, leading fortuitously to the recovery of the $n\pi^* - n\pi^* - \pi\pi^* - \pi\pi^*$ ordering characteristic of formic acid dimer.

3.1.3 HCSOH–HCSOH

An inspection of Table 4, containing the vertical excitation energies and oscillator strengths for HCSOH–HCSOH,

reveals that both the lone pairs and π^* orbitals describing the lower-lying excitations in this dimer appear as linear combinations of the former orbitals sitting on the two monomers, preventing the discrimination between intra-monomer and CT transitions, similarly to what observed for the dimer described in Sect. 3.1.2. Also in this case, the pattern $n\pi^* - n\pi^* - \pi\pi^* - \pi\pi^*$ was found to govern the two regions in which can be divided the spectrum, typically associated to intra- and intermonomer regions. The organization of the transitions according to this pattern is attributed, in this case, to the high symmetry of this dimer, C_{2h}, which leads to the degeneration or near degeneration (recall that all the calculations were performed under C_s symmetry constraints) between pairs of transitions with the same character. The first pairs of $n\pi^*$ and $\pi\pi^*$ transitions peak at ca. 3.6 and 5.4 eV very close

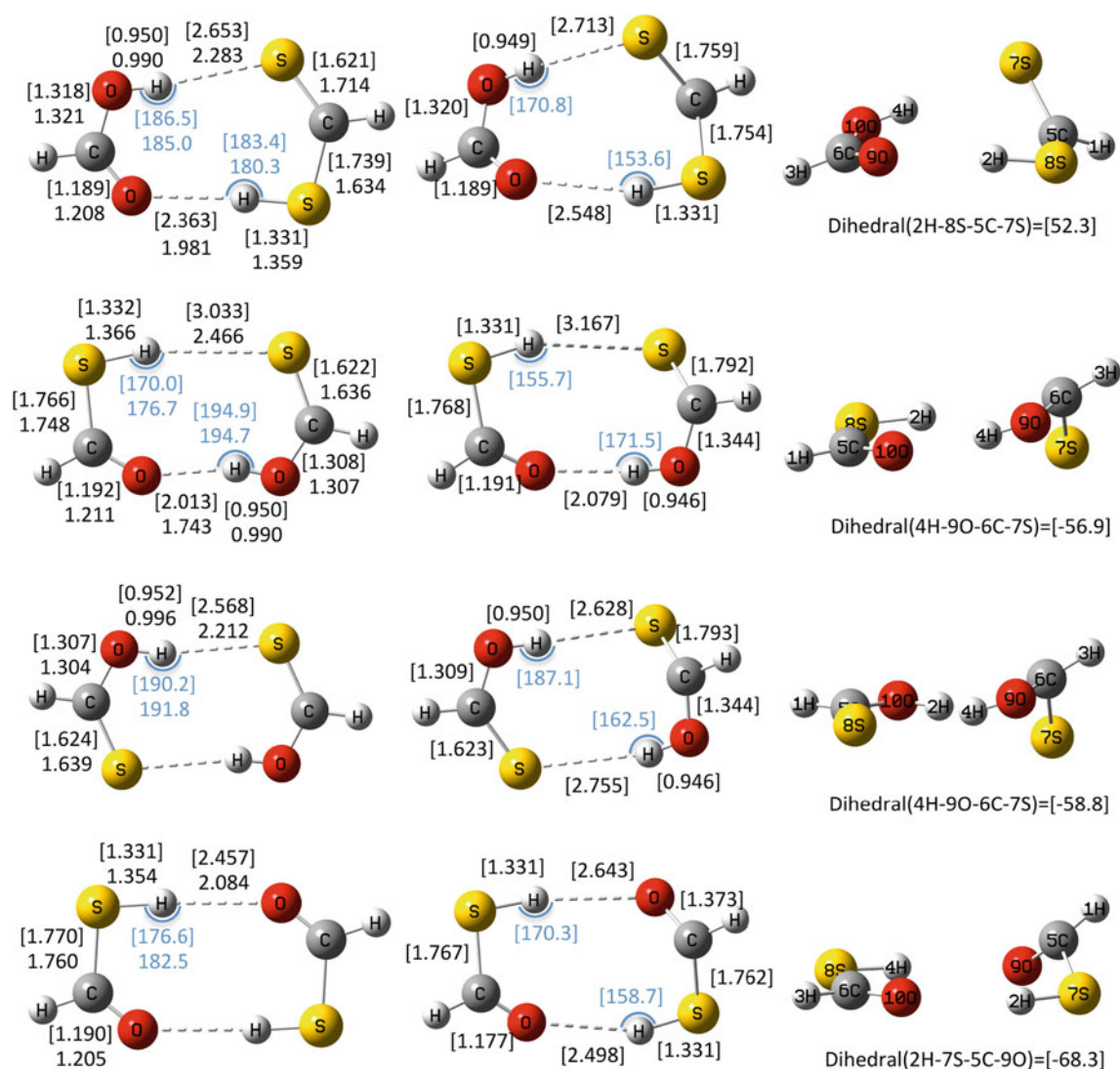**Fig. 3** Optimized geometries for the ground (S_0, first column) and the first excited (S_1, second and third columns) states of di-thiosubstituted derivatives of formic acid dimer. Values in normal font correspond to B3LYP/6-311++G(3df,2p) optimized parameters and CASSCF(8,6)/aug-cc-pVTZ optimized parameters are in *square brackets*. Bond distances are given in Angstrom and angles in degrees
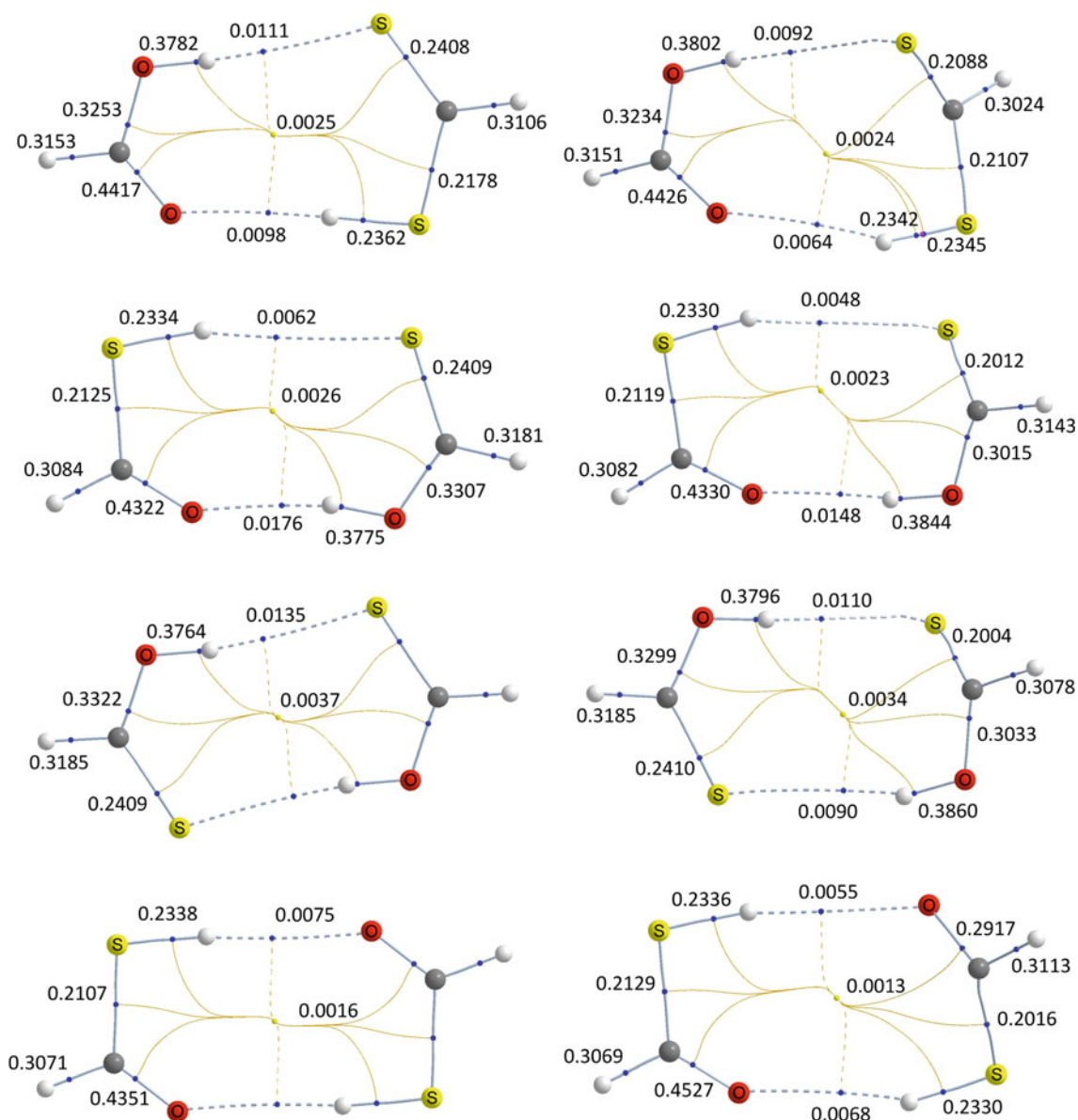
Fig. 4 Molecular graphs for the ground (*left*) and first excited state minima (*right*) of HCSSH–HCOOH, HCOSH–HCSOH, HCSOH–HCSOH, and HCOSH–HCOSH. Electron densities at the BCPs (*blue*) and RCPs (*yellow*) are in a.u

to the wavelengths where the same electronic states absorb in the dimer monosubstituted in the carbonyl position, pointing to the very little influence of the nature of the monomer where the excitation does not take place. As expected, this is not the case of S_7–S_8 $n\pi^*$ and S_{11}–S_{12} $\pi\pi^*$ transitions which experience red-shifts that can amount up to 2.5 eV.

Also interesting, the comparison of the present spectrum with that calculated for the di-thiosubstituted dimer in the hydroxyl and carbonyl position in Table 3 allows establishing the effect of the position of di-thio-substitution—carbonyl or hydroxyl—on the electronic energies. Thus, from the generalized red-shift of the electronic spectrum of

HCSOH–HCSOH compared with that of HCOSH–HCSOH, we conclude that a second oxygen-by-sulfur substitution occurring in carbonyl position has a larger effect on the electronic energies.

Introducing a second sulfur atom in carbonyl position has also a great impact in the oscillator strengths and therefore in the shape of the electronic spectrum. In this respect, the intramonomer $\pi\pi^*$ transitions, which were predicted to contribute to the greatest extent to the UV spectrum of the HCSOH–HCOOH dimer, become darker states in this dimer, while the opposite is observed for intramonomer $n\pi^*$ excitations which become the brightest transitions in the di-thiosubstituted monomer.

3.1.4 HCOSH–HCOSH

Finally, this section analyzes the calculated UV absorptions for the di-thiosubstituted dimer in the hydroxyl positions.

This species only differs from the dimer discussed in the previous section, HCSOH–HCSOH, in a double proton transfer, being the former species 3.8 kcal/mol more stable. At 298 K, this process involves an energy barrier of 5.5 kcal/mol that would prevent interconversion between both species.

Due to its C_{2h} symmetry, this dimer shares some of the features already described for the electronic spectrum of the thiosubstituted dimer in both carbonyl positions. As for HCSOH–HCSOH, the lowest lying $n\pi*$ and $\pi\pi*$ transitions in this dimer appear grouped in pairs of doubly degenerate excited states absorbing at the same energy as their monosubstituted analog HCOSH–HCOOH, whereas the remaining excitations are shifted to lower energies with respect to the same species. Contrary to the expectations, however, these doubly degenerated transitions do not follow the $n\pi* - n\pi* - \pi\pi* - \pi\pi*$ sequence observed for the other C_{2h} system. Indeed, the only calculated CT $n\pi*$ transition, S_7, (the other was omitted from the multiconfigurational calculations to avoid problems in the perturbation treatment) appears above the least stable $\pi\pi*$ transitions, breaking the pattern.

Finally, the comparison of this spectrum with those of the other two di-thiosubstituted dimers in different monomers allows confirming our conclusions on the larger effect which has the position of thio-substitution on the shift of the electronic energies. In fact, the electronic spectrum for this species is significantly shifted to higher energies than that calculated for the hydroxyl and carbonyl substituted dimer, which in turn is blue shifted compared to that of HCSOH–HCSOH.

Neither the position nor the dimer in which second thiosubstitution occurs seem to have any effect on the oscillator strength since as for the other species studied here the most intense absorption corresponds to an intramonomer $n\pi*$ transition.

At this point, it is worth stressing that HCOSH–HCOSH is the only dimer among all the studied not absorbing above 280 nm. This, coupled to the fact that interconversion between HCOSH–HCOSH and HCSOH–HCSOH is not likely to occur, turns this dimer into a potential linker to be used in photovoltaic devices working in the visible/near UV regime.

3.2 Geometries and bonding of the first excited states

Considering that a large change in the geometry of the hydrogen-bonded linker upon excitation might seriously affect the efficiency of the photovoltaic device, hindering the formation of the CT complex, we have optimized the first excited state S_1 of all the dimers. Recall that, according to the results discussed in the previous sections, this state corresponds to the brightest transition in HCSSH–HCOOH and has a non-negligible oscillator strength for HCOSH–HCSOH. Although for the two other dimers, HCSOH–HCSOH and HCOSH–HCOSH, the S_1 does not correspond to the brightest state, due to their symmetry, analogous geometries as the ones optimized here are expected for the states carrying the largest oscillator strengths, that is, S_2, which involve similar excitations.

Figure 3 collects the CASSCF geometries for the ground and first excited states of all the dimers considered in this work. The most remarkable difference between these pairs of structures is the out of plane deviation of the carbonyl/thiocarbonyl group of the monomer where the excitation takes place. This loss of planarity of the molecule is accompanied by a rearrangement of the bond distances and bond angles. Specifically, the population of the $\pi*$ orbital after the promotion of an electron from the lone pair sitting in the same moiety results in a stretching by 0.01–0.04 Å of the C–XH (X=O, S) bond distance and in a considerably larger elongation of the C=X (X=O, S) bond which in average amounts to 0.17 Å.

This excitation also affects the two HBs that hold together the two monomers, which significantly weaken in the excited state, but has no influence in the geometry of the moiety acting as a spectator during the excitation.

These changes in the geometries lead to a reorganization of the electron density of these species, see Fig. 4. Thus, on going from the S_0 to the S_1, we register an average decrease in the electron density which amounts to 0.037, 0.143, 0.008, and 0.029 e.a.u^{-3} for the C=S, C=O, C–SH, and C–OH bonds, respectively.

These results are useful in providing the trends which should be expected in the geometry changes upon excitation. It is important, however, not to forget the well-known poor performance of CASSCF for the description of hydrogen-bonded structures. Hence, the CASSCF geometries and electron densities of Figs. 3 and 4 are only qualitatively interesting. A quantitatively meaningful description requires a method capable of accurately characterizing excited states and at the same time including dynamic correlation.

4 Concluding remarks

This paper reports on the calculated valence excited states of the four di-thiosubstituted formic acid dimer derivatives which differ in the position (carbonyl and hydroxyl) and subunit which occupy the two sulfur atoms.

Due to the dimeric structure of these species, their absorption spectrum can be divided in two regions,

Fig. 5 **a** Simulated absorption spectra based on MS-CASPT2 vertical excitation energies and oscillator strengths for HCOOH–HCOOH (*purple line*), HCOSH–HCOOH (*orange line*), HCSOH–HCOOH (*black line*) (ref. [27]). **b** Simulated absorption spectra based on MS-CASPT2 vertical excitation energies and oscillator strengths for HCOSH–HCOSH (*pink line*), HCOSH–HCSOH (*light blue line*), HCSOH–HCSOH (*green line*), and HCSSH–HCOOH (*red line*)
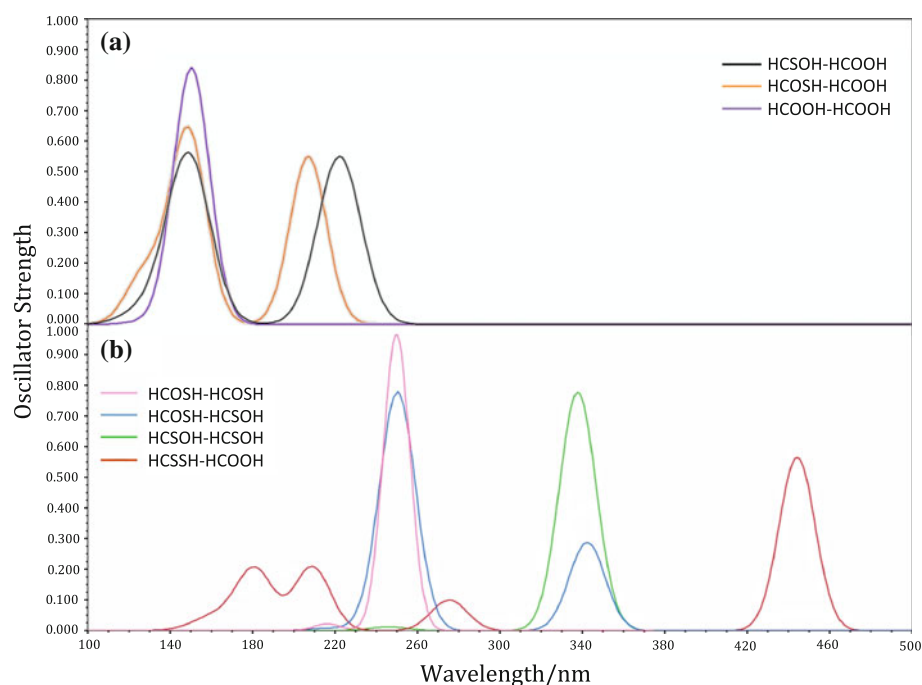
typically ascribed to intra- and intermonomer or charge transfer absorptions. Except in the case of HCSSH–HCOOH where intra- and intermonomer transitions appear mixed, for the rest of the dimers excited states appear organized into two different regions of the spectrum, separated by an energy gap comprised between ca. 0.5 and 2.0 eV.

Interestingly, all the dimers studied here present one or several $n\pi^*$ and $\pi\pi^*$ excitations peaking at the same wavelength as in the mono-substituted dimers with which they share the moiety where the excitation takes place.

This evidences the very little effect the electronic structure of the monomer acting as a spectator has in the intramonomer transitions. Obviously, this does not hold for the highest lying $n\pi^*$ and $\pi\pi^*$ states, typically of CT character, where the nature of the two moieties involved in the excitations determines the position of the absorption. In general, the above transitions shift further to the red upon introducing a second sulfur atom, especially if the second thio-substitution takes place in the carbonyl position. This is consistent with the fact that the calculated transitions for HCSOH–HCSOH appear at lower energies compared with HCOSH–HCSOH, whose spectrum is in turn red-shifted with respect to HCOSH–HCOSH. Finally, these shifts, which affect irregularly the different electronic transitions, lead to the organization of the $n\pi^*$ and $\pi\pi^*$ according to different patterns. Thus, an $n\pi^* - n\pi^* - \pi\pi^* - \pi\pi^*$ pattern similar to the one found for formic acid dimer was registered for HCOSH–HCSOH, HCSOH–HCSOH, and the lowest energy region of the HCOSH–HCOSH spectrum. In contrast, no apparent ordering among the $n\pi^*$ and

$\pi\pi^*$ electronic transitions was found in HCSSH–HCOOH, and an unexpected pattern was observed for the highest energy region of the HCOSH–HCOSH spectrum, where $\pi\pi^*$ excitations appear slightly stabilized with respect to the $n\pi^*$ transition.

Similarly to formic acid dimer and its mono-thiosubstituted derivatives, theory predicts transitions absorbing in the lower energy segment in which can be divided the spectrum to contribute to a larger extent to the UV absorption spectrum. However, whereas in the case of formic acid dimer and its mono-thiosubstituted derivatives, the largest oscillator strength is calculated for $\pi\pi^*$ excitations, introducing a second sulfur atom in the molecule renders these transitions dark and augments significantly the intensity of the $n\pi^*$ transitions which dominate the spectrum of di-thiosubstituted derivatives.

In sum, except for HCOSH–HCOSH which is transparent in the visible, UVA and UVB regions, caution should be exercised when using the remaining dithiosubstituted dimers as hydrogen-bonded linkers in photovoltaic devices with chromophores absorbing above 280 nm, see Fig. 5. Moreover, the corresponding electronic excitations could lead to the out of plane displacement of the carbonyl/thiocarbonyl moiety where the excitation takes place, potentially compromising the formation of the charge transfer complex.

References

1. Balzani V, Credi A, Venturi M (2008) Chemsuschem 1:26
2. D'Souza F, Ito O (2012) Chem Soc Rev 41:86
3. Beckers EHA, Chen Z, Meskers SCJ, Jonkheijm P, Schenning APHJ, Li X-Q, Osswald P, Wuerthner F, Janssen RAJJ (2006) Phys Chem B 110:16967
4. Sanchez L, Sierra M, Martin N, Myles AJ, Dale TJ, Rebek J, Seitz W, Guldi DM (2006) Angew Chem Int Ed 45:4637
5. Deibel C, Strobel T, Dyakonov V (2010) Adv Mater 22:4097
6. Fuoss M, Accascina F (1959) Electrolytic conductance. Interscience, New York
7. Sreearunothai P, Morteani AC, Avilov I, Cornil J, Beljonne D, Friend RH, Phillips RT, Silva C, Herz LM (2006) Phys Rev Lett 96:117403
8. Segura M, Sanchez L, de Mendoza J, Martin N, Guldi DMJ (2003) Am Chem Soc 125:15093
9. Sessler JL, Wang B, Harriman AJ (1993) Am Chem Soc 115:10418
10. Yu M-L, Wang S-M, Feng K, Khoury T, Crossley MJ, Fan Y, Zhang J-P, Tung C-H, Wu L-ZJ (2011) Phys Chem C 115:23634
11. Becke ADJ (1993) Chem Phys 98:5648
12. Lee CT, Yang WT, Parr RG (1988) Phys Rev B 37:785
13. Hehre WJ, Radom L, Schleyer PvR, Pople JA (1986) Ab initio molecular orbital theory. Wiley Interscience, New York
14. Jonas V, Frenking G (1991) Chem Phys Lett 177:175
15. Timoshkin A, Frenking GJ (2000) Chem Phys 113:8430
16. González L, Mó O, Yáñez M (1996) Chem Phys Lett 263:407
17. Roos BO, Taylor PR, Siegbahn PEM (1980) Chem Phys 48:157
18. Widmark PO, Malmqvist PÅ, Roos BO (1990) Theor Chim Acta 77:291
19. Finley J, Malmqvist PA, Roos BO, Serrano-Andres L (1998) Chem Phys Lett 288:299
20. Roos BO, Andersson K (1995) Chem Phys Lett 245:215
21. Woon DE, Dunning TH Jr (1993) J Chem Phys 98:1358
22. Dunning TH Jr (1007) J Chem Phys 1989:90
23. Kendall RA, Dunning TH Jr, Harrison RJJ (1992) Chem Phys 96:6796
24. Bader RFW (1990) Atoms in molecules: a quantum theory. Clarendon Press, Oxford
25. Frisch MJ et al (2009) Gaussian 09, Revision A.1. Gaussian, Inc., Wallingford
26. Aquilante F et al (2010) J Comput Chem 31:224
27. Cimas A, Mó O, Yáñez M, Martin N, Corral I (2010) Phys Chem Chem Phys 12:13037

REGULAR ARTICLE

Electronic structure studies of diradicals derived from *Closo*-Carboranes

Josep M. Oliva · Diego R. Alcoba · Luis Lain · Alicia Torre

Abstract Electronic structure computations have been performed on diradical systems composed of two carborane radicals $CB_{11}H_{12}^{\cdot}$ connected through acetylene, ethylene and ethane bridge units, leading, respectively, to a *linear* and two *trans* structures. Each cage possesses one unpaired electron and the total system can thus be coupled to a singlet or a triplet state. Numerical determinations using the spin-projected method with a hybrid B3LYP functional show that these compounds have singlet ground states with low singlet–triplet energy gaps of 0.004 eV (acetylene bridge), 0.080 eV (ethylene bridge) and 0.0005 eV (ethane bridge). Spin population analyses point out a left/right localized spin distribution in the spin-projected wave function. The possibility of mapping these results onto a Heisenberg spin Hamiltonian is considered, in order to predict low-lying excited states in extended carborane chains.

J. M. Oliva (✉)
Instituto de Química-Física "Rocasolano", Consejo Superior de Investigaciones Científicas, 28006 Madrid, Spain
e-mail: j.m.oliva@iqfr.csic.es

D. R. Alcoba
Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, 1428 Buenos Aires, Argentina

D. R. Alcoba
Instituto de Física de Buenos Aires, Consejo Nacional de Investigaciones Científicas y Técnicas, Ciudad Universitaria, 1428 Buenos Aires, Argentina

L. Lain · A. Torre
Departamento de Química Física. Facultad de Ciencia y Tecnología, Universidad del País Vasco, Apdo. 644, 48080 Bilbao, Spain

1 Introduction

Polyhedral boron chemistry embraces fields of research in inorganic chemistry at both molecular and solid-state levels and, in combination with organic moieties and metals [1], provides applications of interest in material sciences [2], design of pharmacophores [3] and medicine [4]. Thus, a wide variety of molecular and solid-state architectures have been synthesized since the first days of the synthesis of the *closo*-borane cages, particularly with the well-known icosahedral *closo*-borane anion $B_{12}H_{12}^{2-}$ [5]. One of the most challenging aspects of polyhedral boron chemistry is the study of low-lying excited states, as opposed to the (more) well-known excited state chemistry of carbon-derived compounds in organic chemistry—see for instance references [6, 7]. When boron cage atoms are substituted by carbon atoms in the icosahedral $B_{12}H_{12}^{2-}$ dianion, one then obtains the clusters from the series $C_nB_{12-n}H_{12}^{(n-2)}$, the so-called carboranes [8]. In the last few years, we have been interested in the electronic structure of ground states and excited states derived from polyhedral borane and carborane molecules, as isolated units [9–11], or connected in linear [12, 13] and triangular configurations [14, 15]. The versatile combination of charge and spin in these systems offers the possibility of tuning the properties of molecular architectures based on heteroborane cage units.

The $CB_{11}(CH_3)_{12}^{\cdot}$ radical [16]—with one unpaired electron—can be connected in 1D, 2D and 3D architectural

constructions thus involving a polyradical system with a determined number of unpaired electrons. So far to our knowledge, only the diradicals $\cdot(CH_3B)_{11}C–C\equiv C–C(BCH_3)_{11}\cdot$ and trans-$\cdot(CH_3B)_{11}C–CH=CH–C(BCH_3)_{11}\cdot$ have been synthesized [17]. Recently, we have studied the electronic structure of the simplified diradical $\cdot(HB)_{11}C–C\equiv C–C(BH)_{11}\cdot$ [12] using high-level quantum chemical models—CASPT2 [18]—for the calibration of the spin-projected method with the hybrid functional UB3LYP. These studies have shown that the ground state in the diradical is of singlet state nature, having a low-lying triplet state 0.005 eV (CASPT2) higher in energy (k_BT at room temperature is 0.025 eV/0.6 kcal/mol). This energy difference corresponds to the microwave region of the electromagnetic spectrum and therefore one could in principle populate selectively the triplet state using microwave photons, provided that intersystem crossing and spin–orbit interactions are significant. This energy difference also corresponds to rotational modes in gas-phase molecules, molecular motions in liquids and phonons in solids [19].

Let us now consider two icosahedral closo-carborane $CB_{11}H_{12}\cdot$ radicals, each of them with one unpaired electron ($S = \frac{1}{2}$), that might be connected in para position through the carbon atom of the cage with an acetylene, ethylene or ethane bridge unit. The three resulting structures are depicted in Fig. 1.

We will proceed to study the electronic structure of these compounds applying spin-partitioning techniques: Which are the (estimated) singlet–triplet energy gaps—or the corresponding spin–spin coupling constants, J's—in these diradicals? How are the electronic structures of the triplet states as compared to the spin-projected symmetry states? These are the questions we would like to answer in this work.

2 Methodology

All computations in this work have been carried out at the (U)B3LYP/6-31 + G(d) level of theory with the suite of programs Gaussian [20]. Geometry optimizations have been performed for the triplet states and spin-projected states, corresponding all structures to energy minima. The spin-projected method for two electrons was developed in Ref. [21], which reports a spin-unrestricted wave function, $\Psi_{unr,S}$, with both singlet and triplet components:

$$\Psi_{unr,S} = a \cdot \Psi_S + b \cdot \Psi_T; \quad a^2 + b^2 = 1. \tag{1}$$

The wave functions Ψ_S and Ψ_T are "pure" spin states with $S = 0$ (singlet) and $S = 1$ (triplet), respectively. One can then show that
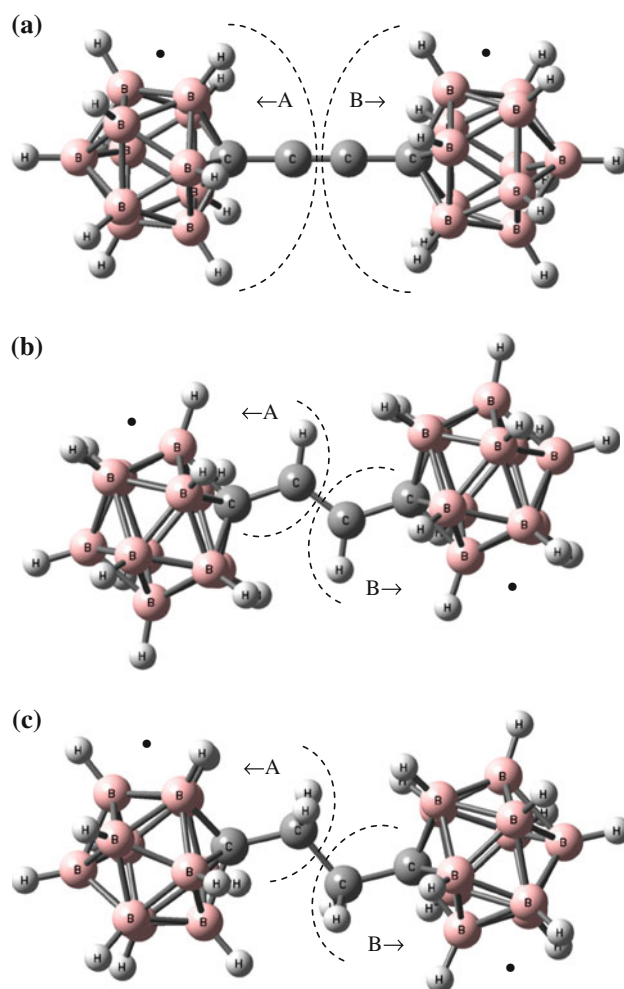
Fig. 1 The three diradicals described in this work: Two closo-carborane radicals $CB_{11}H_{12}\cdot$ connected through the cage carbon atom with **a** an acetylene bridge unit, **b** an ethylene bridge unit and **c** an ethane bridge unit. The dots represent unpaired electrons. The dashed curves divide the molecule into two fragments, denoted as A and B

$$b^2 = \frac{1}{2}\left\langle \Psi_{unr,S}|\widehat{\mathbf{S}}^2|\Psi_{unr,S}\right\rangle \tag{2}$$

and therefore the singlet–triplet energy gap—ΔE_{ST}—can be estimated as:

$$\Delta E_{ST} = \left(E_{unr,S} - E_T\right)/\left(1 - b^2\right) \tag{3}$$

On the other hand, the phenomenological Heisenberg spin Hamiltonian—Eq. 4—predicts the energy of the different spin states of a many-electron system, provided the spin degrees of freedom are independent from the electron (orbital) degrees of freedom

$$\widehat{H} = -2\sum_{A > B} J_{AB}\widehat{\mathbf{S}}_A \cdot \widehat{\mathbf{S}}_B \tag{4}$$

where $\hat{\mathbf{S}}_A$ and $\hat{\mathbf{S}}_B$ are the local spin operators associated with the neighbor molecular units A and B, respectively, which may be atoms or molecular fragments. J_{AB} are the

corresponding Heisenberg coupling constants, which may be related directly to a singlet–triplet energy difference for a two-electron system. Thus, the eigenspectrum of $\hat{H}$ connects experimental and theoretical studies of magnetism in molecular systems.

If we divide a cluster, molecule, etc., into different fragments A, B, …, the information on the spin attributed to these fragments may be obtained from the expectation values of the local spin operators $\left\langle \widehat{\mathbf{S}}_A^2 \right\rangle$ and $\widehat{\mathbf{S}}_A \cdot \widehat{\mathbf{S}}_B$ [22, 23]

$$\left\langle \widehat{\mathbf{S}}^2 \right\rangle = \sum_A \sum_B \left\langle \widehat{\mathbf{S}}_A \cdot \widehat{\mathbf{S}}_B \right\rangle \tag{5}$$

The one-center local spin $\left\langle \widehat{\mathbf{S}}_A^2 \right\rangle$ allows one to determine the spin state of an atom or group of atoms in a molecule or cluster, while the spin correlation between fragments A and B is described by the expectation value $\left\langle \widehat{\mathbf{S}}_A \cdot \widehat{\mathbf{S}}_B \right\rangle$. This value provides an important tool for linking experimental results interpreted in terms of the Heisenberg spin Hamiltonian to quantum chemical calculations, as mentioned above. We will consider the general algebraic expression for $\left\langle \widehat{\mathbf{S}}_A \cdot \widehat{\mathbf{S}}_B \right\rangle$ reported in Refs. [24, 25], which for the case of a Slater determinant wave function has the form

$$\left\langle \widehat{\mathbf{S}}_A \cdot \widehat{\mathbf{S}}_B \right\rangle = \frac{1}{4} \sum_{\mu \in A} \sum_{\nu \in B} (\mathbf{P}^s \mathbf{S})_{\mu\mu} \cdot (\mathbf{P}^s \mathbf{S})_{\nu\nu}$$
$$+ \delta_{AB} \frac{1}{2} \sum_{\mu \in A} \sum_{\nu \in B} (\mathbf{P}^s \mathbf{S})_{\mu\nu} \cdot (\mathbf{P}^s \mathbf{S})_{\nu\mu} \tag{6}$$

where μ, ν,… are the atomic functions used, $\mathbf{P}^s = \mathbf{P}^\alpha - \mathbf{P}^\beta$ the spin density matrix and $\mathbf{S}$ the overlap matrix. In this equation, the sums are restricted to the atomic orbitals assigned to the corresponding fragment.

Several procedures have been proposed for the calculation of the coupling constants J_{AB}. In the Yamaguchi approach (YA) [26], these constants can be calculated as the energy difference between the high-spin ferromagnetic state (hs) and the spin-projected antiferromagnetic state (sp) determinants divided by the difference of their respective $\hat{\mathbf{S}}^2$ operator expectation values, that is

$$J_{AB}(YA) = -\frac{{}^{hs}E - {}^{sp}E}{{}^{hs}\left\langle \widehat{\mathbf{S}}^2 \right\rangle - {}^{sp}\left\langle \widehat{\mathbf{S}}^2 \right\rangle} \tag{7}$$

Alternatively, in the local spin approach (LS), the coupling constants can also be calculated by means of the energy difference between the high-spin ferromagnetic state (hs) and the spin-projected antiferromagnetic state (sp) determinants divided by twice the difference of the two-center local spins in the form [27]

$$J_{AB}(LS) = -\frac{{}^{hs}E - {}^{sp}E}{2\left({}^{hs}\left\langle \widehat{\mathbf{S}}_A \cdot \widehat{\mathbf{S}}_B \right\rangle - {}^{sp}\left\langle \widehat{\mathbf{S}}_A \cdot \widehat{\mathbf{S}}_B \right\rangle\right)} \tag{8}$$

3 Results and discussion

The three carborane diradicals which arise from the connection of two radicals $CB_{11}H_{12}^{\cdot}$ through an acetylene, ethylene and ethane bridge unit are displayed in Fig. 1. While high-level quantum chemical computations are feasible for these dimers in the future, we would like to extend the system with more units to an (in)finite 1D chain and thus obtain the low-lying spin states, through a mapping of the results onto a Heisenberg spin Hamiltonian. A similar mapping has been carried out for many-electron systems using accurate post-HF calculations onto a generalized spin-exchange Hamiltonian, in clusters of hydrogen atoms [28].

As mentioned in the Introduction, we have obtained the singlet–triplet energy gaps and electronic energies for the structure displayed in Fig. 1a, using the spin-projected method and calibrating the results with very high-level quantum–mechanical computations [12]. The results have shown that the UB3LYP/6-31G(d) spin-projected method compares very well with high-level CASPT2/6-31G(d) computations. For this diradical, the ground state is of singlet nature, with a practically degenerate triplet state only 0.005 eV (CASPT2) higher in energy, which corresponds to the far-IR region of the electromagnetic spectrum, as reported. In this work, the UB3LYP/6-31G(d) spin-projected method has also been applied to the other two structures shown in Fig. 1, in order to know how the unpaired electrons couple to each other in these three structures, that is, when the

Table 1 Local spin populations $\langle \widehat{\mathbf{S}}_A \cdot \widehat{\mathbf{S}}_B \rangle$ for the diradicals studied in this work; hs and sp stand for high-spin and spin-projected wave function, respectively

Local spins				
	hs state		sp state	
	A	B	A	B
Acetylene bridge				
A	0.759	0.250	0.662	−0.158
B	0.250	0.749	−0.158	0.662
Ethylene bridge				
A	0.754	0.250	0.488	−0.095
B	0.250	0.753	−0.095	0.485
Ethane bridge				
A	0.752	0.250	0.758	−0.255
B	0.250	0.755	−0.255	0.759

A left moiety, B right moiety—see Fig. 1

Fig. 2 Spin density for the high-spin (triplet) states (*left*) and spin-projected states (*right*) in the three diradicals considered in this work: **a** $^\cdot$(HB)$_{11}$C–C≡C–C(BH)$_{11}^\cdot$ diradical, **b** $^\cdot$(HB)$_{11}$C–CH=CH–C(BH)$_{11}^\cdot$ diradical and **c** $^\cdot$(HB)$_{11}$C–CH$_2$–CH$_2$–C(BH)$_{11}^\cdot$ diradical. Spin density isovalue $\rho_s = \pm\, 0.001$ in all plots. For **a** and **b** two orientations of the diradical are considered: the one above corresponds to the same orientation as in Fig. 1. The one below is a 90° rotation of the former around the axis defined by the carbon atoms of the carborane cage

HIGH SPIN **SPIN PROJECTED**

(a)

(b)

(c)
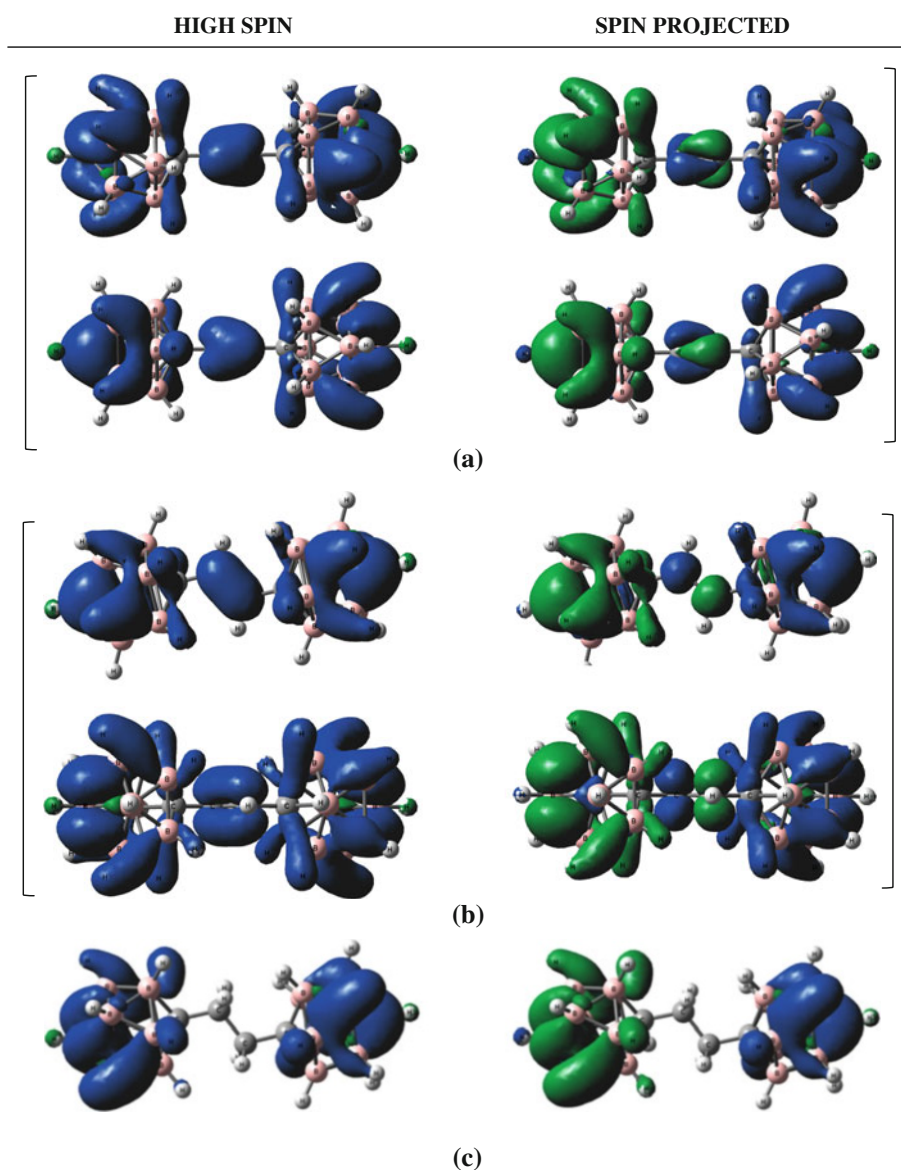
bridge unit connecting the carborane cages is of the acetylene, ethylene and ethane type.

The local spin populations $\langle \widehat{\mathbf{S}}_A \cdot \widehat{\mathbf{S}}_B \rangle$ (Eq. 6) in the studied diradicals are shown in Table 1, where A and B correspond, respectively, to the left and right moieties, as displayed in Fig. 1. The basis set dependence of atomic spin populations has been recently studied in Ref. [29].

The values found for $\langle \widehat{\mathbf{S}}_A^2 \rangle$ quantities in the three studied systems indicate that the hs state presents one-center local spin components close to 0.75 (the canonical value is ½(½ +1)), showing that those states possess a spin distribution corresponding to two well-localized electrons, each one in a moiety. The two-center local spin components $\langle \widehat{\mathbf{S}}_A \cdot \widehat{\mathbf{S}}_B \rangle$ are positive according to the coupling of two electrons to a triplet state and their values are close

to ¼. A slight spin contamination is only observed in the one-center terms. The sp states present values of one-center and two-center local spin quite different in the three systems. In the ethane bridge compound the one-center local spin value 0.76 is very close to the canonical one 0.75, meaning that the spin distribution is again that of two well-localized electrons, one in each fragment. However, in the acetylene bridge system the one-center terms have a larger difference respect of the canonical value, which must be interpreted in terms of two unpaired electrons slightly delocalized. This behavior, with a higher deviation from the corresponding canonical value, is also found in the case of the ethylene bridge diradical where the unpaired electrons are still more delocalized. Contrarily to the case of the hs states, the two-center components of the local spins are negative which corresponds to two unpaired electrons

Table 2 Energies (au), expectation values $\left\langle \widehat{\mathbf{S}}^2 \right\rangle$, singlet–triplet gaps ΔE_{ST} (in eV)—Eq. (3)—and coupling constants J_{AB} (in cm^{-1}) for the diradicals studied in this work; hs and sp stand for high-spin ($S = 1$) and spin-projected, respectively

	Energy	$\left\langle \widehat{\mathbf{S}}^2 \right\rangle$	ΔE_{ST}	J_{AB}(YA)	J_{AB}(LS)
Acetylene bridge					
hs state	−712.56328	2.0075	–	−15.27	−18.71
sp state	−712.56335	1.0079	−0.004		
Ethylene bridge					
hs state	−713.81780	2.0068	–	−321.6	−571.1
sp state	−713.81960	0.7834	−0.080		
Ethane bridge					
hs state	−715.04007	2.0077	–	−1.826	−1.810
sp state	−715.04008	1.0076	−0.0005		

YA and LS stand for Yamaguchi [26] and local spin [27] approaches, respectively

coupled to a singlet state. The values found are close to -0.25 only in the ethane bridge, due to the localization of the electrons in that system.

In order to visualize the spin distribution in these diradicals, Fig. 2 displays the spin density for the high-spin (left) and spin-projected (right) states of the three diradicals, following the same orientations as in Fig. 1. As evident from this Figure, there is clearly a left–right distribution of positive and negative spin density in the spin-projected states corresponding, respectively, to the A and B fragments of each diradical. In the acetylene and ethylene bridge diradicals (Fig. 2a, b), two projections of the spin density are represented—with a rotation of 90° degrees of the top respect to the bottom one, with the rotation around the axis joining the two carbon atoms of the carborane cage—since there is a noticeable contribution from the bridge units to the total spin density, and a different orientation of the molecule is needed in order to highlight the topological differences of the spin density between the acetylene (Fig. 2a) and ethylene (Fig. 2b) bridge units. One could find a topological similarity between the density of a π/π^* molecular orbital and the modulus of the high-spin/spin-projected spin density of the ethylene moiety in Fig. 2b. However, this is not the case for the acetylene bridge diradical, where no nodal planes, but rather nodal surfaces appear to separate the α- and β-spin densities in the bridge moiety. Given the negligible spin density in the ethane bridge diradical, a unique orientation, coinciding with the one from Fig. 1c is displayed in Fig. 2c.

In Table 2 we display the energies, the expectation values $\left\langle \widehat{\mathbf{S}}^2 \right\rangle$ and the coupling constants J_{AB} computed with

the Yamaguchi and local spin approaches, for the three diradicals considered in this work.

As shown in Table 2, the ground state in all three diradicals always corresponds to the estimated singlet state, although the triplet state lies very close in energy with J's in the order (in absolute value): J(ethylene) $\gg J$(acetylene) $> J$(ethane). The ethylene bridge (Fig. 1b) thus provides a "strong" interaction between the unpaired electrons in the isolated carborane cage radicals $CB_{11}H_{12}^{\cdot}$, which agrees with the local spin values above described for this compound. In the case of the ethane bridge (Fig. 1c), one can assume one almost isolated unpaired electron on each carborane cage, given the practical degeneracy of singlet and triplet states; the ethane bridge does not provide an "electronic coupling" between the spins of the unpaired electrons from A and B fragments. The case of the acetylene bridge (Fig. 1a), already studied in Ref. [12], is an intermediate case, with a significant coupling between the spins of the electrons in the fragments A and B. The numerical values found for the coupling constants in both procedures (YA and LS) are very similar in the case of the ethane bridged diradical. However, these values present a difference for the acetylene bridge, which becomes even larger in the case of the ethylene one. These results are in agreement with the degree of electron localization in these three compounds. In fact, only under the assumption of unpaired electrons localized on the radical centers in both hs and sp states, one should expect similar results for both models [22].

4 Concluding remarks and perspectives

In this work, we have presented a detailed electronic structure analysis of three diradicals derived from the connection of two icosahedral carborane radicals $CB_{11}H_{12}^{\cdot}$ through the cage carbon atom with acetylene, ethylene and ethane bridge units. All diradicals have a singlet ground state with very low-lying triplet states. The ethane bridge diradical is practically degenerate, with an energy gap below 1 meV. The acetylene bridge diradical shows a certain coupling between the carborane units, with an energy gap of 4 meV, and the case of ethylene bridge has a higher electron coupling up to 80 meV. These results in terms of energy are consistent with the conclusions arising from the spin population analysis. The local spin values found correspond to the presence of two well-localized electrons in the hs state of each of these compounds, but the analysis for the sp state points out two electrons well-localized (noninteracting) in the case of the ethane bridge, slightly delocalized in the case of acetylene bridge and more delocalized in the ethylene bridge one. The electron

coupling between the carborane cages is due to the ethylene and acetylene bridge unit as clearly shown through spin density plots.

The next challenge is to predict low-lying states in larger polyradical one-dimensional or cyclic chains through a mapping of the current results onto a Heisenberg spin Hamiltonian for a set of carborane clusters. As mentioned above, similar mappings have been performed, for instance in clusters of hydrogen atoms and using accurate post-HF calculations with a further mapping onto a generalized spin-exchange Hamiltonian [28]. This task is currently being performed in our Laboratories.

References

1. Bould J, Baše T, Londesborough MGS, Oro LA, Macías R, Kennedy RJD, Kubát P, Fuciman M, Polívka T, Lang K (2011) Inorg Chem 50:7511
2. van der Vlugt JI (2010) Angew Chem Int Ed 49:252
3. Scholz M, Hey-Hawkins E (2001) Chem Rev 111:7035
4. El-Zaria ME, Ban HS, Nakamura H (2010) Chem Eur J 16:1543
5. Sivaev IB, Bregadze VI, Sjöberg S (2002) Collect Czech Chem Commun 67:679
6. Klessinger M, Michl J (1994) Excited states and photochemistry of organic molecules. Wiley-VCH, New York
7. Valeur B (2002) Molecular fluorescence. Wiley-VCH, Weinheim
8. Grimes RN (2011) Carboranes, 2nd edn. Academic Press, New York
9. Serrano-Andrés L, Klein DJ, Schleyer PVR, Oliva JM (2008) J Chem Theory Comput 4:1338
10. Oliva JM, Serrano-Andrés L (2006) J Comput Chem 27:524
11. Londesborough MG, Hnyk D, Bould J, Serrano-Andrés L, Saurí V, Oliva JM, Kubát P, Polívka T, Lang K (2012) Inorg Chem 51:1471
12. Oliva JM, Serrano-Andrés L, Havlas Z, Michl J (2009) J Mol Struct (Theochem) 912:13
13. Oliva JM (2012) Adv Quantum Chem 64:105
14. Oliva JM (2011) Cyclic spin architectures built from carborane radicals, XXXIII reunión bienal de la Real Sociedad Española de Física, Santander, Spain, 19–23 Septiembre 2011, p 45
15. Oliva JM (2011) 11th international conference on computational and mathematical methods in science and engineering, CMMSE 2011, Alicante, Spain, 16–20 June 2011, p 78
16. King BT, Noll BC, McKinley AJ, Michl J (1996) J Am Chem Soc 118:10902
17. Eriksson L, Vyakaranam K, Ludvík J, Michl J (2007) J Org Chem 72:2351
18. Roos BO, Andersson K, Fulscher MP, Malmqvist PA, Serrano-Andrés L, Pierloot K, Merchán M (1996) Adv Chem Phys 93:219
19. Atkins PW, de Paula J (2009) Physical chemistry, 9th edn. Oxford University Press, Oxford
20. Frisch MJ et al (2004) Gaussian 03, revision C.02. Gaussian Inc., Wallingford
21. Ovchinnikov AA, Labanowski JK (1996) Phys Rev A 53:3946
22. Clark A, Davidson ER (2001) J Chem Phys 115:7382
23. Podewitz M, Herrmann C, Malassa A, Westerhausen R, Reiher M (2008) Chem Phys Lett 451:301
24. Alcoba DR, Lain L, Torre A, Bochicchio RC (2009) Chem Phys Lett 470:136
25. Alcoba DR, Torre A, Lain L, Bochicchio RC (2011) J Chem Theory Comput 7:3560
26. Soda T, Kitagawa Y, Onishi T, Takano Y, Shigeta Y, Nagao H, Yoshioka Y, Yamaguchi K (2000) Chem Phys Lett 319:223
27. Herrmann C, Yu L, Reiher M (2006) J Comput Chem 27:1223
28. Ciofinia I, Adamo C, Barone V, Berthier G, Rassat A (2005) Chem Phys 309:133
29. Philips JJ, Hudspeth MA, Browne PM Jr, Peralta JE (2010) Chem Phys Lett 495:146

REGULAR ARTICLE

A theoretical investigation of the CO_2-philicity of amides and carbamides

**Luis Miguel Azofra · Muhannad Altarsha ·
Manuel F. Ruiz-López · Francesca Ingrosso**

Abstract The knowledge of the interactions taking place at a molecular level can help the development of new technological procedures in Chemistry with low environmental impact. In organic, biochemical and pharmaceutical synthesis and in analytical chemistry, important advances in this domain are related to the use of solvents that can be valid alternatives to hazardous organic solvents. In the last decades, a large emphasis has been given to the use of carbon dioxide under supercritical conditions, since the mild temperature and pressure conditions of the fluid can easily be controlled to improve its capacity to solubilize small organic compounds. On the other hand, the solubility of larger molecules and of polar compounds in this medium is generally very low. This has motivated recent theoretical and experimental studies with the purpose of reaching a better understanding of the so-called CO_2-philicity of molecules and materials, and very encouraging results have been reported. In this paper, we present an ab initio study of the intermolecular interactions between CO_2 and amide and carbamide derivatives, performed on model 1:1 complexes at the MP2/aug-cc-pVTZ//MP2/aug-cc-pVDZ level. Our findings shed some light on the key points to be considered in the design of large CO_2-philic molecules, hinting at the use of solubilizer groups in which amide or urea bonds could be involved.

Published as part of the special collection of articles derived from the 8th Congress on Electronic Structure: Principles and Applications (ESPA 2012).

Electronic supplementary material The online version of this article (doi:10.1007/s00214-012-1326-4) contains supplementary material, which is available to authorized users.

L. M. Azofra · M. Altarsha (✉) · M. F. Ruiz-López (✉) ·
F. Ingrosso (✉)
Université de Lorraine, SRSMC UMR 7565,
54506 Vandœuvre-lès-Nancy Cedex, France
e-mail: Muhannad.Altarsha@univ-lorraine.fr

M. F. Ruiz-López
e-mail: Manuel.Ruiz@univ-lorraine.fr

F. Ingrosso
e-mail: Francesca.Ingrosso@univ-lorraine.fr

L. M. Azofra · M. Altarsha · M. F. Ruiz-López · F. Ingrosso
CNRS, SRSMC UMR 7565, 54506 Vandœuvre-lès-Nancy
Cedex, France

L. M. Azofra
Instituto de Química Médica (C.S.I.C.), c/Juan de la Cierva,
3, 28006 Madrid, Spain

1 Introduction

Supercritical carbon dioxide (scCO₂) is becoming an important commercial and industrial solvent and has attracted increasing attention for the development of green chemical processes. In addition to its low toxicity and environmental impact, scCO₂ is readily available since its critical point is characterized by a critical temperature of 31.1 °C and a critical pressure of 7.4 MPa [1–4].

However, a strong limitation to a wider use of industrial technologies based on this solvent is related to the low solubility of large molecules and of polar compounds. A clear picture of the intermolecular interactions that can be exploited to trigger a better solubility is therefore necessary to assist in the design of organic, biochemical and pharmaceutical synthesis/separation procedures taking place in

scCO$_2$. In the past years, different studies have tried to propose new routes to an improved CO$_2$-philicity, a concept that was introduced by analogy with the properties of aqueous systems and which has been related to Lewis acid/base (LA–LB) interactions. The first success in the design of CO$_2$-philic materials was achieved with the development of fluorinated polymers [5]. However, the technology used for fluorination is quite expensive, and it can be problematic from the environmental viewpoint [6]. The interpretation of the CO$_2$-philic character of fluorinated compounds, based on experimental and theoretical investigations, has been reviewed in Ref. [6]. Some specific interactions between the F atom and the electron-poor C atom of CO$_2$ have been pointed out, in addition to an influence of F on the acidity of neighboring H atoms, which makes them H-bond donors with respect to the O atoms of CO$_2$.

The search for non-fluorous molecules soluble in scCO$_2$ was stimulated by some work pointing out LA–LB interactions between CO$_2$ and polymers possessing electron-donating functional groups such as the carbonyl group [7]. In the following years, a great deal of work has been devoted to the CO$_2$-philicity of carbonyl derivatives, including the synthesis of functionalized silicones [8], of diglicolyc acid esters [9], and of amide derivatives [10, 11]. A high CO$_2$-solubility has been found for sugar derivatives and for poly(ether-carbonate) copolymers [12, 13]. Oligomeric surfactants based on glycol ethers have been developed with a different purpose, such as CO$_2$ capture, to improve the absorption of excess CO$_2$ from the atmosphere [14]. It is worth mentioning that some non-fluorous, non-carbonyl compounds have also been shown to be soluble in CO$_2$, among which bipyridine derivatives [15], polycyclic aromatic hydrocarbons [16], and some recently synthesized hybrid surfactants [17].

To complement the experimental knowledge of these interactions, quantum chemistry studies have been carried out, in particular for complexes formed by CO$_2$ with carbonyl derivatives [5, 18–28]. The references here reported were discussed in depth elsewhere [26]. Complexes of CO$_2$ with ethylene and acetylene have also been described [29]. In general, it has been shown that these complexes are stabilized by LA–LB interactions and that the CO$_2$ molecule behaves as a Lewis acid, in accord with the usual chemical concepts. However, in a recent study [30], we discovered that unconventional four-membered ring structures exist for CO$_2$-carbonyl compound complexes, in which CO$_2$ behaves cooperatively as both a Lewis base and a Lewis acid and which are at least as stable as the traditional structures. In subsequent work [26], we have reported a systematic investigation at the MP2 and CCSD(T) levels of complexes between CO$_2$ and aldehydes, ketones and esters together with some fluorinated

derivatives. We have shown that the LB character of CO$_2$ is inoperative in the interaction with aldehydes, while it plays a key role in the interactions with ketones and esters, especially in the case of fluorinated derivatives. Experimental data on some of these 1:1 complexes were available, thus allowing us to validate our theoretical procedure.

Some recent and encouraging experimental work has reported high scCO$_2$ solubilities for newly synthesized amide derivatives [10, 11]. Moreover, the organic synthesis of carbamide derivatives has been successfully carried out in scCO$_2$ through a reaction scheme in which CO$_2$ is at the same time a reactant and the reaction medium, instead of the usual method in organic solvents that employs hazardous reagents such as phosgene [31]. Motivated by these new findings, in this work we provide a theoretical investigation about the nature of solute–solvent interactions for amides and carbamides. Indeed, these compounds display a conjugated π system involving the p_z orbitals of the O, C and N atoms, and the interactions with CO$_2$ might present significant differences with respect to carbonyl derivatives studied in Ref. [26] that need to be addressed. The basic molecules formamide and urea together with some derivatives obtained from the latter by methylation were included in our study.

The electronic features that can explain the CO$_2$-phylicity of the amide and urea bonds were investigated by means of ab initio calculations for a set of 1:1 complexes, comprising geometry optimization, calculation of the interaction energies, natural bond analysis and a study of the molecular orbitals involved in the intermolecular interactions.

2 Computational methodology

The relative energetic stability and the electronic properties of the complexes formed by CO$_2$ with the following molecules were studied: formamide, acetamide, N-methylacetamide, N,N-dimethylacetamide, azetidin-2-one, N-methylazetidin-2-one, urea, N-methylurea and N,N'-dimethylurea. When different cis/trans isomers were possible, all the corresponding complexes were taken into account (see Section III for further details). The whole list of investigated molecules and cis/trans isomers is summarized in Fig. 1.

Geometry optimization for all the monomers and the complexes were carried out at the second-order perturbation theory level (MP2 [32]) using the aug-cc-pVDZ basis set [33, 34]. Harmonic frequency calculations were performed to confirm the nature of the potential energy surface minima. Single-point energies were computed using the MP2/aug-cc-pVTZ level on the geometries that were optimized at the MP2/aug-cc-pVDZ level. The interaction energies of the complexes were then calculated as the
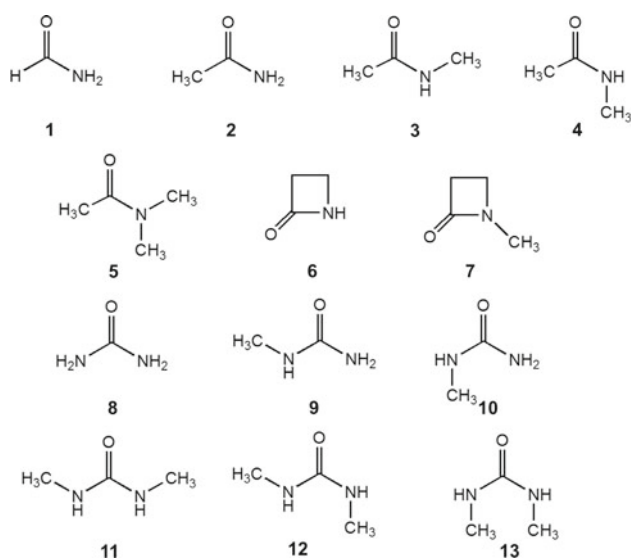
Fig. 1 Molecular structure of the molecules whose 1:1 complexes with CO_2 were analyzed in this work

difference between the energy of the complex and the sum of the energies of the relaxed isolated monomers. In a previous work [26], we showed that this computational scheme provides results close to the much more costly CCSD(T)/aug-cc-pVDZ calculations. Basis set superposition errors were estimated in the same work and were shown to be small; hence, they will not be calculated here. We also showed that the geometry of this type of complexes can be obtained through DFT calculations using the M06-2x [35] and wB97XD [36] functionals, although the results are not very accurate.

Natural bonding orbitals (NBO) calculations [37, 38] were performed (using SCF densities and MP2/aug-cc-pVDZ optimized geometries) to examine the electronic properties of the complexes, in particular to examine the nature of the interactions. All calculations were run using Gaussian 09 [39].

3 Results and discussion

3.1 Structures

We start the discussion by presenting a description of the structure of different types of complexes between the molecules that were considered in this study (Fig. 1) and a CO_2 molecule. We have limited the analysis to complexes of type I and III, according to the nomenclature proposed in our previous work [26], and to complexes of type IV, specific for amides and carbamides. These structures are summarized in Scheme 1. T-shaped complexes (type II in Ref. [26]) and structures involving only hydrogen bonds

are also possible, but in principle, they should be less stable than I and have not been considered here.

As it has been already discussed [26–28], in complexes of type I, CO_2 behaves as a LA, whereas the carbonyl compound plays the role of a LB. Besides, a weak hydrogen bond is formed between an H atom on the molecule (in α or β position) and the O atom of CO_2. In the present study, the Ia complex can only be found in the case of formamide and the Ib complex may involve a hydrogen atom attached to the a N or C atoms. These two situations are indicated as Ib(NH) and Ib(CH), respectively. As shown below, in one asymmetric urea derivative (molecule 10), two distinct Ib(NH) complexes are possible; we distinguish the two of them by using the labels Ib(NH) and Ib(NH′). Complexes of type IIIa have a particular geometry in which four atoms form a roughly planar interaction site (a four-membered ring), and two kinds of LA–LB interactions are active, with CO_2 acting as either a donor or an acceptor of electron density. Complexes of type IV are reported here for the first time. They involve interactions between CO_2 and the π system of the amide/urea moiety, as described below.

The optimized geometries for the complexes that we treated are summarized in Fig. 2: for all reported minima, we obtained real frequencies only. We observe that in the case of formamide, only I type complexes could be found. This result agrees with what has been found in the case of aldehydes [26] and suggests that the out-of-plane structures are not stable when an H atom is directly attached to the carbonyl group. All complex types (I, III and IV) were found for the other compounds, apart from the following cases: a) no IIIa complexes were observed in acetamide or in urea derivatives having at least one N–H bond in the syn periplanar position with respect to the C=O bond, and b) no IV complexes were found with β-lactams 6 and 7.

3.2 Energies

In the following, we present an analysis of the relative stabilities of different complexes. The interaction energies calculated at the MP2/aug-cc-pVDZ level are shown in Table 1 together with the MP2/aug-cc-pVTZ single-point energies calculated on MP2/aug-cc-pVDZ optimized geometries. Although some slight differences can be observed, the general trends predicted by the two methods are in reasonable agreement. The interaction energies vary between −3.88 and −5.87 kcal/mol at the MP2/aug-cc-pVTZ level and between −3.33 and −6.79 kcal/mol at the MP2/aug-cc-pVDZ level. These values are comparable to those previously found for aldehydes, ketones and esters, though in some cases, the interaction energies are slightly higher (in absolute value). The largest interaction energy with CO_2 occurs for the IV(NCH_3) complex formed by
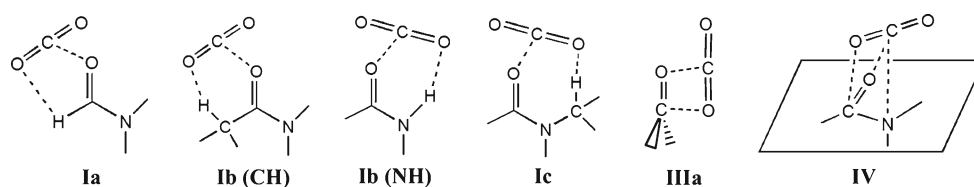
Scheme 1 Schematic representation of the interactions taking place in the different complexes studied in this work
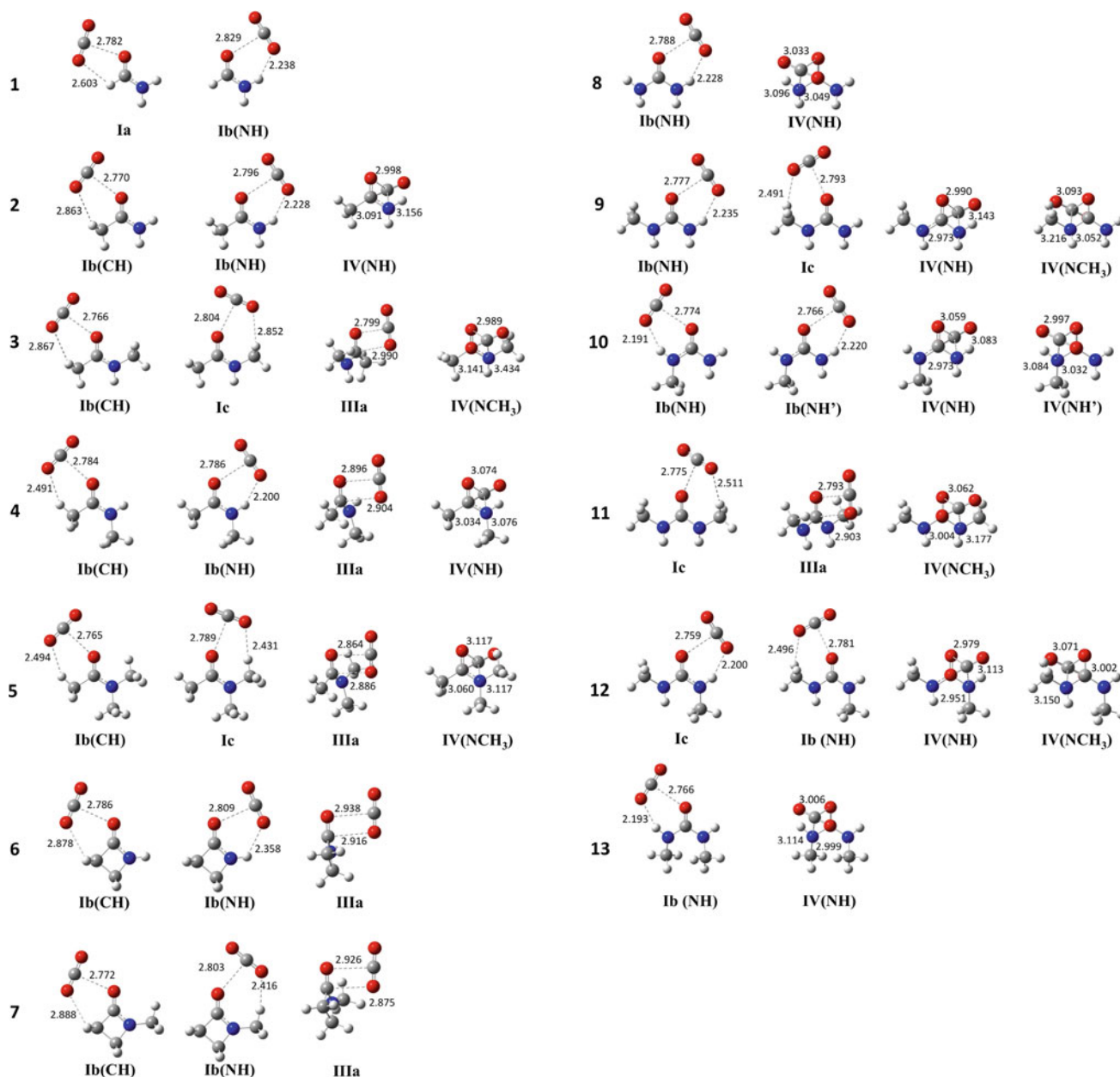
Fig. 2 Optimized structures (at the MP2/aug-cc-pVDZ level) for the complexes that were studied in this work. Distances characterizing the main interactions between the molecules shown in Fig. 1 and CO_2 are reported in Å

compound 11, that is *N,N'*-dimethylurea with both methyl groups in syn periplanar position with respect to C=O. We report as Supporting Information the interaction energies

corrected by the zero point energy (ZPE) contribution, enthalpies and entropies (MP2/aug-cc-pVDZ level) as obtained from the harmonic vibrational frequencies

Table 1 Calculated interaction energies (in kcal/mol) for the complexes in Fig. 2

Amides	Complex type	ΔE		Carbamides	Complex type	ΔE	
		DZ	TZ			DZ	TZ
1	Ia	−4.05	−3.88	8	Ib(NH)	−5.14	−5.01
	Ib(NH)	−4.93	−4.76		IV(NH)	−4.30	−3.78
2	Ib(NH)	−5.21	−5.01	9	Ib(NH)	−5.23	−5.09
	Ib(CH)	−4.20	−3.98		Ic	−4.23	−4.02
	IV(NH)	−4.27	−3.81		IV(NCH$_3$)	−4.83	−4.19
3	Ib(CH)	−4.43	−4.16		IV(NH)	−5.06	−4.43
	Ic	−4.28	−4.06	10	Ib(NH)	−5.44	−5.30
	IIIa	−4.83	−4.29		Ib(NH')	−5.65	−5.46
	IV(NCH$_3$)	−4.73	−4.09		IV(NH)	−5.18	−4.51
4	Ib(NH)	−5.49	−5.25		IV(NH')	−4.78	−4.24
	Ib(CH)	−4.39	−4.17	11	Ic	−5.52	−5.12
	IIIa	−4.74	−4.18		IIIa	−5.55	−4.89
	IV(NH)	−4.74	−4.16		IV(NCH$_3$)	−6.79	−5.87
5	Ib(CH)	−4.56	−4.33	12	Ib(NH)	−5.74	−5.54
	Ic	−4.45	−4.18		Ic	−4.47	−4.26
	IIIa	−5.75	−4.99		IV(NCH$_3$)	−5.81	−5.02
	IV(NCH$_3$)	−5.45	−4.61		IV(NH)	−5.57	−4.87
6	Ib(NH)	−5.10	−4.84	13	Ib(NH)	−5.66	−5.44
	Ib(CH)	−4.55	−4.18		IV(NH)	−5.41	−4.74
	IIIa	−4.37	−3.90				
7	Ib(CH)	−4.55	−4.22				
	Ic	−4.73	−4.36				
	IIIa	−5.25	−4.53				

The labels DZ and TZ stand for the MP2/aug-cc-pVDZ and for the MP2/aug-cc-pVTZ//MP2/aug-cc-pVDZ calculations, respectively

analysis carried out at 298 K. The corresponding values for the acetone complexes are reported as a comparison. Note that the ZPE contribution is similar for all the complexes (1.0–1.5 kcal/mol) and does not affect the stability trends discussed below, which are based on electronic energies. Note also that the computed entropies are comparable to those reported for 1:1 complexes having similar interaction energies.

As a rule, we can say that carbamide-CO$_2$ complexes are slightly more stable than the amide-CO$_2$ ones. Indeed, for each carbamide-CO$_2$ complex, there is at least one arrangement for which the interaction energy is greater than (in absolute value) 5 kcal/mol, whereas the interaction energy in amide-CO$_2$ complexes is in general below that value (the exceptions are the 2 Ib(NH) and 4 Ib(NH) complexes that display interactions energies of −5.01 and −5.25 kcal/mol, respectively). Some other general trends are (1) whenever the Ib(NH) interaction type is possible for a given complex (either in amide or carbamide derivatives), such complex type is the most favorable one and (2) when the Ib(NH) interaction type is not possible, then the IIIa interaction type becomes the most favorable one

(except in 11 where, as said above, IV(NCH$_3$) leads to the largest interaction type in the whole series). The average interaction energy for Ib(NH) complexes is −5.17 kcal/mol, and for IIIa complexes −4.46 kcal/mol (similarly, one obtains average values of −3.88 kcal/mol for Ia, −4.17 kcal/mol for Ib(CH), −4.33 kcal/mol for Ic and −4.49 kcal/mol for IV complexes; the average energy, obtained by taking into account all the structures, is −4.56 kcal/mol).

3.3 NBO analysis

To investigate the nature of the interactions between CO$_2$ and the different molecules that we considered, and to connect the geometrical properties of the complexes with the electronic and energetic properties, we carried out natural bond orbital analysis using SCF densities and optimized geometries at the MP2/aug-cc-pVDZ level. In Table 2, we analyze the main contributions to donor–acceptor second-order energies, also providing the nature of the molecular orbitals involved in the interactions. The shape of the orbitals involved in all the complexes of 3 with

Table 2 NBO analysis: results obtained by using second-order perturbation theory to characterize intermolecular interactions

Complex		Donor/acceptor	Orbitals	E^2	Complex		Donor/acceptor	Orbitals	E^2
1	Ia	Substrate/CO_2	$n\pi^*$	1.86	8	Ib(NH)	Substrate/CO_2	$n\pi^*$	1.54
		CO_2/Substrate	$n\sigma_{CN^*}$	0.26			CO_2/Substrate	$n\sigma_{NH^*}$	1.99
		CO_2/Substrate	$n\sigma_{CH^*}$	0.28			CO_2/Substrate	$n\sigma_{NH^*}$	1.62
		CO_2/Substrate	$n\sigma_{CH^*}$	0.37		IV(NH)	Substrate/CO_2	$\pi\pi^*$	0.49
	Ib(NH)	Substrate/CO_2	$n\pi^*$	1.39			Substrate/CO_2	$n\pi^*$	0.46
		CO_2/Substrate	$n\sigma_{NH^*}$	1.75			CO_2/Substrate	$n\pi^*$	0.86
		CO_2/Substrate	$n\sigma_{NH^*}$	1.72	9	Ib(NH)	Substrate/CO_2	$n\pi^*$	1.53
2	Ib(NH)	Substrate/CO_2	$n\pi^*$	1.53			CO_2/Substrate	$n\sigma_{NH^*}$	1.55
		CO_2/Substrate	$n\sigma_{NH^*}$	1.55			CO_2/Substrate	$n\sigma_{NH^*}$	1.96
		CO_2/Substrate	$n\sigma_{NH^*}$	1.96		Ic	Substrate/CO_2	$n\pi^*$	0.99
	Ib(CH)	Substrate/CO_2	$n\pi^*$	1.76			CO_2/Substrate	$n\sigma_{CH^*}$	0.98
		CO_2/Substrate	$n\sigma_{CH^*}$	0.23		IV(NCH$_3$)	Substrate/CO_2	$\pi\pi^*$	0.69
	IV(NH)	Substrate/CO_2	$\pi\pi^*$	0.63			Substrate/CO_2	$n\pi^*$	0.22
		Substrate/CO_2	$n\pi^*$	0.24			CO_2/Substrate	$n\pi^*$	0.87
		CO_2/Substrate	$n\pi^*$	0.64			CO_2/Substrate	$n\sigma_{CH^*}$	0.48
3	Ib(CH)	Substrate/CO_2	$n\pi^*$	1.71		IV(NH)	Substrate/CO_2	$\pi\pi^*$	0.78
		CO_2/Substrate	$n\sigma_{CH^*}$	0.23			Substrate/CO_2	$n\pi^*$	0.19
	Ic	Substrate/CO_2	$n\pi^*$	1.00			CO_2/Substrate	$n\pi^*$	1.23
	IIIa	Substrate/CO_2	$\pi\pi^*$	1.52	10	Ib(NH)	Substrate/CO_2	$n\pi^*$	1.62
		CO_2/Substrate	$n\pi^*$	0.57			CO_2/Substrate	$n\sigma_{NH^*}$	2.40
	IV(NCH$_3$)	Substrate/CO_2	$\pi\pi^*$	0.96			CO_2/Substrate	$n\sigma_{NH^*}$	2.42
		Substrate/CO_2	$n\pi^*$	0.13		Ib(NH')	Substrate/CO_2	$n\pi^*$	1.62
		CO_2/Substrate	$\pi\sigma_{CH^*}$	0.61			CO_2/Substrate	$n\sigma_{NH^*}$	1.66
		CO_2/Substrate	$n\pi^*$	0.34			CO_2/Substrate	$n\sigma_{NH^*}$	2.06
4	Ib(NH)	Substrate/CO_2	$n\pi^*$	1.58		IV(NH)	Substrate/CO_2	$\pi\pi^*$	0.60
		CO_2/Substrate	$n\sigma_{NH^*}$	2.38			Substrate/CO_2	$n\pi^*$	0.34
		CO_2/Substrate	$n\sigma_{NH^*}$	2.38			CO_2/Substrate	$n\pi^*$	1.25
	Ib(CH)	Substrate/CO_2	$n\pi^*$	1.46		IV(NH')	Substrate/CO_2	$\pi\pi^*$	0.47
		CO_2/Substrate	$n\sigma_{CH^*}$	0.65			Substrate/CO_2	$n\pi^*$	0.66
		CO_2/Substrate	$n\sigma_{CH^*}$	0.95			CO_2/Substrate	$n\pi^*$	0.64
	IIIa	Substrate/CO_2	$\pi\pi^*$	1.25	11	Ic	Substrate/CO_2	$n\pi^*$	1.10
		CO_2/Substrate	$n\pi^*$	1.72			CO_2/Substrate	$n\sigma_{CH^*}$	0.94
	IV(NH')	Substrate/CO_2	$\pi\pi^*$	0.55		IIIa	Substrate/CO_2	$\pi\pi^*$	1.03
		Substrate/CO_2	$n\pi^*$	0.54			CO_2/Substrate	$n\pi^*$	0.99
		CO_2/Substrate	$n\pi^*$	0.92		IV(NCH$_3$)	Substrate/CO_2	$\pi\pi^*$	0.79
5	Ib(CH)	Substrate/CO_2	$n\pi^*$	1.44			Substrate/CO_2	$n\pi^*$	0.22
		CO_2/Substrate	$n\sigma_{CH^*}$	0.65			CO_2/Substrate	$n\pi^*$	1.03
		CO_2/Substrate	$n\sigma_{CH^*}$	0.98			CO_2/Substrate	$n\sigma_{CH^*}$	0.45
	Ic	Substrate/CO_2	$n\pi^*$	1.08	12	Ib(NH)	Substrate/CO_2	$n\pi^*$	1.63
		CO_2/Substrate	$n\sigma_{CH^*}$	1.42			CO_2/Substrate	$n\sigma_{NH^*}$	2.32
	IIIa	Substrate/CO_2	$\pi\pi^*$	1.41			CO_2/Substrate	$n\sigma_{NH^*}$	2.36
		CO_2/Substrate	$n\pi^*$	1.08		Ic	Substrate/CO_2	$n\pi^*$	1.09
	IV(NCH$_3$)	Substrate/CO_2	$\pi\pi^*$	0.58			CO_2/Substrate	$n\sigma_{CH^*}$	0.96
		Substrate/CO_2	$n\pi^*$	0.54		IV(NCH$_3$)	Substrate/CO_2	$\pi\pi^*$	0.72
		CO_2/Substrate	$n\pi^*$	0.70			Substrate/CO_2	$n\pi^*$	0.31
		CO_2/Substrate	$\pi\sigma_{CH^*}$	0.24			CO_2/Substrate	$n\pi^*$	1.00
6	Ib(NH)	Substrate/CO_2	$n\pi^*$	1.58			CO_2/Substrate	$\pi\sigma_{CH^*}$	0.42
		CO_2/Substrate	$n\sigma_{NH^*}$	1.65		IV(NH)	Substrate/CO_2	$\pi\pi^*$	0.65

Table 2 continued

Complex		Donor/acceptor	Orbitals	E^2	Complex		Donor/acceptor	Orbitals	E^2
		CO$_2$/Substrate	$n\sigma_{NH^*}$	1.25			Substrate/CO$_2$	$n\pi^*$	0.46
	Ib(CH)	Substrate/CO$_2$	$n\pi^*$	1.58			CO$_2$/Substrate	$n\pi^*$	1.03
		CO$_2$/Substrate	$n\sigma_{CC^*}$	0.48	13	Ib(NH)	Substrate/CO$_2$	$n\pi^*$	1.61
	IIIa	Substrate/CO$_2$	$\pi\pi^*$	1.08			CO$_2$/Substrate	$n\sigma_{NH^*}$	2.31
		CO$_2$/Substrate	$n\pi^*$	1.65			CO$_2$/Substrate	$n\sigma_{NH^*}$	2.46
7	Ib(CH)	Substrate/CO$_2$	$n\pi^*$	1.80		IV(NH)	Substrate/CO$_2$	$\pi\pi^*$	0.75
		CO$_2$/Substrate	$n\sigma_{CC^*}$	0.48			Substrate/CO$_2$	$n\pi^*$	0.35
	Ic	Substrate/CO$_2$	$n\pi^*$	1.07			CO$_2$/Substrate	$n\pi^*$	1.14
		CO$_2$/Substrate	$n\sigma_{CH^*}$	1.25					
	IIIa	Substrate/CO$_2$	$\pi\pi^*$	1.14					

The second-order energetic contributions (E^2) to the most important donor/acceptor interactions are reported in kcal/mol

Fig. 3 Shape of the natural molecular orbitals involved in the main donor–acceptor intermolecular interactions of different type of structures for all the complexes formed by *N*-methylacetamide with CO$_2$
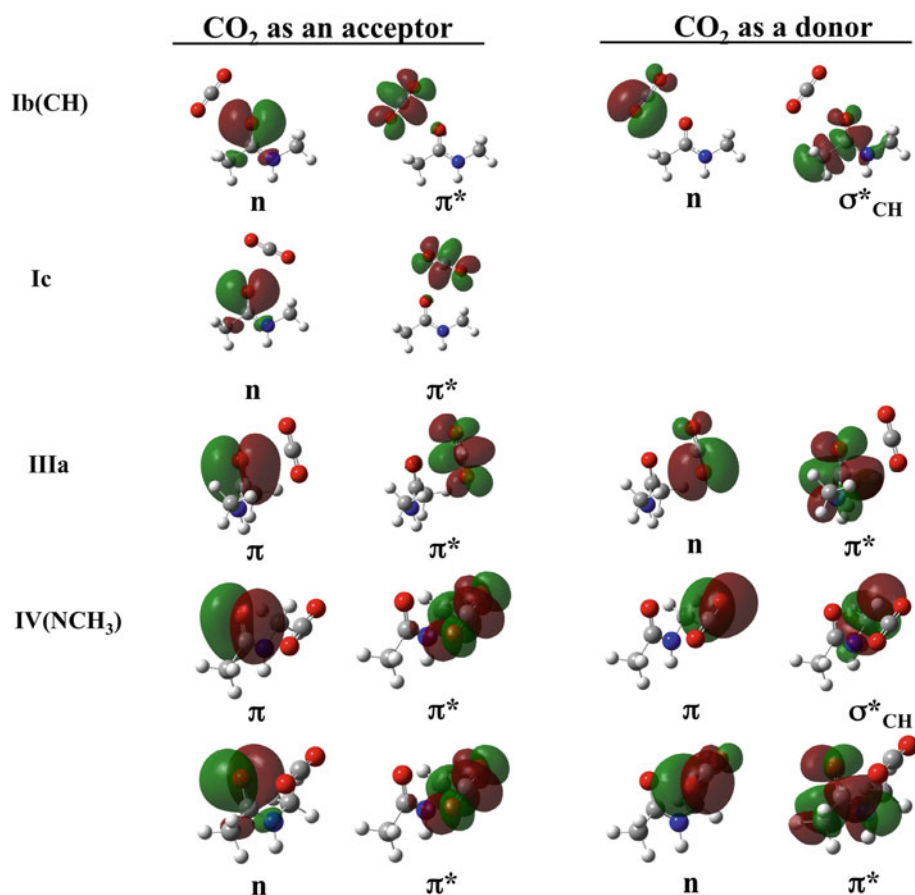
CO$_2$ is represented in Fig. 3. For simplicity, the amide or carbamide derivative interacting with CO$_2$ will be referred to as "substrate".

In the case of complexes I and III, our results generalize what we had already found for aldehydes (only I complexes are formed in that case), ketones and esters.

In complexes of type I, the main interactions are (1) a typical LB–LA donor–acceptor interaction, in which the *n* orbital of the carbonyl group of the substrate interacts with the π^* orbital of CO$_2$ and (2) one or two interactions in which the in-plane *n* orbitals of CO$_2$ interact with σ^* orbitals of the substrate. In the case of the Ia complex, two different σ^* orbitals are involved, one for the C–H bond and one for the C–N bond. For Ib complexes, we need to distinguish between Ib(CH) and Ib(NH) interactions. In the former case, only the σ^* orbital of a C–H bond is involved

but for the four-membered cyclic molecules 6 and 7, where we interestingly observe a weak interaction involving the σ* orbital for the C–C bond (note in Fig. 2 that the C–H bond in 6 and 7 is out-of-plane). In Ib(NH) complexes, we observe two interactions between n orbitals of CO_2 and the σ* orbital of the N–H bond of the substrate. Finally, for Ic complexes, the σ* orbital involved is that of the C–H bond of the *N*-methyl group. The interactions in which CO_2 behaves as an acceptor and those in which it behaves as a donor have usually the same order of magnitude, with a few exceptions in complexes Ib(NH) of urea derivatives, where the two $n\sigma^*$ interactions related to the O···H–N hydrogen bond have, not surprisingly, a much larger contribution. This fact is connected to the acidity of the NH amide proton and explains the large stability of the Ib(NH) structures discussed above.

In IIIa complexes, we observe (1) a contribution due to the π orbital of the substrate interacting with the π* orbital of CO_2 and (2) a contribution due to the n orbital of CO_2 interacting with a π* orbital of the substrate. In the latter case, the CO_2 molecule behaves as a Lewis base.

Let us now consider IV complexes, which are specifically formed in the case of amides and carbamides. As shown in Table 2: (1) both a π and an n orbital localized on the carbonyl group of the substrate can interact with a π* orbital of CO_2, and (2) an n orbital of CO_2 interacts with a π* orbital of the substrate in all IV complexes; it is worth emphasizing that this π* orbital has a substantial contribution from the p_z orbital on the N atom. In addition, in the case of IV(NCH₃) complexes (i.e., when a methyl group is in the syn periplanar position with respect to the carbonyl group), there is a weak hydrogen bond interaction that involves a π orbital of CO_2 and the C–H σ* orbital of the *N*-methyl group of the substrate.

It is apparent now to what extent the presence of the a nitrogen atom(s) in amides and carbamides influences the interaction of carbonyl groups with carbon dioxide. Compared to carbonyl compounds in aldehydes, ketones or esters, the 1:1 complexes with CO_2 can be significantly stabilized due to the interaction of CO_2 with the acidic H–N protons. However, this is not the only important feature, and in fact, the presence of the N atom, through its participation to the delocalized π system, allows for a direct interaction with CO_2. In the latter case, CO_2 behaves as a LB, one of its n orbitals donating electron charge to a delocalized π* antibonding orbital in the amide/urea unit.

4 Conclusions

The study reported in this paper shows that amide and carbamide derivatives have significant affinity for carbon dioxide. They therefore appear as suitable and promising functionalizing groups for increasing the solubility of large molecules in supercritical CO_2. The case of amides is particularly interesting because our results suggest that the interactions of CO_2 with the backbone of peptides and proteins might provide a substantial contribution to the stabilization of such systems in supercritical CO_2, a topic that has high biotechnological relevance.

Moreover, the results reported above confirm and extend the previous conclusion obtained in the study of aldehydes, ketones and esters, stating that, beyond its usual Lewis acid character, CO_2 may behave as a Lewis base when it interacts with carbonyl compounds. In the case of amides and carbamides, we have seen that such an interaction is characterized by the participation of a π* orbital delocalized over the O, C and N atoms (structures IV). However, when the amide or carbamide derivative bears an NH proton in the syn periplanar position with respect to the carbonyl group, the main interaction corresponds to an in-plane arrangement (structures Ib(NH)) in which there is cooperativity between the usual LA(CO_2)–LB(carbonyl compound) interaction and a hydrogen bond involving the acidic N–H proton. The fact that CO_2 may behave as both a Lewis acid and a Lewis base suggests that cooperative interactions, already pointed out in the case of 1:1 complexes (see the Introduction), might play a role in the case of 1:2 complexes. This is an important topic that needs to be addressed, and further work is being done in this direction.

The findings reported in this work open new perspectives in the design of green chemical procedures using supercritical CO_2 as a solvent, so far a challenging task due to solubility limitations.

References

1. Eckert CA, Knutson BL, Debenedetti PG (1996) Nature 383:313–318
2. Wells SL, Desimone J (2001) Angew Chem Int Educ 40:518–527
3. Mesiano AJ, Beckman EJ, Russell AJ (1999) Chem Rev 99:623–634
4. Jessop PG, Leitner W (1999) Chemical synthesis using supercritical fluids. Wiely-VCH Verlag GmbH & CO, KGaA, Weinheim
5. DeSimone JM, Guan Z, Elsbernd CS (1992) Science 257:945–947
6. Beckman EJ (2004) Chem Commun (17):1885–1888
7. Kazarian SG, Vincent MF, Bright FV, Liotta CL, Eckert CA (1996) J Am Chem Soc 118:1729–1736
8. Fink R, Hancu D, Valentine R, Beckman EJ (1999) J Phys Chem B 103:6441–6444

9. Yang H-J, Tian J, Kim H (2010) J Chem Eng Data 55:4130–4139
10. Liu J-F, Yang H-J, Wang W, Li Z (2008) J Chem Eng Data 53:2189–2192
11. Yang Z, Yang H-J, Tian J, Guo C-Y, Kim H (2011) J Chem Eng Data 56:1191–1196
12. Potluri VK, Xu J, Enick R, Beckman E, Hamilton AD (2002) Org Lett 4:2333–2335
13. Sarbu T, Styranec T, Beckman EJ (2000) Nature 405:165–168
14. Miller MB, Luebke DR, Enick RM (2010) Energy Fuels 24:6214–6219
15. Wang W, Yang H-J, Hu J, Guo C-Y (2009) J Supercrit Fluids 51:181–187
16. Miller DJ, Hawthorne SB, Clifford AA, Zhu S (1996) J Chem Eng Data 41:779–786
17. Mohamed A, Sagisaka M, Hollamby M, Rogers SE, Heenan RK, Dyer R, Eastoe J (2012) Langmuir 28:6299–6306
18. Raveendran P, Wallen SL (2002) J Am Chem Soc 124:7274–7275
19. Blatchford MA, Raveendran P, Wallen SL (2002) J Am Chem Soc 124:14818–14819
20. Danten Y, Tassaing T, Besnard M (2002) J Phys Chem A 106:11831–11840
21. Raveendran P, Wallen SL (2003) J Phys Chem B 107:1473–1477
22. Rivelino R (2008) J Phys Chem A 112:161–165
23. de Lange KM, Lane JR (2011) J Chem Phys 135:064304
24. Trung NT, Hung NP, Hue TT, Nguyen MT (2011) Phys Chem Chem Phys 13:14033–14042
25. Zhang X, Han X, Xu W (2011) J Theor Comput Chem 10:483–508
26. Altarsha M, Ingrosso F, Ruiz-Lopez MF (2012) Chem Phys Chem 13:3397–3403
27. Kim KH, Kim Y (2008) J Phys Chem A 112:1596–1603
28. Raveendran P, Wallen SL (2002) J Am Chem Soc 124:12590–12599
29. Alkorta I, Blanco F, Elguero J, Dobado JA, Ferrer SM, Vidal I (2009) J Phys Chem A 113:8387–8393
30. Altarsha M, Ingrosso F, Ruiz-Lopez MF (2012) J Phys Chem B 116:3982–3990
31. Scondo A, Dumarcay-Charbonnier F, Marsura A, Barth D (2009) J Supercrit Fluids 48:41–47
32. Møller C, Plesset MS (1934) Phys Rev 46:0618–0622
33. Pople JA, Headgordon M, Raghavachari K (1987) J Chem Phys 87:5968–5975
34. Bartlett RJ, Purvis GD (1978) Int J Quantum Chem 14:561–581
35. Zhao Y, Truhlar D (2008) Theoret Chem Acc 120:215–241
36. Chai J-D, Head-Gordon M (2008) Phys Chem Chem Phys 10:6615–6620
37. Reed AE, Curtiss LA, Weinhold F (1988) Chem Rev 88:899–926
38. Weinhold F, Carpenter JE (1988) In: Naaman R, Vager Z (eds) The structure of small molecules and ions. Plenum Press, New York, p 227
39. 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, Jr. JAM, 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 Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09, Revision B.01. Gaussian, Inc., Wallingford CT

REGULAR ARTICLE

Br$_2$ dissociation in water clusters: the catalytic role of water

**J. J. Santoyo-Flores · A. Cedillo ·
M. I. Bernal-Uruchurtu**

Abstract In this work, we explore the effect halogen
bonds and hydrogen bonds have on the activation of the
chemical bonds modified in the course of the Br$_2$ +
(H$_2$O)n → HOBr + HBr reaction and the effect the num-
ber of water molecules present in the cluster has on the
barrier and the stability of the products formed. Using
ab initio MP2 methods, we have analyzed the local elec-
trodonating and electroaccepting power functions that
revealed the differentiation between bromine atoms in Br$_2$
due to interaction with water molecules, thus resulting in
an increase in the donor power that stabilizes the proton
transfer process from a water molecule in the next step of
the reaction. This situation is confirmed by the lowering in
the corresponding activation barrier found. Our results
show that the catalytic effect of water is already important
for clusters with three water molecules and the effect
additional molecules have is related to the acid dissociation
of the products.

Keywords Halogen bond · Atmospheric chemistry ·
Halogens · Water clusters · Catalytic effects

J. J. Santoyo-Flores · M. I. Bernal-Uruchurtu (✉)
Centro de Investigaciones Químicas, Universidad Autónoma
del Estado de Morelos, Av. Universidad 1001,
62209 Cuernavaca, Morelos, México
e-mail: mabel@uaem.mx

A. Cedillo
Departamento de Química, Universidad Autónoma
Metropolitana-Iztapalapa, San Rafael Atlixco 186,
09340 México D. F., México

1 Introduction

The intermolecular forces acting between halogens and a
wide variety of molecules have attracted a significant
amount of attention in the recent years. From the nature of
the interaction [1–3], passing through their spectroscopic
properties [4], to their effect in supramolecular systems and
crystal engineering [5], several authors have commented
and explored the similarities between hydrogen bonds and
halogen bonds [6–8] aimed to design supramolecular sys-
tems with tunable forces since the donor, the halogen-
containing moiety, might be selected from a vast set of *R–
X* molecules and so the acceptor [9, 10]. However, little has
been said about the dihalogens in aqueous environments
where the simultaneous and concurrent presence of halo-
gen and hydrogen bonds might be of relevance for under-
standing water cluster mediated atmospheric chemistry. In
contrast, not only the catalytic effects of water but also the
hydration of halogenated species and halide ions has been
extensively studied with experimental and theoretical
methods [11–16].

Some years ago, a DFT study of halogens in water
clusters suggested the ionic pair formation as a conse-
quence of the charge separation between the two halogen
atoms in aqueous clusters [17]. It is difficult to predict the
ionic pair formation from the observed charge separation,
however, MP2 calculations confirmed the large polariza-
tion occurring in the halogen molecule leading to a partial
ionicity in the halogens, thus contributing to the strength of
the interactions playing for these structures [2]. The upper
limit of this ionicity, the ion pair formation, lies at the
bottom of the H$_2$O + Br$_2$ → HOBr + HBr reaction.

The chemistry of hipohalous acids (HOF, HOCl and
HOBr) is relevant for atmospheric phenomena where their
oxidative properties are involved in stratospheric processes

related to the consumption of ozone in this region [18, 19]. In particular, the bromine-containing systems have been found to be more reactive than the chlorine ones. Voegele et al. have investigated the reverse reactions, that is, $HOX + HY + nH_2O \rightarrow Br_2 + (H_2O)_{n+1}$, with $X = Cl, Br$ and $Y = Cl, Br$ using DFT and variational transition state theory with tunneling corrections. Their work concludes that barriers for the reaction with $X = Y = Br$ are consistently smaller than for chlorine and proceed barrierless when at least two water molecules catalyze the reaction [20, 21].

On the other hand, the hypobromous acid formation in liquid water is a process limited by two factors: first, the limited solubility of Br_2 in liquid water, 33.6 g L^{-1} and second, the small equilibrium constant for the reaction, $K = 5.2 \times 10^{-9}$. Nonetheless, the occurrence of this reaction can be easily observed by the spectroscopic absorption signal of the tribromide ion at 267 nm as described by Kerenskaya et al. [22]. The tribromide ion is present in solution as a byproduct of a series of equilibria as follows:

$$Br_2 + H_2O \leftrightarrow HBr + HOBr$$
$$HBr + H_2O \leftrightarrow H_3O^+ + Br^-$$
$$Br_2 + Br^- \leftrightarrow Br_3^-$$
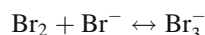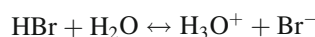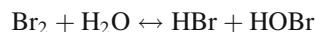
The intensity of the tribromide anion signal is quite large in the early stages of bromine clathrate formation and gets smaller as the amount of Br_2 available in solution decreases when the hydrate is formed. From the small equilibrium constant of the first step in these equilibria, it seems difficult to explain the significant amounts of Br_3^- spectroscopically found. This might suggest that the interaction between water and bromine during the early stages of hydrate formation might be somehow different from the one occurring in liquid solution.

In this work, we explore the effect halogen bonds and hydrogen bonds have on the activation of the chemical bonds modified in the course of the $Br_2 + (H_2O)_n \rightarrow HOBr + HBr$ reaction, and the effect the number of water molecules present in the cluster and the hydrogen-bond–halogen-bond network formed have on the barrier and the stability of the products formed.

2 Methodology

The electronic structure calculations reported in this work were obtained using the Gaussian98 suite of programs [23]. The optimized structures of the clusters used as models of the reactants are those reported in a previous work [2] where all structures were optimized at the MP2/aug-cc-pVDZ level. A harmonic frequency analysis was performed confirming that these stationary points correspond to minima or first-order saddle points in the potential

energy surface. In our previous work, the charge transfer and polarization analysis occurring in the clusters was done using the NBO-derived charges. This analysis is used here to understand the modifications on the charge distribution as a consequence of the rupture of the bromine–bromine bond.

In Fig. 1, we present the structure of the clusters selected in this work to study the Br–Br bond dissociation. In a previous work by one of us [2], these structures were fully characterized along several other configurations. A preliminary study on all those structures showed that the Br–Br dissociation depends in a crucial way on the intermolecular network established between the dihalogen and water molecules, that is, the hydrogen-bond–halogen-bond interactions. This behavior has been previously observed for some other proton-coupled activated processes in water clusters [24]. We briefly resume here some of the relevant properties of the clusters where the role of water in the dissociation process was found (See Table 1).

The structure with two water molecules (2W) corresponds to the most stable configuration found for this size of cluster, in it one of the halogen atoms is simultaneously engaged in both a halogen bond (XB) and a hydrogen–halogen (XH) bond with the two ends of a water dimer, a structure favored by the dipole–quadrupole interactions between Br_2 and the water molecules. A quite similar pattern can be observed for the structure with three water molecules (3W) resulting from one of the halogen atoms inserting into the water trimer ring. In this structure, cooperative effects are important; the XB and the XH distances are shorter than the analogous interactions for the 1:1 complex. The cluster with four water molecules used
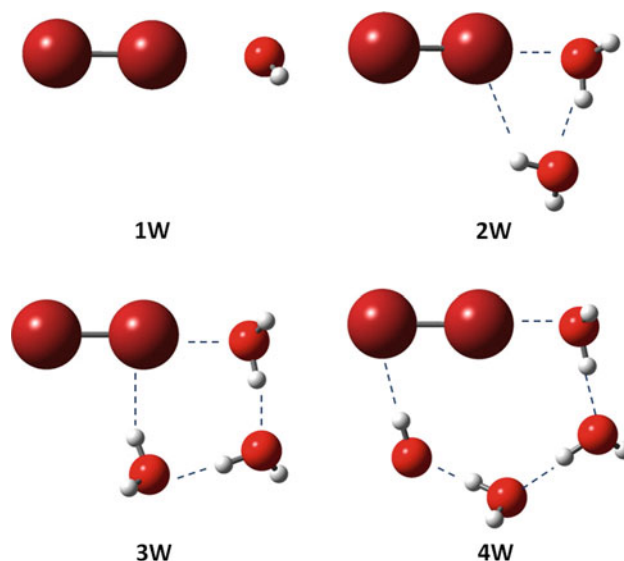

Fig. 1 MP2/aVDZ optimized structures of $Br_2(H_2O)n$ clusters used as reactant geometries (from Ref. [2])

Table 1 Selected properties of the Br$_2$–water clusters used in this work reported in Ref. [2]

Cluster		E_{int} (kJ/mol) MP2/aVDZ BSSE corrected	R(Br–Br)	r(Br···O) θ (Br–Br–O)	r(Br···H) θ (Br–Br–H)
1W	Br$_2$–H$_2$O	−15.5	2.338	2.769 (180)	
2W	Br$_2$–(H$_2$O)$_2$	−42.3	2.351	2.674 (175)	2.880 (124)
3W	Br$_2$–(H$_2$O)$_3$	−77.5	2.364	2.568 (173)	2.570 (155)
4W	Br$_2$–(H$_2$O)$_4$	−111.8	2.358	2.470 (179)	2.255 (176)

Distances are in Å and angles are in degrees

here (4W) is 2.1 kcal mol^{-1} less stable than the structure in which the halogen lies above a cyclic water cluster but results of particular relevance to this study due to the shorter *X*B interactions leading to a slightly larger charge transfer from the *X*B water molecule to the halogen as it will be further discussed.

The molecular electrostatic potential was obtained from the electronic density coming from MP2/aug-cc-pVDZ calculations. To better understand, the effect intermolecular forces have on the propensity of donating or receiving fractional amounts of charge on reactive processes as the one we are interested here, the local electroaccepting and electrodonating powers are calculated as $\omega^{\pm}(r) \equiv f^{\pm}(r)$ $(\mu^{\pm})^2/2\eta$, as proposed by Gázquez et al. [25]. Here, $f^{\pm}(r) \equiv \pm[\rho_{N\pm1}(r) - \rho_N(r)]$ are the acceptor/donor Fukui functions, where ρ_N is the electron density of the electronic system with N electrons; $\eta = \frac{1}{2}(I - A)$ is the chemical hardness, where I and A are the ionization potential and electron affinity, respectively, and $\mu^{\pm}$ are the chemical potentials for the process of accepting/donating charge and are given by $\mu^{+} \equiv \frac{1}{4}(3A + I)$ and $\mu^{-} \equiv \frac{1}{4}(3I + A)$, respectively.

3 Results and discussion

3.1 Charge separation and/or polarization as function of the number of water molecules in the cluster

In Table 2, we present the polarization occurring in the bromine calculated with two different methods. There is a reasonable agreement in the trends obtained with these methodologies: charge transfer (CT) from water to the dibromine molecule is small in contrast with the large internal charge transfer occurring in it. From the methods used, it is clear that Mulliken predicts a much larger polarization than NBO. The CT occurring along the Br–Br···O line is small, and it clearly shows the dependence on the Br···O distance as it has been suggested by some experimental studies [4, 26].

3.2 The donor acceptor behavior of molecules in the clusters

To gain a deeper understanding of the role each water molecule in the cluster has on the dibromine dissociation in

Table 2 Molecule polarization calculated as the difference of atomic charges on each atom (δq) and total charge q(Br$_2$) for bromine in the Br$_2$–(H$_2$O)$_n$ clusters in a.u. obtained using the density from the MP2/aug-cc-pVDZ level

Cluster	Mulliken		NBO	
	δq^a	q(Br$_2$)	δq	q(Br$_2$)
1W				
Br$_2$–H$_2$O	0.137	−0.028	0.078	−0.027
2W				
Br$_2$–(H$_2$O)$_2$	0.151	−0.034	0.095	−0.043
3W				
Br$_2$–(H$_2$O)$_3$	0.200	−0.063	0.132	−0.054
4W				
Br$_2$–(H$_2$O)$_4$	0.408	−0.081	0.082	−0.031

^{a} All reported values are in a.u.

these clusters, the local electrodonating and electroaccepting powers functions were calculated. In Fig. 2, we present the local indices of the [Br$_2$(H$_2$O)n] clusters projected onto a molecular surface. For the models 2W, 3W and 4W, one can notice a higher donating capacity on the bromine atom opposite to the oxygen–bromine interaction contact. This activation is due to the intramolecular charge redistribution process resulting from the interaction with the water cluster, which is a consequence of a small, but non-negligible, CT that contributes to the stability of these systems [26]. This differentiated increase in the donor power stabilizes the proton transfer process from a water molecule in the next step of the reaction. This effect is consistent with the lowering in the corresponding activation barrier, as it is shown below. In opposition, the local indices have a negligible value on the water molecules, except for that which interacts with the bromine molecule through its oxygen atom, which presents a marginal contribution.

3.3 The energetic cost of bromine dissociation in water clusters

The studied isodesmic reaction: Br$_2$ + H$_2$O → HOBr + HBr is an endothermic process according to the formation enthalpies differences between the products and the reactants. Accurate predictions of the heat of formation of the compounds involved in this reaction are available in the literature [27] except for HOBr. For this latter, it was only
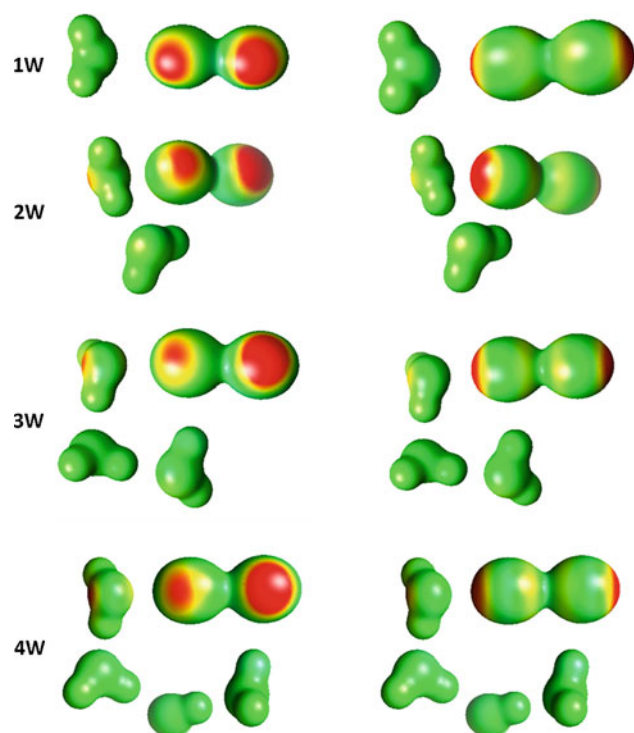
Fig. 2 Electrodonating (*right column*) and electroaccepting (*left column*) power functions projected on the electrostatic potential surface. In all cases, the *red* regions correspond to the maximum values of the analyzed property

few years ago when a refined set of ab initio calculations by Denis [28] converged to the experimentally obtained value of -259.07 kJ mol^{-1} that Lock et al. [29] obtained several years ago in photodissociation experiments for the heat of formation of HOBr. Using this value and the corresponding reported values for the other compounds [27], the heat of this reaction is found to be 112.68 kJ mol^{-1}. From the theoretical MP2/aug-cc-pVDZ level predictions of the heat of formation for the participating moieties, the heat of reaction is underestimated by 25 kJ mol^{-1} with respect to the value obtained using experimental data. We found that the energy difference is only 10.5 kJ mol^{-1} if a larger basis set (aug-cc-pVTZ) is used. Considering the computational cost of optimizing cluster geometries and exploring their PES with this latter basis set and the fact that the structure of each molecule, HBr, HOBr and H$_2$O as well as the interaction energies of the clusters obtained with the aug-cc-pVDZ basis set are in excellent agreement with the ones obtained with the larger basis set, we decided to perform this study with the aug-cc-pVDZ basis set.

To explore the region of the potential energy surface of the clusters involved in the reaction, we performed relaxed scans of the dibromine coordinate, that is, starting from the optimized structure of [Br$_2$(H$_2$O)$_n$] with $n = 1$–4, the bromine–bromine distance was stepwise increased from 2.33 to 3.15 Å and fixed, while the others degrees of freedom were fully relaxed. In Fig. 3, we present the approximate reaction profiles. These profiles do not correspond to the reaction coordinate but are helpful to locate the region of the transition state in processes where several covalent bonds are modified. As it can be seen in Fig. 3, the energetic cost of distorting the Br–Br covalent bond in these conditions is almost unaffected by the number of water molecules present in the cluster. However, it is clear that for the system 4W a different behavior emerges as the Br–Br distance is about 2.90 Å. With this information, we looked for the transition states structures starting from geometries in which this bond was smaller than that value.

a. [Br$_2$–H$_2$O]:

The TS structure for this system has a single imaginary frequency (v_i) at -855.5 cm^{-1} and is shown in Fig. 4a. At this level of calculation, the potential energy barrier for the bromine dissociation in presence of a single water molecule is 244.5 kJ mol^{-1}. This structure has a very long Br–Br distance confirming a late TS for the reaction. Furthermore, this structure closely matches the one described by Voegele et al. [21] for the reverse reaction where they found a 130.2 kJ mol^{-1} barrier at the MPW1K/6-31+G(d,p) level. The difference between the barriers for the forward and reverse reaction is a good match of the energy difference between reactants and products (114.3 kJ mol^{-1}), albeit the different methodologies used thus confirming the good performance of both approaches.

There is a large charge redistribution occurring in the system. If this geometry is considered to be closer to the reactants than to the products, the fragment corresponding to the original Br$_2$ molecule is negatively charged by 0.18 a.u. A similar analysis, now comparing with the expected products, shows that the fragments
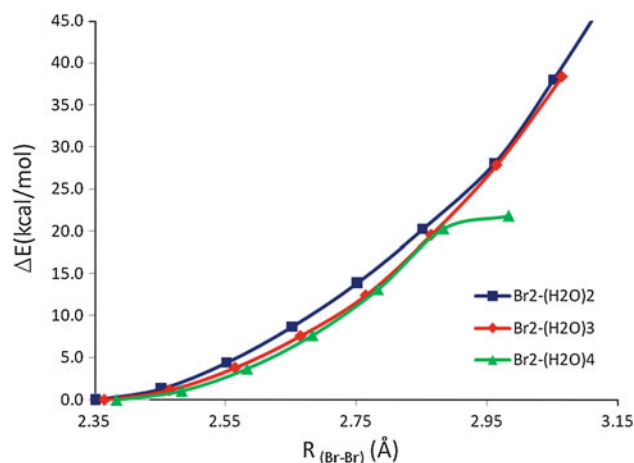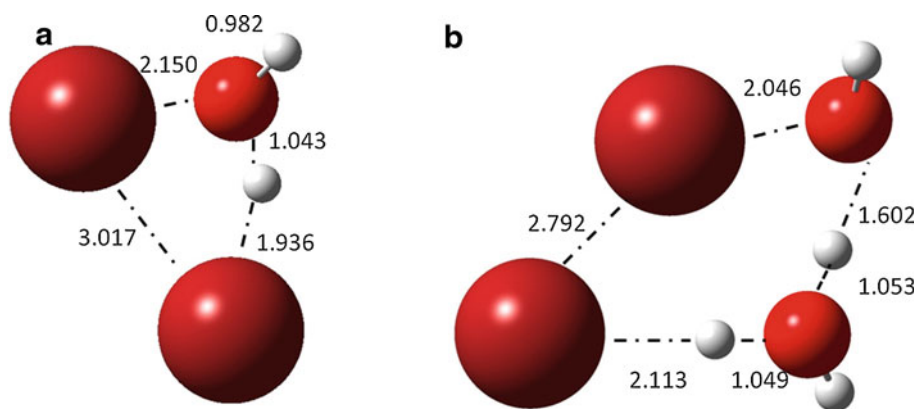

Fig. 3 Energetic profiles of the Br–Br elongation in the clusters. Each symbol corresponds to a value of the Br–Br distance where all other degrees of freedom was optimized

Fig. 4 Transition state structures found for the dissociation of Br$_2$ in clusters containing one (**a**) and two water (**b**) molecules. All distances are in Å

corresponding to HOBr and HBr are far from neutral, c.a. 0.3 a.u. It is possible to relate the high energetic cost of this reaction with the electrostatic work that must accompany the formation of these new moieties.

b. [Br$_2$(H$_2$O)$_2$]: The catalytic role of water molecules.
The transition state structure identified for the Br$_2$ + 2H$_2$O reaction ($v_i = -262.2$ cm^{-1}) lies 202.2 kJ mol^{-1} above the structure corresponding to the reactants. In Fig. 4b, it is possible to observe that in this TS structure, the role of one of the water molecules is to mediate the proton transfer between a water molecule and the anionic bromide fragment, a behavior that can be considered as catalytic. It is possible to understand the role of the catalyst as a charge mediator. In a similar way than the TS for the reaction with a single water molecule, in this case the fragments corresponding to the products, HOBr, HBr and H$_2$O, are all partially charged being the H$_2$O set the one bearing the largest charge, c.a. 0.4 a.u., whereas HBr and HOBr only 0.2 a.u. Furthermore, the distance between the two bromine atoms is considerably shorter than for the previous cluster thus facilitating the charge redistribution or delocalization.

For the reverse reaction in a previous work [21], two different TS structures were located, one corresponding to a catalyzed process with a barrier of 85 kJ mol^{-1} and another ascribed to an uncatalyzed reaction with a 134 kJ mol^{-1} barrier. The transition state structure found in this work is closer in geometry to the former and the combined barriers match again the change in enthalpy for the reaction. All efforts aimed to locate the so-called, uncatalyzed TS were unsuccessful, thus suggesting that the uncatalyzed channel is not as important as the other one.

c. [Br$_2$(H$_2$O)$_3$]: A stable products cluster is found.
The potential energy barrier for bromine dissociation in a three water molecule cluster was found to be 103.4 kJ mol^{-1}; this is at least 96 kJ mol^{-1} smaller than for the cluster with two water molecules. The transition state structure ($v_i = -214.7$ cm^{-1}), shown in Fig. 5a, is only 5.9 kJ mol^{-1} less stable than the cluster containing the products (Fig. 5b), in line with the idea advanced by Voegele et al., of a barrierless process in the reverse sense.
There is an interesting modification in the cluster as the reaction proceeds to products. As it is previously

Fig. 5 Transition state structure found for the dissociation of Br$_2$ in clusters containing three water molecules (**a**) and stable structure of the reaction products (**b**). All distances are in Å
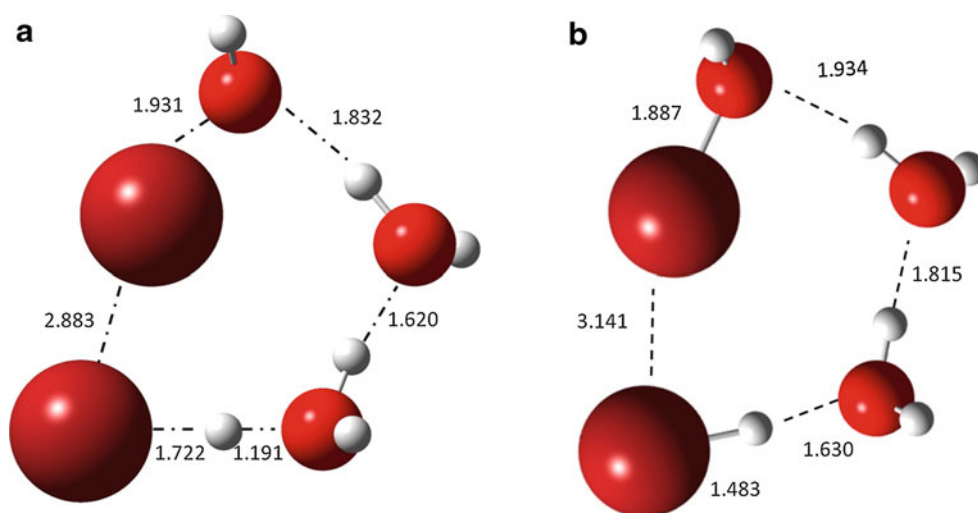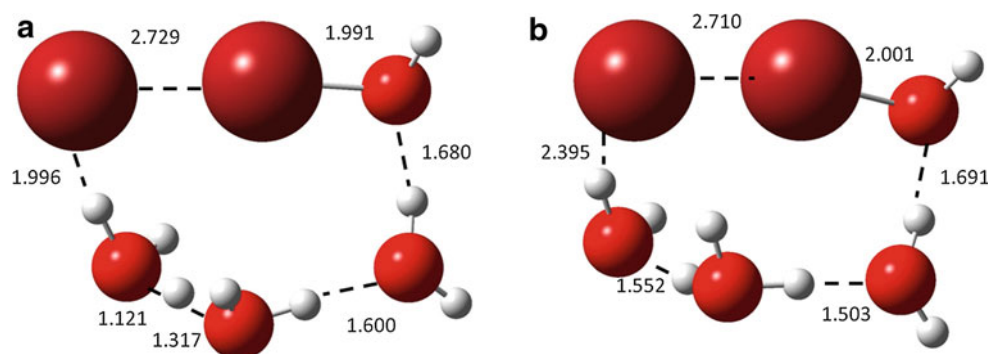
Fig. 6 Transition state structure found for the dissociation of Br_2 in clusters containing four water molecules (**a**) and stable structure of the reaction products (**b**). All distances are in Å

mentioned on the reactants structure, it is possible to see the same halogen atom involved in a XB and a XH interaction. This latter interaction breaks as soon as the bromine–bromine bond is slightly larger than the equilibrium distance. The proton formerly engaged in the XH interactions points to the anionic side of the halogen molecule forming a new XH interaction, thus enabling the concerted proton transfer. Again, the catalytic role of water molecules is due to their ability to separate the charges through their proton donor/acceptor properties and, in this case, the combined effect with their halogen acceptor ability. From the charge distribution of the TS structure, it is possible to see that the HBr and the HOBr fragments bear a small fractional charge, ~ 0.1 a.u.

d. $[Br_2(H_2O)_4]$: HBr is a strong acid in aqueous clusters. In Fig. 6, the structure of the transition state ($v_i = -125.6$ cm^{-1}) and products for the reaction are shown. In contrast to the structures for smaller clusters, the transition state is not very close in geometry to the products; the relative planarity of the initial reactant cluster is not conserved. This TS structure is -0.96 kJ mol^{-1} below the energy of the initial cluster, and the products lay 19.9 kJ mol^{-1} below this value. This is the only case in the series where the TS lies lower in energy than the reactant structure. It is possible that besides pointing toward a barrierless dissociation of dibromine, this barrier is more closely related with the acid/base equilibrium established between HBr and the water molecules than with the bromine dissociation itself. Several attempts to locate a precedent TS structure and/or possibly some intermediary structures were not successful possibly due to the flatness of the PES in that region, that is, the hydrogen bonds stabilizing this structure are very short thus promoting the proton delocalization between them. The TS structure (6a) is possible to observe that HBr dissociates in a bromide anion and a proton forming a hydronium, that is, a contact ion pair. This structure

evolves to a solvent separated ionic pair product structure through a proton migration to a neighbor water molecule. Goursot et al. [30] studied the behavior of HBr in 4 and 5 water molecule clusters finding that in all cases, four water molecules were needed to dissociate HBr and that contact ionic pair structures evolve to solvent separated ones being these latter slightly more stable for the cases with four and five water molecules. Our observations match not only qualitatively their results but also the energy difference between both ionic structures although our system only has three water molecules. This might be due to the presence of HOBr in the cluster and the role it plays in the charge redistribution.

The delocalization of the electronic density leads to a very polarized cluster, in which all fragments bear a partial charge being negatively charged Br–, HOBr and one water molecule and positively charged the hydronium ion and one water molecule. This situation is similar to what has been described for the dissociation of HCl in water clusters [31], there is a significant charge transfer within the cluster that prevents its breaking apart thus enabling the reverse reaction to readily take place.

4 Conclusions

In this work, we have looked into the effect local interactions have on the reactivity of bromine. Our results confirm that the combined effect of strong intermolecular interactions like halogen bonds and hydrogen bonds is behind the catalytic role of water molecules in this dissociation process. These interactions favor the process balancing the polarization of Br_2 in the reactant cluster and the charge transfer from water to it and later on act as mediators promoting the charge delocalization that favors the bond breaking of Br_2. Depending on the size of the cluster, the role of water molecules as proton donors and proton

acceptors are reinforced by the halogen bond and the hydrogen–halogen interactions presents in the cluster. The double catalytic role of water molecules has already been described for some other reactive processes in the gas phase [24, 32] and in the liquid [16].

In particular, the analysis of the local electrodonating and electroaccepting powers made evident that the interactions between the dihalogens and water molecules do not have a strong effect on the halogen. However, this interaction stabilizes the proton transfer process involved in the chemical reaction through an enhancement of the electrodonating power of the halogen-bonded water molecule.

An interesting finding is the dissociation of HBr in the presence of only three water molecules. Probably, this is due to the possibility of forming a H_3O^+ ion with a triple coordination to strongly polarized and partially charged water molecules. As found for some other reactive systems, the extent of formation of the acid, HBr in this case, depends sensitively on the arrangement of water molecules in the network. Furthermore, the charge transfer between the protonated water molecule and its first solvation shell might be a phenomenon closely related to the lack of barrier on the reverse sense, thus explaining the small equilibrium constant of this reaction. However, on the feasibility of this reaction occurring in the atmosphere, several considerations must be taken into account: first, the occurrence of this structures in the atmosphere is not entropically favored; second, photochemical processes might be relevant, either conducing to molecular water evaporation thus impeding the reaction to take place or by modifying the quantum tunneling rates in these clusters, situations that deserve to be carefully analyzed. The fact that this process is characterized by a very small equilibrium constant in liquid solution might point to a different environment around bromine. Work is in progress to address the hydration structure of bromine in condensed phases.

References

1. Politzer P, Murray JS, Concha MCJ (2008) Mol Model 14:659–665
2. Bernal-Uruchurtu MI, Hernández-Lamoneda R, Janda KC (2009) J Phys Chem A 113:5496–5505
3. Riley KE, Hobza P (2008) J Chem Theory Comput 4:232–242
4. Bernal-Uruchurtu MI, Kerenskaya G, Janda KC (2009) Int Rev Phys Chem 28:223–265
5. Metrangolo P, Resnati G (2008) Science 321:918–919
6. Lankau T, Wu YC, Zou JW, Yu CH (2008) J Theor Comput Chem 7:13–35
7. Politzer P, Murray JS, Lane P (2007) Int J Quantum Chem 107:3046–3052
8. Aakeroy CB, Fasulo M, Schultheiss N, Desper J, Moore C (2007) J Am Chem Soc 129:13772–13773
9. Brammer L, Espallargas GM, Libri S (2008) CrystEngComm 10:1712–1727
10. Parker AJ, Stewart J, Donald KJ, Parish CA (2012) J Am Chem Soc
11. Bondibey VE, Beyer MK (2001) Int Rev Phys Chem 21:277–306
12. Arshadi M, Yamdagni R, Kebarle P (1970) J Phys Chem 74:1475
13. Cooke SA, Cotti G, Evans CM, Holloway JH, Kisiel Z, Legon AC, Thumwood JMA (2001) Chem Eur J 7:2295–2305
14. Lu Y, Li H, Zhu X, Zhu W, Liu H (2011) J Phys Chem A 115:4467–4475
15. Ohmine I, Saito S (1999) Acc Chem Res 32:741–749
16. Rivail J-L, Antonczak S, Chipot C, Ruiz-López MF, Gorb LG (2009) Structure and reactivity in aqueous solution. American Chemical Society, Washington, DC, pp 154–167
17. Pathak AK, Mukherjee T, Maity DK (2008) J Phys Chem A 112:744–751
18. Yang X, Cox RA, Warwick NJ, Pyle JA, Carver GD, O'Connor FM, Savage NH (2005) J Geophys Res Atmos 110
19. Grenfell JL, Lehmann R, Mieth P, Langematz U, Steil B (2006) J Geophys Res Atmos 111
20. Voegele AF, Tautermann CS, Loerting T, Liedl KR (2002) J Phys Chem A 106:7850–7857
21. Voegele AF, Tautermann CS, Loerting T, Liedl KR (2003) Chem Phys Lett 372:569–576
22. Kerenskaya G, Goldschleger IU, Apkarian VA, Janda KC (2006) J Phys Chem A 110:13792–13798
23. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JJA, Stratmann RA, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi, R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Rega N, Salvador P, Dannenberg JJ, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Baboul AG, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Andres JL, Gonzalez C, Head-Gordon M, Replogle ES, Pople JA (2002) Revision A.11.3. Gaussian, Inc., Pittsburgh, PA
24. Relph RA, Guasco TL, Elliott BM, Kamrath MZ, McCoy AB, Steele RP, Schofield DP, Jordan KD, Viggiano AA, Ferguson EE, Johnson MA (2010) Science 327:308–312
25. Gázquez JL, Cedillo A, Vela A (2007) J Phys Chem A 111:1966–1970
26. Cappelletti D, Ronca E, Belpassi L, Tarantelli F, Pirani F (2012) Acc Chem Res
27. Haynes WM (ed) (2012) CRC Handbook of Chemistry and Physics, 92nd edn (Internet Version 2012). CRC Press/Taylor and Francis, Boca Raton, FL
28. Denis PA (2006) J Phys Chem A 110:5887–5892
29. Lock M, Barnes RJ, Sinha A (1996) J Phys Chem 100:7972–7980
30. Goursot A, Fischer G, Lovallo CC, Salahub DR (2005) Theor Chem Acc 114:115–123
31. Arillo-Flores OI, Bernal-Uruchurtu MI (2010) J Phys Chem A 114:8975–8983
32. Vaida V (2011) J Chem Phys 135:020901

REGULAR ARTICLE

Isodesmic reaction for pK_a calculations of common organic molecules

Sebastián Sastre · Rodrigo Casasnovas ·
Francisco Muñoz · Juan Frau

Abstract Three quantum chemistry methods (B3LYP, M05-2X and CBS-4B3*) have been used, in combination with SMD and CPCM continuum solvent models, to calculate the aqueous pK_a values of common organic compounds (aliphatic alcohols, carboxylic acids, amines, phenols, benzoic acids and pyridines) by using an isodesmic reaction. Good precision is found for all the studied functional groups, resulting mean absolute deviations of 0.5–1 pK_a units (equivalent to the best results obtained with thermodynamic cycles). It is worthy to note that no explicit water molecules were needed with the isodesmic reaction. In addition, the quality of the results is not strongly dependent on the combination of quantum chemistry method, solvent model and reference species. Therefore, the isodesmic reaction could be successfully used when dealing with gas-phase unstable species, with species that undergo large conformational changes between gas-phase and solution-phase or other difficult cases for the thermodynamic cycles.

Keywords pK_a calculation · Isodesmic reaction · Continuum solvent model

S. Sastre · R. Casasnovas · F. Muñoz · J. Frau (✉)
Institut d'Investigació en Ciènces de la Salut (IUNICS),
Departament de Química, Universitat de les Illes Balears,
07122 Palma de Mallorca, Spain
e-mail: juan.frau@uib.es

1 Introduction

Great efforts have been devoted during the last years to develop computational strategies for the accurate prediction of pK_a values [1].

The traditional approach in theoretical pK_a calculations is based on the use of thermodynamic cycles that combine gas-phase deprotonation free energies (ΔG_{gas}) and the solvation energies (ΔG_{solv}) of the involved species. The use of such cycles is due to the fact that it is not rigorously correct to obtain the free energy of deprotonation in solution (ΔG_{soln}) by calculating the free energies of the involved species according to the rigid rotor-harmonic approximation in the continuum solvent, as discussed previously in detail [2, 3]. In fact, the rigorous calculation would require much more costly simulation methods in which the solvent is explicitly considered, together with a free energy calculation procedure.

The use of thermodynamic cycles shows problems for those species that are gas-phase unstable or undergoes large conformational changes between gas-phase and solution-phase. However, it is possible to calculate the approximate free energies of deprotonation in solution by using the isodesmic reaction, since gas-phase energies are not required for the pK_a calculation (Scheme 1). Additionally, this reaction scheme, also known as relative pK_a calculation or proton exchange reaction, benefits from the

$$AH^q_{(soln)} + B^{m-1}_{(soln)} \xrightarrow{\Delta G_{(soln)}} A^{q-1}_{(soln)} + BH^m_{(soln)}$$

Scheme 1 Isodesmic reaction employed for the pKa calculation of AH^q. ΔG_{soln} stands for the free energy in solution of the acid–base reaction. The global charge of the acid species and its conjugate base are represented by q and $q-1$, respectively. The charges of the reference species and its conjugate base are represented by m and $m-1$, respectively

absence of the proton free energy of solvation, which is a potential source of error in the calculated free energies (2).

The pK_a of the acid AH can be calculated from Eqs. 1–3 and the experimental pK_a value of the reference species BH

$$\Delta G_{soln} = G_{soln}(BH^m) + G_{soln}(A^{q-1}) - G_{soln}(AH^q) \\ - G_{soln}(B^{m-1}) \tag{1}$$

$$G_{soln} = E_{soln} + G_{nes} + \Delta G_{corr\_soln} \tag{2}$$

$$pK_a(AH^q) = \frac{\Delta G_{soln}}{2,303RT} + pK_a(BH^m) \tag{3}$$

As previously noted, it is not possible to calculate the exact free energies in solution because the partition functions of the involved species (AH^q, BH^m, A^{q-1} and B^{m-1}) are unknown. Our group [3] recently proposed to calculate approximate free energies in solution under the assumption that gas-phase vibrational partition functions of the solutes as an approximation for the solution-phase and, secondly, that rotational and translational contributions to the free energies corresponding to conjugate acid and base species are of similar magnitude so, according to Eq. 1, they mostly cancel out. According to Eq. 2, the approximate free energies of each species in solution (G_{soln}) are calculated as the sum of the potential energy of the solute (E_{soln}) which includes the electric response of the continuum solvent, all the non-electrostatic contributions to the solute–solvent interaction (namely dispersion, repulsion and cavitation energies) which are all included in the term G_{nes}, and the contribution to the free energies from the vibrational motion of the nuclei at 298 K (ΔG_{corr_soln}).

It should be noted that both solvation energies of neutral species and gas-phase free energies (of neutral and charged species) can be determined with an accuracy of 1 kcal/mol [1], however, solvation energies of charged species calculated with continuum solvent models show, at best, average errors of 4 kcal/mol [4, 5]. According to the isodesmic reaction scheme, such errors should not be present in the pK_a calculation as solvation energies are not required. In any case, a higher accuracy in the pK_a calculation is expected for the isodesmic reaction since a good cancellation of errors is expected if both the reference species (BH^m) and the acid species (AH^q) present the same electric charge and similar structure.

So far the isodesmic reaction has mainly been used for the calculation of enthalpies of formation [6] but, concerning pK_a calculations, it has been less used than the thermodynamic cycles. Li et al. [7] reported that relative instead of absolute pK_a values are predicted with higher precision when combining ab initio methods and continuum solvent models for the study of methylimidazoles in aqueous solution. A recent combined experimental and theoretical study performed by Ruiz-López et al. [8] also

shows that remarkably low errors are obtained when using the isodesmic reaction for diprotic species. Govender and Cukrowski [9, 10] have employed this scheme, together with PCM solvent model and UA0 cavities for the calculation of the successive dissociation constants of nitrilopropanoic and nitrilotriacetic acids, obtaining errors of 0.2–3 pK_a units.

Our group [3] has recently employed the isodesmic reaction (Scheme 1) to calculate the pK_a of substituted pyridines and carbon acids. The results obtained for pyridines were equivalent to those of thermodynamic cycles [11] but no explicit water molecules were needed to increase the accuracy [3]. Concerning carbon acids, the mean absolute deviation of 2 pK_a units was also equivalent to the best results obtained with thermodynamic cycles [12, 13] with the advantage that the electric charge of the reference species was not determinant for a good precision [3]. We also reported the pK_a values of amino acids with very low errors (i.e., 0.22 and 0.19 pK_a units for the dissociation of carboxylic and amino groups, respectively) which are one of the paradigmatic difficult cases for thermodynamic cycles as the major tautomers in solution-phase and gas-phase are different [3].

The main objective in this work is to evaluate the robustness of the isodesmic reaction in the calculation of pK_a values of common acid–base organic functionalities and to compare such results with those calculated with thermodynamic cycles previously reported in the literature. Two different continuum solvent models were employed (namely CPCM [14–16] and SMD [17, 18]) in combination with three quantum chemistry methods, two common DFT functionals (B3LYP [19] and M05-2X [20, 21]) and the composite method CBS-4B3* [3, 22].

2 Methodology

The pK_a values of aliphatic alcohols, carboxylic acids, aliphatic amines, benzoic acids, phenols and pyridines calculated with the isodesmic reaction (Scheme 1) according to Eqs. 1–3 are displayed in Table 1.

The criterion followed to choose the reference species BH was the similitude of chemical structure with the studied species AH so; for alcohols, carboxylic acids and amines, the reference species are, respectively, ethanol, acetic acid and ethylamonium ion, while phenol, benzoic acid and pyridinium ion were the references for phenols, benzoic acids and pyridines, respectively.

Three quantum chemistry methods, specifically B3LYP [19], M05-2X [20, 21] and CBS-4B3* [3, 22], have been used to calculate the free energies of Eq. 2. In all cases, the geometries were optimized and characterized as energy minima by the absence of imaginary frequencies according

Table 1 Pyridines, aliphatic alcohols, carboxylic acids, amines, phenols and benzoic acids studied in this work

Pyridines	Alcohols	Carboxylic acids	Amines	Phenols	Benzoic acids
2-Methylpyridine	Methanol	Chloroacetic	Methylamine	*p*-Cyanophenol	*o*-Chlorobenzoic
3-Methylpyridine	2-Chloroethanol	Formic	Propylamine	*o*-Chlorophenol	*m*-Chlorobenzoic
4-Methylpyridine	Propanol	3-Chlorobutanoic	*i*-Propylamine	*m*-Cyanophenol	*p*-Chlorobenzoic
2,3-Dimethylpyridine	*i*-Propanol	Benzoic	Butylamine	*m*-Chlorophenol	*p*-Methylbenzoic
2,4-Dimethylpyridine	2-Butanol	4-Chlorobutanoic	2-Butylamine	*m*-Fluorophenol	*m*-Methylbenzoic
3-Bromopyridine	tert-butanol	Hexanoic	*tert*-Butylamine	*p*-Chlorophenol	*p*-Fluorobenzoic
3-Fluoropyridine		Propanoic	Trimethylamine	*p*-Fluorophenol	
3-Cyanopyridine		Pentanoic	Dimethylamine	*m*-Methylphenol	
3-Chloropyridine		Trimethylacetic		*p*-Methylphenol	
				o-Methylphenol	

to the 6-311++G(d,p) basis set. The CBS-4B3* composite method is a simplification of the CBS-QB3 method in which the CCSD(T) calculations and the energetic corrections of spin contamination and empirical corrections are suppressed [22]. This method benefits of lower computational costs and provides as accurate deprotonation free energies [22] and pK_a values [3] as the original CBS-QB3.

The solvent effects have been taken into account in all geometry optimizations and also energy calculations by using the CPCM model [13–15] with the UAKS cavities [23] as implemented in Gaussian 03 [23] and the SMD model [17, 18] as implemented in Gaussian 09 [24].

2.1 Results and discussion

Table 2 shows the mean absolute deviations (MAD) between the experimental and calculated pK_a values of each functional group according to the combination of solvent model (CPCM or SMD) and the quantum chemistry method (B3LYP, M05-2X or CBS-4B3*).

Several important points can be extracted after analyzing the values of Table 2. First of all, it should be noted that no combination of quantum chemistry method and continuum solvent model stand out by providing much lower errors than the other combinations. In fact, the highest difference between MAD values is 0.67 pK_a units for the CPCM/M05-2X and SMD/M05-2X calculations of aliphatic alcohols.

Concerning all the organic functionalities, the MAD values are lower than 1 pKa units in all cases with the exception of the pK_a values predicted with SMD for the aliphatic alcohols, which show MAD values lower than 1.3 pK_a units. The best results correspond to the primary amines and benzoic acids with MAD values lower than 0.35 (Table T3 of Supporting Information) and 0.5 (Table 2) pK_a units, respectively, for the SMD calculations.

Table 2 Mean absolute deviations (MAD), standard deviation (SD) and maximum absolute deviation (AD max) of aqueous pK_a values calculated using different methods (CBS-4B3*/6-311++G(d,p), B3LYP/6-311++G(d,p) and M05-2X/6-311++G(d,p)), and solvent models (CPCM and SMD)

	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Pyridines						
MAD	0.57	0.75	0.83	0.62	0.80	0.78
SD	0.48	0.74	0.55	0.36	0.39	0.44
AD max	1.44	2.17	1.87	1.29	1.34	1.55
Alcohols						
MAD	0.85	0.68	0.62	1.20	1.01	1.29
SD	0.49	0.46	0.51	0.96	0.75	0.97
AD max	1.54	1.19	1.22	2.84	2.27	2.85
Carboxylic acids						
MAD	0.78	0.98	0.79	0.57	0.64	0.67
SD	0.42	0.36	0.47	0.36	0.54	0.59
AD max	1.32	1.44	1.51	1.13	1.54	1.89
Amines						
MAD	0.87	0.72	0.90	0.24	0.35	0.27
SD	0.84	0.74	1.00	0.19	0.25	0.20
AD max	2.51	2.20	2.89	0.63	0.83	0.71
Phenols						
MAD	0.90	1.08	1.02	0.70	0.89	0.87
SD	0.57	0.68	0.68	0.48	0.57	0.56
AD max	2.16	2.71	2.57	1.66	2.18	2.10
Benzoic acids						
MAD	0.45	0.57	0.41	0.41	0.50	0.34
SD	0.29	0.41	0.30	0.35	0.48	0.29
AD max	0.95	1.36	0.82	1.05	1.39	0.69

53

Considering the MAD values obtained with both CPCM and SMD solvent models, the best results correspond to the SMD model for carboxylic acids, amines, phenols and benzoic acids independently of the quantum chemistry method, whereas CPCM provides lower MAD values for the aliphatic alcohols. Finally, the results of both solvent models for pyridines and primary amines (Table T3 of Supporting Information) show no significant differences.

Regarding the quantum chemistry method, the best results are provided by the CBS-4B3* method for 7 out of 12 combinations, specifically for pyridines, carboxylic acids and phenols with CPCM (MAD values of 0.57, 0.78 and 0.90 pK_a units, respectively) and pyridines, carboxylic acids, amines and phenols with SMD (MAD values of 0.62, 0.57, 0.24 and 0.70 pK_a units, respectively). The M05-2X gives the best results for the alcohols and benzoic acids with CPCM and also benzoic acids with SMD, while the B3LYP functional gives the best results for the amines with CPCM and alcohols with SMD (Table 2).

As noted, the pK_a values obtained with the isodesmic reaction exhibit low errors for all the functionalities, independently of the combination of method and solvent model so, we considered important to compare our results with previously reported theoretical pK_a values calculated with thermodynamic cycles.

Theoretical pK_a values of aliphatic alcohols have been reported in several studies [25–28] among which stand out the one of Pliego and Riveros [25], which reports the pK_a values of methanol and ethanol with MAD values of 0.5 pK_a units with the IPCM model and including 2–3 explicit water molecules. Namazian and Heidary [26] obtained the pK_a of methanol, ethanol, propanol and isopropanol with maximum errors lower than 0.5 pK_a units by using the CPCM model to calculate the solvation energies. In close agreement with such studies, MAD values of 0.62 pK_a units are obtained for the CPCM/M05-2X calculations when using the isodesmic reaction (Table 2).

Carboxylic acids have been also extensively studied in theoretical pK_a calculations [28–38]. Namazian and Halvani [32] calculated the pK_a of 66 acids with the PCM model and B3LYP functional and obtained average errors of 0.5 pK_a units. In another study, Toth et al. [31] calculated the pK_a values of five carboxylic acids at the CPCM/HF/6-31+G(d)//HF/6-31G(d) level with average errors of 0.5 pK_a units. Such average errors are similar to those obtained in this study with the SMD solvent model (Table 2).

The pK_a values of amines were calculated with good accuracy [39–43]. For example, average errors of 0.46 pK_a units calculated with the PCM model at the B3LYP/6-31+G*//MP2/6-31+G* level including an explicit water molecule were reported in the work of Behjatmanesh-Ardakani et al. [42]. In the present work, an accuracy of ∼0.3 pK_a units is found for the CPCM model when considering just the primary amines (Table T3 of Supporting Information), although higher errors are obtained when including secondary and tertiary amines (Table 2). Nevertheless, the SMD model provides, in all cases, MAD errors lower to 0.3 pK_a units (Table 2).

Scmidt am Busch and Knapp [37] calculated the pK_a of benzoic, p-methylbenzoic, m-methylbenzoic and p-chlorobenzoic acids with MAD values of 0.5 pK_a units. As seen in Table 2, the errors obtained from the isodesmic reaction are even lower, particularly those of SMD/M05-2X calculations (i.e., 0.34 pK_a units).

The theoretical prediction of pK_a values of substituted phenols has also been reported in different studies [28, 37, 44, 45]. Liptak et al. [44] reported RMSD errors between 0.4 and 3.9 pK_a units depending on the calculated solvation energies. As can be seen in Table 2, in our case, the MAD values are slightly over 0.5 pK_a units and under 1 pK_a unit for most of the method and solvent model combinations.

Concerning pyridines, mean deviations lower than 1 pK_a unit were reported [11, 46–50]. In a previous work of our group [11], we evaluated the accuracy of different thermodynamic cycles, the importance of explicit water molecules and the influence of using gas-phase or solution-phase optimized geometries when calculating the solvation energies with the CPCM model. In the best cases, mean errors of 0.5 pK_a units were reported when using a single explicit water molecule [11]. The results provided by the isodesmic reaction (Table 2) show the same precision than the best obtained with thermodynamic cycles (i.e., MAD values between 0.57 and 0.83 pK_a units) without requiring explicit solvent molecules.

The lowest MAD values obtained with each solvent model, together with other MAD values of previous works that used thermodynamic cycles are depicted in Fig. 1. As can be seen, the accuracy of the isodesmic reaction predicted pK_a values is of the same order of those predicted with thermodynamic cycles for pyridines, benzoic acids and carboxylic acids for both CPCM and SMD models. Concerning amines, the SMD/CBS-4B3* calculations outperform the errors of thermodynamic cycles. Besides, in the worst-case scenario, the difference between MAD values of isodesmic reaction and thermodynamic cycles is lower than 0.6 pK_a units.

We have chosen the carboxylic acids and amines to evaluate the influence of the reference species on the calculated pK_a values. Tables 3 and 4 show the MAD values obtained when using different references for carboxylic acids and amines, respectively. The resulting MAD values for carboxylic acids vary between 0.5 and 1.0 pK_a units unless chloroacetic and formic acids are used as reference species. Regarding the amines, the MAD values fluctuate between 0.28 and 0.78 pK_a units regardless of the reference is a primary, secondary or tertiary amine.

Fig. 1 Comparison between the best mean absolute deviations (MAD) obtained with CPCM and SMD solvent models and MAD values reported in previous studies that used thermodynamic cycles
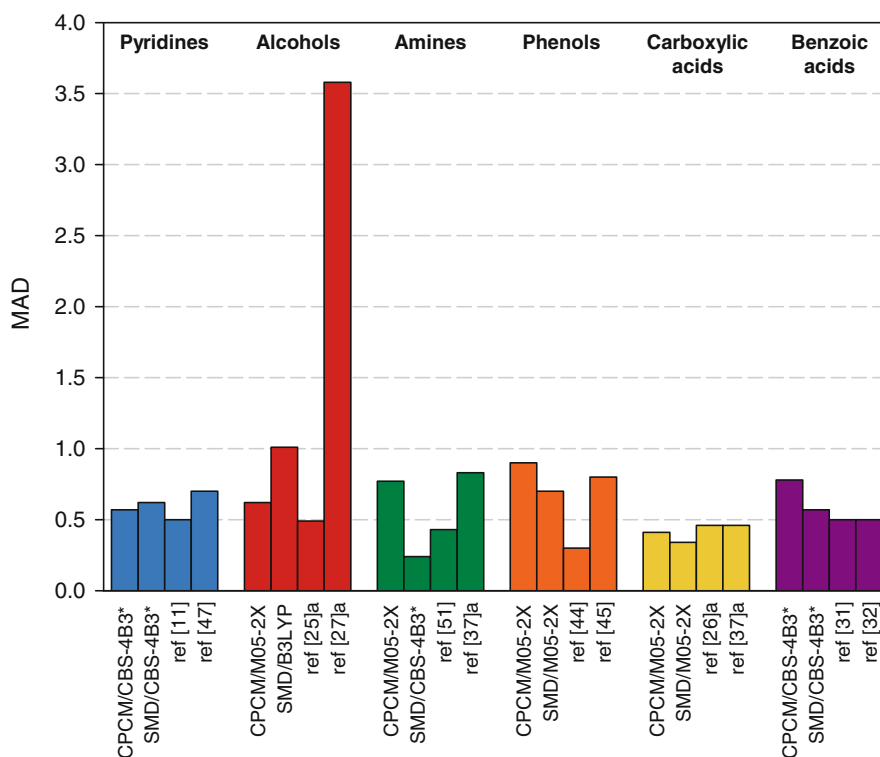
[a] MADs calculated with the species belonging to the functional group of interest.

Rebollar-Zepeda and Galano [51] had recently reported a comprehensive assessment of reaction schemes and density functionals for the pK_a calculation of amines. Their evaluation of the isodesmic reaction points out that the precision of the calculated pK_a values is strongly dependent on the reference species [51]. It is well known that the

Table 3 Effect of the reference species on the accuracy of the calculated pK_a values of carboxylic acids expressed as mean absolute deviations (MAD)

Reference	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Acetic ac.	0.78	0.98	0.79	0.57	0.64	0.67
Chloroacetic ac.	1.45	1.72	1.50	0.84	1.42	1.70
Formic ac.	1.11	1.51	1.13	1.02	1.45	1.17
3-Chlorobutanoic ac.	0.74	1.25	0.80	0.57	0.66	0.79
Benzoic ac.	0.86	0.89	0.86	0.78	0.64	0.88
4-Chlorobutanoic ac.	0.74	0.89	0.79	0.58	0.64	0.67
Hexanoic ac.	0.88	1.09	0.98	0.69	0.82	0.72
Propanoic ac.	0.84	1.00	0.89	0.91	0.96	0.82
Pentanoic ac.	0.88	1.06	0.92	0.66	0.73	0.74
Trimethylacetic ac.	1.07	1.25	1.31	0.87	1.01	0.81

acidity of amines is significantly influenced by the solvent accessibility to the amino group. Since continuum solvent models do not reproduce the solvent structure around the solute, it is not possible to model the steric effects on the solvation of the amino group caused by its substituents. Consequently, not only the electrostatic characteristics have to be taken into account for the selection of the reference species but also the steric behavior of the substituents should be similar to the studied amine species. However, it is not trivial to model such effects by the introduction of explicit water molecules in the calculation because the dynamic contribution of the substituents to the solvation structure would require statistical treatment by monte carlo or molecular dynamics simulations.

To get a deeper insight into the influence of the reference species, the pK_a of some functional groups has been calculated by using a structurally different species (Table 5). Specifically, the pK_a of benzoic acids, phenols and pyridines were calculated by considering acetic acid, ethanol and ethylamine, respectively, as reference species. In the first case, MAD values of 0.5 pK_a units result when using acetic acid which is similar to the precision obtained if the reference species were the benzoic acid. However, the errors in the pK_a values of phenols are significantly larger (i.e., MAD values between 4.0 and 6.5 pK_a units) when using ethanol as reference species. Such different behavior for benzoic acids and phenols can be attributed to

Table 4 Effect of the reference species on the accuracy of the calculated pK_a values of amines expressed as mean absolute deviations (MAD)

Reference	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Ethylamine	0.87	0.72	0.90	0.24	0.35	0.27
Methylamine	0.61	0.93	0.80	0.61	0.93	0.80
Propylamine	0.35	0.32	0.17	0.35	0.32	0.17
i-Propylamine	0.23	0.34	0.38	0.23	0.34	0.38
Butylamine	0.28	0.35	0.48	0.28	0.35	0.48
2-Butylamine	0.21	0.35	0.22	0.21	0.35	0.22
tert-Butylamine	0.31	0.63	0.30	0.31	0.63	0.30
Trimethylamine	0.27	0.41	0.27	0.27	0.41	0.27
Dimethylamine	0.36	0.58	0.33	0.36	0.58	0.33

the delocalization of the negative charge to the π-system in the phenoxyde species, which causes significant differences in the electrostatic interactions of the deprotonated oxygens of phenoxydes and aliphatic alcoxydes. However, delocalization of the negative charge in benzoate anions is low so the charge distribution of carboxylate groups in benzoates is similar to that of aliphatic carboxylates. When using ethylamine as reference, the MAD values in the pK_a of pyridines increase up to 2.8 pK_a units for the CPCM calculations but remain closer to the MAD values when using pyridine as reference for the SMD calculations (Table 5). Therefore, it is worthy to note that the chemical structure should not be the only criterion to consider when choosing the reference species and that the cancellation of errors due to similarities in the solute–continuum interactions is not always trivial.

Table 5 Effect of the reference species on the accuracy of the calculated pK_a values of aromatic molecules expressed as mean absolute deviations (MAD)

Reference	CPCM			SMD		
	CBS-4B3*	B3LYP	M05-2X	CBS-4B3*	B3LYP	M05-2X
Benzoic acids						
Benzoic acid	0.45	0.57	0.41	0.41	0.50	0.34
Acetic acid	0.51	0.47	0.48	1.16	0.65	1.16
Phenols						
Phenol	0.90	1.08	1.02	0.70	0.89	0.87
Ethanol	4.01	4.31	5.01	5.21	5.59	6.46
Pyridines						
Pyridine	0.57	0.75	0.83	0.62	0.80	0.78
Ethylamine	1.86	2.77	1.92	1.22	0.82	1.42

As a suggestion of one of the referees, we considered the inclusion of the rotational and translational contributions, calculated from gas-phase partition functions, to the free energies in solution. Such contributions had a minor effect of 0.01 to 0.1 pK_a units for all the studied functional groups. Only for two species, methanol and methylamine, these energies had higher and significant contribution to the pK_a values (i.e., 0.54 and 0.43 pK_a units, respectively). However, it has to be noted that this result only stresses the fact that rotational and translational free energies of the conjugate acid–base pairs calculated with gas-phase partition functions are very similar and, therefore, largely cancel out in Eq. 1. In addition, this result does not entail the actual rotational and translational free energies of the conjugate acid–base species to be necessarily similar.

In summary, it has been shown that the isodesmic reaction scheme provides reliable results in the pK_a calculation of common organic functionalities, with pK_a values comprised between 1 and 19. The accuracy of the isodesmic reaction predicted pK_a values with combinations of common quantum chemistry methods and continuum solvent models is similar to that provided by thermodynamic cycles. Besides, no explicit water molecules are required to obtain a good accuracy. Although the reference species influences the precision of the calculated pK_a values, the cancellation of errors in the isodesmic reaction allows more flexible criteria in the choice of the reference species. So, taken everything into account, it is worth to consider this procedure for theoretical pK_a predictions, especially for those cases where the thermodynamic cycles show problems related to gas-phase calculations.

References

1. Ho J, Coote ML (2010) A universal approach for continuum solvent pK_a calculations: are we there yet? Theor Chem Acc 125:3–21
2. Ho J, Klamt A, Coote ML (2010) Comment on the correct use of continuum solvent models. J Phys Chem A 114:13442–13444
3. Casasnovas R, Fernández D, Ortega-Castro J, Frau J, Donoso J, Muñoz F (2011) Avoiding gas-phase calculations in theoretical pK_a predictions. Theor Chem Acc 130:1–13
4. 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

5. Cramer CJ, Truhlar DG (2008) A universal approach to solvation modeling. Acc Chem Res 41:760–768

6. Wang L, Heard DE, Pilling MJ, Seakins P (2008) A Gaussian-3X prediction on the enthalpies of formation of chlorinated phenols and Dibenzo-p-dioxins. J Phys Chem A 112:1832–1840

7. Li GS, Ruiz-López MF, Maigret B (1997) Ab initio study of 4(5)-Methylimidazole in aqueous solution. J Phys Chem A 101:7885–7892

8. Derbel N, Clarot I, Mourer M, Regnouf-de-Vains J, Ruiz-López MF (2012) Intramolecular interactions versus hydration effects on p-Guanidinoethyl-phenol structure and pK_a values. J Phys Chem A 116:9404–9411

9. Govender KK, Cukrowski I (2009) Density functional theory in prediction of four stepwise protonation constants for nitrilotri-propanoic acid (NTPA). J Phys Chem A 113:3639–3647

10. Govender KK, Cukrowski I (2010) Density functional theory and isodesmic reaction based prediction of four stepwise protonation constants, as log $K_H^{(n)}$, for nitrilotriacetic acid. the importance of a kind and protonated form of a reference molecule used. J Phys Chem A 114:1868–1878

11. Casasnovas R, Frau J, Ortega-Castro J, Salvà A, Donoso J, Muñoz F (2009) Absolute and relative pK_a calculations of mono and diprotic pyridines by quantum methods. J Mol Struct Theochem 912:5–12

12. Gao DQ, Svoronos P, Wong PK, Maddalena D, Hwang J, Walker H (2005) pK_a of acetate in water: a computational study. J Phys Chem A 109:10776–10785

13. Ho J, Coote ML (2009) pKa calculation of some biologically important carbon acids—an assessment of contemporary theoretical procedures. J Chem Theory Comput 5:295–306

14. 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

15. Barone V, Cossi M, Tomassi J (1998) Geometry optimization of molecular structures in solution by the polarizable continuum model. J Comput Chem 19:404–417

16. Cossi M, Rega M, Scalmani G, Barone V (2003) Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. J Comput Chem 24:669–681

17. Marenich AV, Cramer CJ, Truhlar DG (2009) Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. J Phys Chem B 113:6378–6396

18. Marenich AV, Cramer CJ, Truhlar DG (2009) Performance of SM6, SM8, and SMD on the SAMPL1 test set for the prediction of small-molecule solvation free energies. J Phys Chem B 113:4538–4543

19. Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys 98:5648–5652

20. Zhao Y, Schultz NE, Truhlar DG (2006) Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. J Chem Theory Comput 2:364–382

21. Zhao Y, Truhlar DG (2008) Functionals with broad applicability in chemistry. Acc Chem Res 41:157–167

22. Casasnovas R, Frau J, Ortega-Castro J, Salvà A, Donoso J, Muñoz F (2010) Simplification of the CBS-QB3 method for predicting gas-phase deprotonation free energies. Int J Quantum Chem 110:323–330

23. Gaussian 03, Revision E.01, Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery Jr JA, Vreven T, 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, Bakken V, 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 D K, 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 D J, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill P MW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA, Gaussian, Inc., Wallingford CT, (2004)

24. Gaussian 09, Revision B.01, 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 Jr JA, 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 Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ, Gaussian, Inc., Wallingford CT (2009)

25. Pliego JR, Riveros JM (2002) Theoretical calculation of pK_a using the cluster-continuum model. J Phys Chem A 106:7434–7439

26. Namazian M, Heidary H (2003) Ab initio calculations of pK_a values of some organic acids in aqueous solution. J Mol Struct-Theochem 620:257–263

27. Silva CO, da Silva EC, Nascimento MAC (2000) Ab initio calculations of absolute pK_a values in aqueous solution II. Aliphatic alcohols, thiols, and halogenated carboxylic acids. J Phys Chem A 104:2402–2409

28. 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

29. Liptak MD, Shields GC (2001) Accurate pK_a calculations for carboxylic acids using complete basis set and Gaussian-n models combined with CPCM continuum solvation methods. J Am Chem Soc 123:7314–7319

30. Liptak MD, Shields GC (2001) Experimentation with different thermodynamic cycles used for pK_a calculations on carboxylic acids using complete basis set and gaussian-n models combined with CPCM continuum solvation methods. Int J Quantum Chem 85:727–741

31. Toth AM, Liptak MD, Phillips DL, Shields GC (2001) Accurate relative pK_a calculations for carboxylic acids using complete basis set and Gaussian-n models combined with continuum solvation methods. J Chem Phys 114:4595–4606

32. Namazian M, Halvani S (2006) Calculations of pK_a values of carboxylic acids in aqueous solution using density functional theory. J Chem Thermodyn 38:1495–1502

33. Namazian M, Halvani S, Noorbala MR (2004) Density functional theory response to the calculations of pK_a values of some carboxylic acids in aqueous solution. J Mol Struct Theochem 711:13–18

34. Schuurmann G, Cossi M, Barone V, Tomasi J (1998) Prediction of the pK_a of carboxylic acids using the ab initio continuum-solvation model PCM-UAHF. J Phys Chem A 102:6706–6712

35. da Silva CO, da Silva EC, Nascimento MAC (1999) Ab initio calculations of absolute pK_a values in aqueous solution I. Carboxylic acids. J Phys Chem A 103:11194–11199

36. Saracino GAA, Improta R, Barone V (2003) Absolute pK_a determination for carboxylic acids using density functional theory and the polarizable continuum model. Chem Phys Lett 373: 411–415

37. Schmidt am Busch M, Knapp EW (2004) Accurate pK_a determination for a heterogeneous group of organic molecules. Chem Phys Chem 5:1513–1522

38. Jia Z, Du D, Zhou Z, Zhang A, Hou R (2007) Accurate pK_a determinations for some organic acids using an extended cluster method. Chem Phys Lett 439:374–380

39. da Silva EF, Svendsen HF (2003) Prediction of the pK_a values of amines using ab initio methods and free-energy perturbations. Ind Eng Chem Res 42:4414–4421

40. Eckert F, Klamt A (2006) Accurate prediction of basicity in aqueous solution with COSMO-RS. J Comput Chem 27:11–19

41. Blanco SE, Ferretti FH (2005) Calculation of ionization constants of methylamines in aqueous solution. J Mol Struc Theochem 722: 197–202

42. Behjatmanesh-Ardakani R, Karimi MA, Ebady A (2009) Cavity shape effect on pK_a prediction of small amines. J Mol Struc Theochem 910:99–103

43. Khalili F, Henni A, East ALL (2009) Entropy contributions in pK_a computation: application to alkanolamines and piperazines. J Mol Struc Theochem 916:1–9

44. Liptak MD, Gross KC, Seybold PG, Feldgus S, Shields GC (2002) Absolute pK_a determinations for substituted phenols. J Am Chem Soc 124:6421–6427

45. Rebollar-Zepeda AM, Campos-Hernández T, Ramirez MT, Rojas-Hernández A, Galano A (2011) Searching for computational strategies to accurately predict pK_as of large phenolic derivatives. J Chem Theory Comput 7:2528–2538

46. Caballero NA, Melendez FJ, Muñoz-Caro C, Niño A (2006) Theoretical prediction of relative and absolute pK_a values of aminopyridines. Biophys Chem 124:155–160

47. Chen IJ, MacKerell AD (2000) Computation of the influence of chemical substitution on the pK_a of pyridine using semiempirical and ab initio methods. Theo Chem Acc 103:483–494

48. Guven A (2005) Acidity study on 3-substituted pyridines. Int J Mol Sci 6:257–275

49. Blanco SE, Ferretti FH (2005) Calculation of acidity constants of pyridines in aqueous solution. Chem Phys Lett 403:400–404

50. Ögretir C, Tay NF, Özturk II (2007) A theoretical study on protonation of some halogen substituted pyridine derivatives. J Mol Graph Model 26:740–747

51. Rebollar-Zepeda AM, Galano A (2012) First principles calculations of pk_a values of amines in aqueous solution: application to neurotransmitters. Int J Quantum Chem 112:3449–3460

Cooperativity of hydrogen and halogen bond interactions

Sławomir J. Grabowski

Abstract The cooperativity effects in the $Cl^-\cdots HCCH$ $\cdots HF$, $Cl^-\cdots ClCCH\cdots HF$ and $F^-\cdots ClCCH\cdots HF$ complexes are analyzed here. The results show that the formation of the hydrogen and halogen bonds is ruled by the same mechanisms and that the cooperativity enhances these interactions. The MP2(full)/6-311 ++G(d,p) calculations were performed for the above triads and the corresponding sub-units; dyads linked by the hydrogen or halogen bonds and monomers. The NEDA scheme of the decomposition of the interaction energy was applied here. It was found that for the halogen bonded systems, the most important is the polarization term of the energy of interaction while for the hydrogen bonds the charge transfer interaction energy and next the electrostatic contribution. The interaction between orbitals is also analyzed here in terms of the Natural Bond Orbitals method.

Keywords Halogen bond · Hydrogen bond · Non-covalent interaction · Bent's rule · Natural bond orbitals method · NEDA scheme · The electron charge distribution

S. J. Grabowski
Faculty of Chemistry, University of the Basque Country
UPV/EHU and Donostia International Physics Center (DIPC),
P.K. 1072, 20080 Donostia, Euskadi, Spain

S. J. Grabowski (✉)
IKERBASQUE, Basque Foundation for Science,
48011 Bilbao, Euskadi, Spain
e-mail: s.grabowski@ikerbasque.org

1 Introduction

The hydrogen bond is a very well-known phenomenon often analyzed and discussed due to its importance in numerous chemical, physical and biological processes [1–4]. However, in recent years, the increase of studies on other non-covalent interactions is observed as for example hydride bond [5–8], dihydrogen bond [9–12], halogen bond [13–16], etc. There are various approaches for the uniform description of all non-covalent interactions; different studies are concentrated on differences and similarities between them. It seems that the σ-hole concept [16–18] explains characteristics and mechanisms of formation of the halogen bond and numerous other non-covalent interactions. According to that concept, some of non-covalent interactions exist owing to the contact between two centers characterized by the opposite electrostatic potentials. The σ-hole concept [16] explains why the halogen atom (usually connected with the carbon atom) may act at the same time as the Lewis base center forming connections through such interactions as for example hydrogen or halide bond and as the Lewis acid center forming for example halogen bond—the interaction analyzed intensively in recent years. Such dual character of the halogen atom is manifested in the positive electrostatic potential along the extension of the A–X bond (X designates Cl, Br or I halogen atom and A is the Lewis acid center—usually the carbon atom) and a belt of negative electrostatic potential around the X-atom, i.e. in the direction perpendicular, or nearly so, to the A–X bond.

The natural bond orbitals (NBO) method [19–21] is often used to characterize non-covalent interactions. It was pointed out that the A–H⋯B hydrogen bond phenomenon is a combination of two effects [22]. The hyperconjugative interaction is connected with the electron charge transfer

from the B-lone pair to the σ*(A–H) antibonding orbital. It leads to the increase of the population of the antibonding σ* orbital and in consequence of the weakening and elongation of the A–H bond. The second effect, the rehybridization, leads to the increase in s-character of A-hybrid orbital in the A-H bond. It is in line with Bent′s rule [23] which states that atoms maximize the s-character in hybrid orbitals aimed toward electropositive substituents and the atoms minimize such character (maximize the p-character) toward more electronegative substituents. It was pointed out that both effects, the increase in s-character of A-orbital and the hyperconjugation, are characteristic for all hydrogen bonds [22]. It was found very recently that the mechanism of the formation of numerous non-covalent interactions may be described by the balance of the effects described above [24, 25]. In particular, the halogen bond formation is ruled by the mechanisms which are similar to those of the hydrogen bond interaction. For example, if the A–Y bond of the Lewis acid moiety interacts with the B center of the Lewis base (the A–Y···B interaction), then the complexation leads to the increase in s-character of A-hybrid orbital in the A–Y bond and to the increase of the polarization of this bond (Scheme 1). For example, this concerns the O–H, C–Cl and N–H bonds of the O–H···O hydrogen bond, C–Cl···O halogen bond and N–H···H–Be dihydrogen bond, respectively. The reverse situation is observed if the B–Y bond of the Lewis base participates in the B–Y···A interaction. For the latter, the complexation leads to the decrease in s-character of B hybrid orbital in the B–Y bond and to the decrease of the polarization of this bond (Scheme 1). For example, this is observed for the C–Cl, Be–H and Mg–H bonds in the C–Cl···H–O hydrogen bond, Be–H···H–N dihydrogen bond and Mg–H···Na$^+$ hydride bond, respectively (Scheme 1). Such changes in the s-character and polarization are in line with Bent's rule since the formation of A–Y···B is connected with the increase of the positive charge of Y-atom and the formation of B–Y···A with the decrease of this charge. In other words, Y-atom after complexation is less or more electronegative as it was before.

The cooperativity effect is another phenomenon often observed for hydrogen bonded systems [1–4] and analyzed in recent years for the other non-covalent interactions [26–29]. One can mention numerous studies on cooperativity effects in H-bonded complexes such as those on the H_3N···HF and H_3N···HF···HF complexes which were investigated with the use of experimental microwave and theoretical ab initio techniques [30], on the chain of moieties connected by OH bonds [2], the analysis of two-dimensional cyclic networks containing three-centered hydrogen bond interactions [31], amide hydrogen bonding chains [32], intramolecular bifurcated hydrogen bonds [33], in homo (HCN)$_n$ and hetero (HCN)$_n$···HF chains [34], the decomposition of the interaction energy performed for
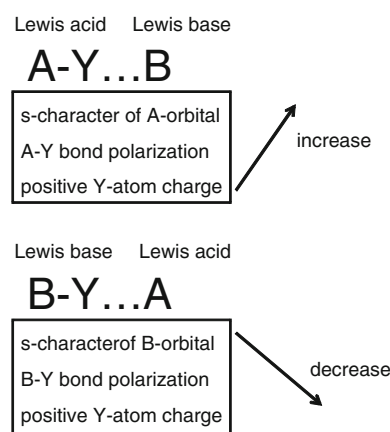
Scheme 1 The changes of the Lewis acid and Lewis base units as a result of complexation

$H_2C=O$···(H–F)$_n$ clusters (n up to 9) [35] and numerous others. In a case of other non-covalent interactions one can mention the studies on cooperativity effects between halogen and hydrogen bond [36, 37], between halogen and dihydrogen bond [38], between halogen-hydride and dihydrogen bond [38] or between halogen bonds [39]. There are numerous others, this is worth to mention that there is the review article where cooperativity in multiple weak bonds is considered for different pairs of non-covalent interactions [40].

Hence, the aim of this study is to compare two phenomena, from one side the cooperativity between halogen bond and hydrogen bond and from the other side the cooperativity between two hydrogen bonds for comparison. The goal of this study is also the presentation of parameters which are analyzed rarely for non-covalent interactions and they are not applied for the analysis of cooperativity effects, for example, the mentioned earlier here s-character and the polarization of the corresponding bond participating in the non-covalent interaction. The cooperativity effect is also discussed here in terms of Bent's rule.

2 Computational details

The MP2(full)/6-311++G(d,p) calculations for the complexes analyzed here were performed with the Gaussian09 set of codes [41]. The optimizations led to energetic minima since no imaginary frequencies were observed. The following triads are considered here: FH···HCCH···Cl$^-$, FH···HCCCl···Cl$^-$ and FH···HCCCl···F$^-$. For the first complex, there are two hydrogen bonds F–H···π and C–H···Cl$^-$. For the next two complexes, there is the cooperativity between hydrogen bond and halogen bond since the F–H···π and C–Cl···Cl(F)$^-$ interactions exist. The full optimizations were also performed for all monomers of

complexes considered here and for the following dyads: FH⋯HCCH (F–H⋯π hydrogen bond), FH⋯HCCCl (F–H⋯π hydrogen bond), HCCH⋯Cl$^-$ (C–H⋯Cl$^-$ hydrogen bond), HCCCl⋯Cl$^-$ (C–Cl⋯Cl$^-$ halogen bond) and HCCCl⋯F$^-$ (C–Cl⋯F$^-$ halogen bond).

It is worth to mention that complexes related to the hydrogen and halogen bonded dyads analyzed here were often analyzed experimentally. Numerous A–H⋯π [3, 4] hydrogen bonds and C–X⋯B halogen bonds [14] were found in crystal structures. The gas phase studies were also performed on such systems, for example, the microwave spectrum of the acetylene–HF complex in the ground vibrational state has been assigned using a pulsed Fourier transform microwave spectrometer [42]. However, it is interesting to know if the triads considered here have any experimental analogs. The Cambridge Structural Database (CSD) [43] search was performed here for the A–H⋯Y–C ≡ C–X1⋯X2 fragment (A is the non-metal atom, Y is any substituent, X1 and X2 designate halogen atoms). In other words, the fragments where the A–H⋯π hydrogen bond and C–X1⋯X2 halogen bond exist were searched. The following search criteria were used: no disordered structures, no structures with unresolved errors, no powder structures and $R \leq 7$ %. Only two fragments of two crystal structures fit these criteria, i.e. C–H⋯C≡C–Br⋯Br$^-$ for the tetraphenylphosphonium tetrakis(1-bromo-2-phenyla-cetylene) bromide structure [44] and the C–H⋯C≡C–I⋯Cl$^-$ fragment for the structure of bis(bis(Ethylenedi-thio)tetrathiafulvalene) 2,5-dimethyl-1,4-bis(iodoethynyl) benzene dichloride [45]. However, the e.s.d.'s values for CC bonds for these structures amount 0.01 Å and more. The Cambridge Structural Database (CSD) search was also performed for the A–H⋯Y–C≡C–H⋯X2 fragment related to the system analyzed here (the difference between this fragment and the previous one is that the H-atom is inserted here instead of the X1 atom). The same search criteria were fixed here as for the former fragment and only 15 entries were found. However, the e.s.d.'s for CC bonds are smaller than 0.005 Å only in 7 cases. The number of fragments found in CSD and the accuracy of the related structures are not sufficient to perform the statistical analysis. One can see that triads related to those analyzed theoretically here are not common in the crystal structures (note that the number of crystal structures collected in CSD is close to 6,00,000). However, these findings show that the real interactions are analyzed in this study.

The energies expressed by Eqs. 1–3 were calculated here according to the earlier studies on cooperativity effects [36].

$$\Delta E'_{AB} = E_{ABC} - (E_A + E_{BC}) \qquad (1)$$

$$\Delta E'_{BC} = E_{ABC} - (E_C + E_{AB}) \qquad (2)$$

$$\Delta E_{ABC} = E_{ABC} - (E_A + E_B + E_C) \qquad (3)$$

The subscripts A, B and C designate Cl(F)$^-$ ion, H–C≡C–Cl(–H) and H–F molecules, respectively. Hence, $\Delta E_{AB}'$ and $\Delta E_{BC}'$ correspond to the interaction energies in Cl(F)$^-$⋯Cl(H)–CCH and FH⋯HCCH (F–H⋯π), respectively. However, these energies correspond to geometries of sub-units (dyads) in the triads. E_{ABC} is the energy of the triad, and ΔE_{ABC}, the interaction energy in the triad. All systems (ABC, AB, BC, A, B and C) were optimized separately thus the corresponding interaction energies include the deformation effects connected with the complexation [46]. All energies were corrected for BSSE [47, 48]. The interaction energies for the optimized dyads are also considered here, and they were calculated according to the following relations.

$$\Delta E_{AB} = E_{AB} - (E_A + E_B) \qquad (4)$$

$$\Delta E_{BC} = E_{BC} - (E_B + E_C) \qquad (5)$$

Also for these energies, the BSSE correction and the deformation energy mentioned above are included. The energetic cooperativity parameter has been calculated according to the Eq. 6 proposed before by Alkorta and co-workers [40].

$$E_{coop} = \Delta E_{ABC} - (\Delta E_{AB} + \Delta E_{BC} + \Delta E_{AC}) \qquad (6)$$

Here, as before, ΔE_{ABC} is the interaction energy in the triad, ΔE_{AB} and ΔE_{BC} are the interaction energies of the isolated dyads in their minima configurations (Eqs. 4 and 5) and ΔE_{AC} is the interaction energy between A and C in the geometry they possess in the corresponding triad. One can see that for the last term, the unreal system is considered since there are two molecules here separated by the B spacer. For the energy expressed by Eq. 6, the BSSE correction is also taken into account since all right-hand-side energies of this equation are corrected for BSSE.

The Natural Bond Orbitals (NBO) method [19, 20] was applied to analyze interactions in the triads and dyads analyzed here. For the A–H⋯B hydrogen bond, the $n_B \rightarrow \sigma^*_{AH}$ is often considered as its characteristic interaction. It is related to the maximum $n_B \rightarrow \sigma^*_{AH}$ overlapping, n_B designates the lone electron pair of the proton acceptor while σ^*_{AH} is an antibonding orbital of the A–H proton donating bond. The $n_B \rightarrow \sigma^*_{AH}$ interaction is calculated as the second-order perturbation theory energy (Eq. 7):

$$\Delta E(n_B \rightarrow \sigma^*_{AH}) = -2\langle n_B|F|\sigma^*_{AH}\rangle/(\varepsilon(\sigma^*_{AH}) - \varepsilon(n_B)) \qquad (7)$$

$\langle n_B|F|\sigma^*_{AH}\rangle$ designates the Fock matrix element and $(\varepsilon(\sigma^*_{AH})-\varepsilon(n_B))$ is the orbital energy difference. Different non-covalent interactions are analyzed here thus the latter

equation may be modified to take into account the A–H···π hydrogen bond (Eq. 8).

$$\Delta E\left(\pi_B \rightarrow \sigma_{AH}^*\right) = -2\left\langle \pi_B | F | \sigma_{AH}^* \right\rangle^2 / \left(\varepsilon(\sigma_{AH}^*) - \varepsilon(\pi_B)\right) \tag{8}$$

or to consider the halogen bond interaction (Eq. 9).

$$\Delta E\left(n_B \rightarrow \sigma_{AX}^*\right) = -2\left\langle n_B | F | \sigma_{AX}^* \right\rangle^2 / \left(\varepsilon(\sigma_{AX}^*) - \varepsilon(n_B)\right) \tag{9}$$

where AX designates the C–Cl bond since the C–Cl···Cl$^-$ and C–Cl···F$^-$ interactions are considered here. The energies expressed by Eqs. 7–9 will be designated later here as E_{NBO}.

The polarization of bonds is calculated in this study within the NBO scheme in the following way. The natural bond orbitals (NBOs) are combinations of the natural hybrid orbitals (NHOs) [20]. For example, the NBO for a localized σ bond between atoms A and B, σ_{AB}, is formed from directed orthogonal hybrids (NHOs) h_A and h_B.

$$\sigma_{AB} = c_A h_A + c_B h_B \tag{10}$$

The natural hybrids in turn are composed from a set of effective valence-shell atomic orbitals, natural atomic orbitals (NAOs). The percentage of the NBO on each hybrid, $100|c_A|^2$, may be calculated. This is the percentage of the electron density on the A-atom, and it is treated as the polarization of the A–B bond.

The natural energy decomposition analysis (NEDA) [49–51] is applied here. The NEDA interaction energy partitioning is expressed by Eq. 11.

$$\Delta E^{HF} = ES + PL + CT + EX + DEF \tag{11}$$

For the AB complex of two interacting moieties, A and B, Eq. 11 may be described in the following way. ΔE^{HF} is the Hartree–Fock interaction energy. ES corresponds to the classical electrostatic interaction. PL, polarization term, arises from the extra electrostatic interaction connected with the polarization of the unperturbed molecular orbitals of the separated fragments (A and B) to those of the complex (AB). CT, the charge transfer contribution, is the stabilizing component which arises from the delocalization of the electrons between A and B components of the complex. The exchange component, EX, is an attractive contribution to the interaction energy arising from the exchange interactions of electrons on A with those on B. DEF designates the deformation energy; DEF has a contribution from each fragment.

$$DEF = DEF(A) + DEF(B) \tag{12}$$

where

$$DEF(A) = E\left(\psi_A^{def}\right) - E(\psi_A) \tag{13}$$

DEF(A) is the repulsive (positive) quantity since ψ_A the wave function converged for A is of lower energy than ψ_A^{def} in the complex.

The NEDA decomposition (Eq. 11) may be expressed in the other form (Eq. 14).

$$\Delta E^{HF} = EL + CT + CORE \tag{14}$$

where:

$$EL = ES + PL + SE \tag{15}$$

and

$$CORE = EX + DEF - SE \tag{16}$$

EL is named as the electrical interaction, SE is a portion of the DEF component arising from the self-polarization energies of the monomers; this is the cost to form the induced charge density.

The NEDA scheme was extended [52] to treat charge densities from the density functional theory (DFT) calculations. In this extended approach (DFT/NEDA), the intermolecular exchange component (EX) of HF method is replaced by the exchange–correlation (XC) component. The XC component accounts for intermolecular electron exchange and correlation. The DFT interaction energy within DFT/NEDA scheme may be expressed in the following form.

$$\Delta E^{DFT} = ES + PL + CT + XC + DEF \tag{17}$$

Both approaches of the decomposition were applied in this study (Eqs. 11 and 17). For DFT/NEDA calculations, the B3LYP functional was applied here.

For NBO, NEDA and DFT/NEDA calculations, the NBO 5.0 program [53] incorporated into GAMESS set of codes [54] was used.

3 Results and discussion

3.1 Geometrical and energetic parameters

Figure 1 presents the dyads linked by the C–H(Cl)···Cl(F)$^-$ interactions. Figure 2 shows the HCCH···HF and HCCCl···HF complexes linked by the F–H···π hydrogen bond. Figures 3, 4, 5 present the triads analyzed in this study. The HCCH, HCCCl and HF monomers are presented in these figures as well as parameters discussed later here. Table 1 presents selected geometrical parameters of the complexes mentioned above. One can observe the positive (favorable) cooperativity for all triads. For example, the H···Cl$^-$ intermolecular distance for the Cl$^-$···HCCH dyad amounts 2.299 Å while this distance for the Cl$^-$···HCCH···HF triad is equal to 2.172 Å. Such shortening of the intermolecular distance is usually attributed to

the strengthening of the interaction; here, it concerns the cooperativity between two hydrogen bonds. The Cl···Cl⁻ distance in the Cl⁻···ClCCH complex is equal to 3.146 Å and it decreases to 3.063 Å in the Cl⁻···ClCCH···HF triad. The latter decrease is the result of the cooperativity between the halogen and hydrogen bonds. Similarly, for the halogen bond in the F⁻···ClCCH and F⁻···ClCCH···HF complexes, the Cl···F⁻ distance amounts 2.461 and 2.385 Å, respectively. One can see that for the hydrogen and halogen bonds, the additional F–H···π hydrogen bond enhances the former interaction. One can also conclude that similar cooperativity processes are observed for the hydrogen and halogen bonds. The same effects may be observed from the other side. The H···C distances for F–H···π interactions in HCCH···HF and HCCCl···HF dyads decrease in the corresponding triads: Cl⁻···HCCH···HF and Cl(F)⁻···ClCCH···HF, respectively (see Table 1).

The elongation of the F–H and C–H proton donating bonds is observed in the H-bonded dyads if they are compared with the corresponding monomers. This elongation is greater in the corresponding triads (Table 1). The slight elongation of the C–Cl bond in the Cl(F)⁻···ClCCH dyads is observed as an effect of the halogen bond formation. Practically, there are no changes of the C–Cl bond lengths for the Cl(F)⁻···ClCCH···HF triads if they are
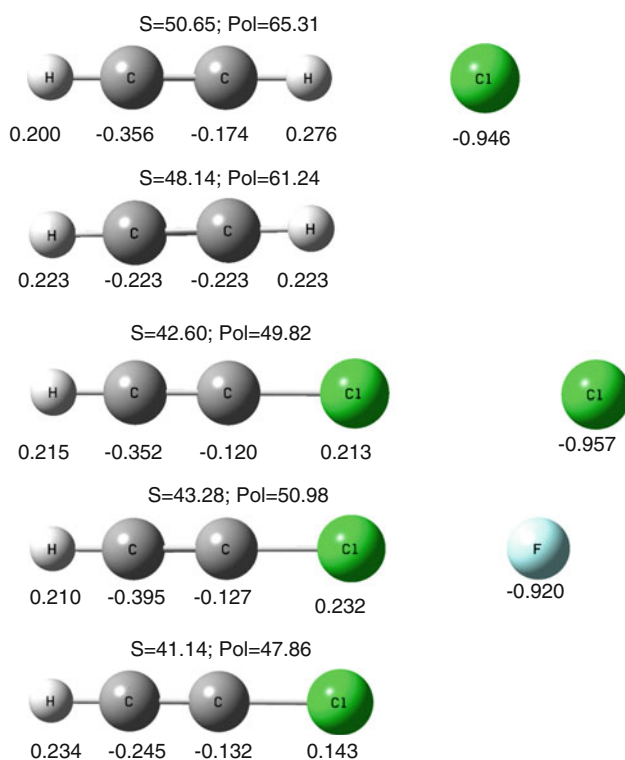

Fig. 1 The s-character in the C-hybrid orbital (s), the polarization of the C–H(Cl) bond (Pol) and the NBO atomic charges for the following dyads and monomers: HCCH···Cl⁻, HCCH, HCCCl···Cl⁻, HCCCl···F⁻ and HCCCl

compared with the corresponding dyads. Even for the Cl⁻···ClCCH···HF triad, the slight shortening of the C–Cl bond is observed in comparison with the ClCCH monomer.

The energies presented in Table 2 are in agreement with the changes of the geometrical parameters discussed earlier here. The interactions corresponding to the hydrogen and halogen bonds in the Cl(F)⁻···H(Cl)CCH dyads (see E_{AB}) are enhanced in the corresponding triads by approximately 4–5 kcal/mol (see E_{AB}'). Similarly, the F–H···π hydrogen bonds in the HCCCl(H)···HF dyads are enhanced in the corresponding triads (see E_{BC} and E_{BC}', respectively). For the F⁻···ClCCH···HF complex, the strongest interactions are observed (E_{ABC}). For the latter system, the most favorable cooperativity effect is observed since the cooperativity energy (E_{coop}) is equal to −2.5 kcal/mol while for two remaining triads this value is equal to −0.6 and to −1.4 kcal/mol.

For all dyads and triads considered here, the decomposition of the Hartree–Fock interaction energy (the NEDA scheme) and the decomposition of the DFT interaction energy (DFT/NEDA) were performed (Table 3). One can see few tendencies for both approaches, NEDA and DFT/NEDA. For the dyads linked through hydrogen bonds, i.e. Cl⁻···HCCH, FH···HCCH and FH···HCCCl, the most important attractive interaction energy term is the charge transfer energy, CT (see Eqs. 11 and 17 for NEDA and DFT/NEDA approaches). The latter term is often attributed in the literature to the covalency of the hydrogen bond interaction [19, 55]. Sometimes it is stated that all hydrogen bonds are covalent in nature since the CT energy is characteristic for such interactions [19, 56]. The next most important attractive term for the above-mentioned hydrogen bonded complexes is the electrostatic energy, ES. There is the similar situation for the Cl⁻···HCCH···HF complex with two hydrogen bonds where both CT and ES terms are the most important attractive contributions.

For the Cl⁻···ClCCH and F⁻···ClCCH halogen bonded systems, the polarization interaction energy is the most important attractive term. It is in line with the previous studies where the importance of this kind of interaction was claimed for the halogen bonds [24]. Also for two triads Cl⁻···ClCCH···HF and F⁻···ClCCH···HF where there is cooperativity between hydrogen and halogen bond, the polarization term is the most important attractive interaction.

The principal differences in the HF and DFT analysis, NEDA and DFT/NEDA, respectively, appear in EX and XC components. The XC accounts for intermolecular electron exchange and correlation; when electron correlation is neglected in the HF method, this term reduces to the exchange component, EX. One can observe the enhancement of the B3LYP interaction (ΔE, Table 3) if it is compared with the HF energy. The greatest differences

Fig. 2 The s-character in the F-hybrid orbital (s), the polarization of the H–F bond (Pol) and the NBO atomic charges for the following dyads and monomer: HCCH···HF, HCCCl···HF and HF

Fig. 3 The s-character in the F-hybrid orbital (s), the s-character in the C-hybrid orbital (s), the polarization of the H–F bond (Pol), the polarization of the C–H bond (Pol) and the NBO atomic charges for the Cl⁻···HCCH···HF trimer

Fig. 5 The s-character in the F-hybrid orbital (s), the s-character in the C-hybrid orbital (s), the polarization of the H–F bond (Pol), the polarization of the C–Cl bond (Pol) and the NBO atomic charges for the F⁻···ClCCH···HF trimer
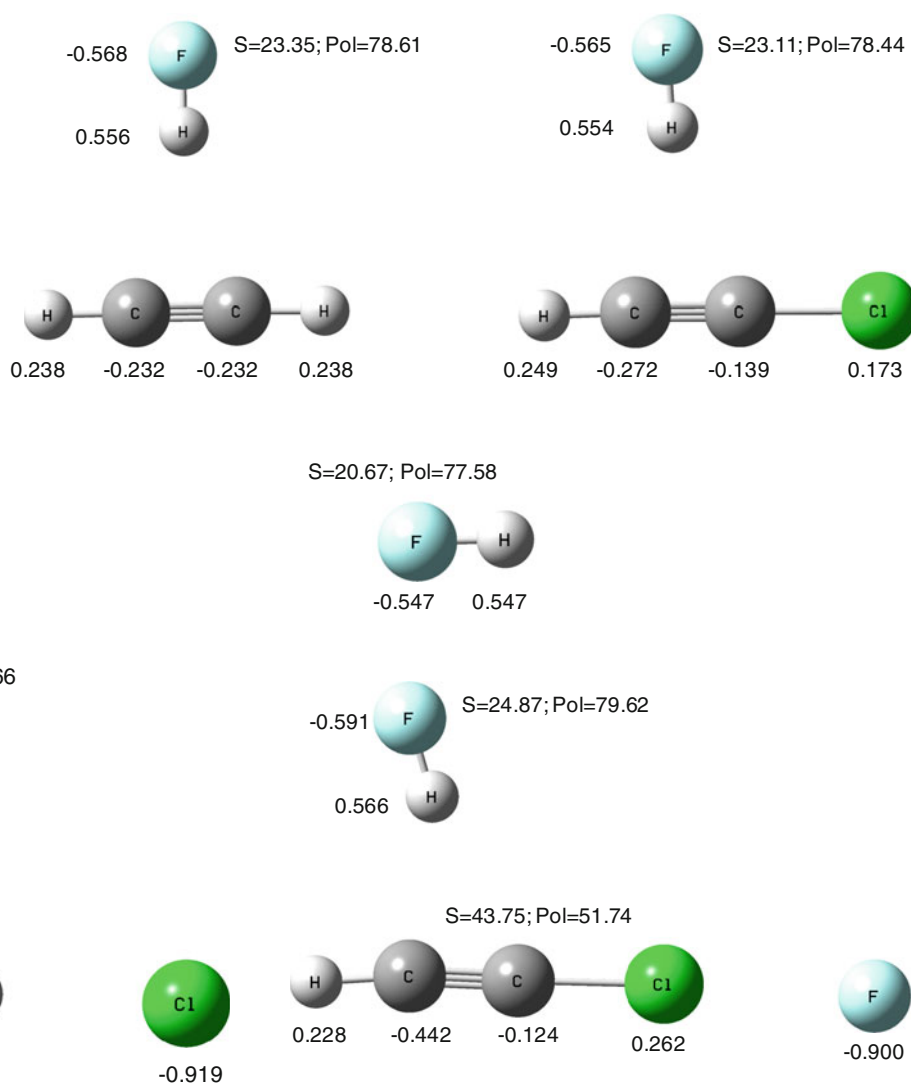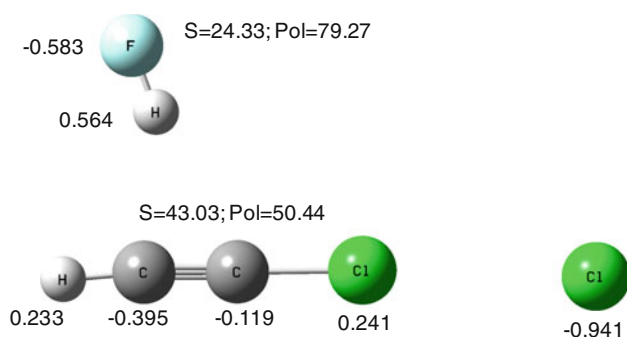
Fig. 4 The s-character in the F-hybrid orbital (s), the s-character in the C-hybrid orbital (s), the polarization of the H–F bond (Pol), the polarization of the C–Cl bond (Pol) and the NBO atomic charges for the Cl⁻···ClCCH···HF trimer

Table 1 Geometrical parameters (distances in Å) of the complexes considered in this study

Complex	H(Cl)···Cl(F)	C–H(Cl)	H···C	H–F
Cl⁻···HCCH	2.299	1.090	–	–
Cl⁻···ClCCH	3.146	1.642	–	–
F⁻···ClCCH	2.461	1.652	–	–
HCCH···HF	–	–	2.267	0.923
HCCCl···HF	–	–	2.227; 2.232	0.923
Cl⁻···HCCH···HF	2.172	1.101	2.146; 2.249	0.930
Cl⁻···ClCCH···HF	3.063	1.639	2.103; 2.346	0.928
F⁻···ClCCH···HF	2.385	1.651	2.058; 2.308	0.932

distances for isolated monomers; HCCH: CH(1.064); for HCCCl: CCl(1.640); for HF: 0.916

Table 2 Energies (corrected for BSSE, in kcal/mol) expressed by Eqs. 1–6

Complex	ΔE_{AB}	ΔE_{BC}	$\Delta E_{AB}{}'$	$\Delta E_{BC}{}'$	ΔE_{ABC}	E_{coop}
Cl$^-$···HCCH	−8.60	–	–	–	–	–
Cl$^-$···ClCCH	−5.66	–	–	–	–	–
F$^-$···ClCCH	−11.53	–	–	–	–	–
HCCH···HF	–	−3.15	–	–	–	–
HCCCl···HF	–	−2.48	–	–	–	–
Cl$^-$···HCCH···HF	–	–	−13.88	−8.50	−16.94	−0.57
Cl$^-$···ClCCH···HF	–	–	−9.54	−6.38	−11.96	−1.44
F$^-$···ClCCH···HF	–	–	−17.13	−7.94	−19.47	−2.48

Table 3 NEDA and DFT/NEDA decomposition terms of the energy of interaction (in kcal/mol, see Eqs. 11 and 17 as well as Eqs. 13, 14 and 15)

Complex	ΔE	El	CT	CORE	ES	PL	SE	XC	EX	DEF
Cl$^-$···HCCH	−8.5	−20.1	−22.0	33.6	−15.5	−8.3	3.7		−2.9	40.2
	−10.3	−20.4	−24.1	34.2	−15.2	−9.5	4.2	−4.4		42.9
Cl$^-$···ClCCH	−3.8	−22.6	−6.7	25.6	−9.8	−25.3	12.4		−4.4	42.4
	−7.2	−26.2	−8.8	27.9	−11.5	−29.0	14.3	−6.2		48.4
F$^-$···ClCCH	−11.1	−45.8	−15.9	50.6	−21.5	−47.9	23.6		−8.3	82.5
	−16.0	−53.3	−22.1	59.3	−26.2	−53.6	26.5	−12.4		98.3
HCCH···HF	−1.1	−7.0	−8.7	14.5	−6.1	−1.9	0.9		−2.0	17.4
	−3.2	−7.2	−10.1	14.1	−5.9	−2.7	1.4	−3.1		18.5
HCCCl···HF	−1.4	−6.7	−7.3	12.6	−5.7	−2.1	1.1		−1.8	15.5
	−3.2	−6.9	−8.6	12.3	−5.5	−2.9	1.5	−2.9		16.7
Cl$^-$···HCCH···HF	−16.8	−32.7	−31.9	47.8	−27.3	−9.7	4.4		−4.7	56.8
	−20.3	−32.3	−35.5	47.5	−26.2	−11.4	5.2	−7.2		60.1
Cl$^-$···ClCCH···HF	−9.5	−32.9	−14.9	38.3	−17.9	−29.8	14.8		−6.2	59.2
	−15.1	−36.4	−18.8	40.1	−19.3	−34.1	16.9	−9.1		66.1
F$^-$···ClCCH···HF	−18.3	−57.6	−26.1	65.4	−30.6	−53.5	26.6		−10.1	102.0
	−25.9	−64.8	−34.5	73.4	−35.0	−59.4	29.5	−15.4		118.3

First line—HF results, second line—DFT ones

between B3LYP and HF energies (ΔE) occur for F$^-$···ClCCH···HF, Cl$^-$···ClCCH···HF and F$^-$···ClCCH complexes. It is difficult to generalize which term is mainly responsible for the above-mentioned differences. Generally, absolute values of the energy components are greater for DFT/NEDA than for NEDA and there are only few exceptions. If one considers the decomposition of the energy of interaction according to Eqs. 11 and 17, then the greatest differences between both methods for the attractive components occur for CT energy followed by POL and ES.

Table 4 presents the interaction energies between orbitals; they were defined in the previous section (Eqs. 7–9). There is the $n_{Cl} \rightarrow \sigma^*_{CH}$ interaction corresponding to the hydrogen bond in the Cl$^-$···HCCH complex as well as the $n_{Cl} \rightarrow \sigma^*_{CCl}$ and $n_F \rightarrow \sigma^*_{CCl}$ interactions corresponding to the halogen bond in the Cl$^-$···ClCCH and F$^-$···ClCCH complexes,

respectively. One can see that for the Cl$^-$···HCCH and F$^-$···ClCCH complexes, the energies of interaction are very close one to each other in spite of the fact that the systems are linked through different interactions, hydrogen bond and halogen bond, respectively. The interaction for the Cl$^-$···ClCCH is not as strong as for two previous systems. For the HCCH···HF and HCCCl···HF complexes, the $\pi_B \rightarrow \sigma^*_{FH}$ energy of interaction amounts $\sim$4–5 kcal/mol. One can see that all these orbital–orbital interactions are enhanced in the corresponding trimers (Table 4).

3.2 Cooperativity effects in terms of Bent's rule

Figures 1 and 2 present selected NBO parameters for the dyads considered here while Figs. 3, 4, 5 concern the analyzed triads. The Lewis acid—Lewis base interaction

Table 4 E_{NBO} energies (in kcal/mol) expressed by Eqs. 7–9

Complex	$\Delta E\ (n_B \to \sigma^*_{CH(Cl)})$	$\Delta E\ (\pi_B \to \sigma^*_{FH})$
$Cl^- \cdots HCCH$	12.4	–
$Cl^- \cdots ClCCH$	4.3	–
$F^- \cdots ClCCH$	10.4	–
$HCCH \cdots HF$	–	4.8
$HCCCl \cdots HF$	–	3.9
$Cl^- \cdots HCCH \cdots HF$	19.5	7.4
$Cl^- \cdots ClCCH \cdots HF$	5.5	6.0
$F^- \cdots ClCCH \cdots HF$	13.0	7.6

leads to systematic changes of numerous parameters of interacting moieties (see Scheme 1). It was stated before that for the A–H⋯B hydrogen bond, there is the balance between two effects: the hyperconjugative interaction and the rehybridization [22]. The latter effect is connected with the increase in the s-character of A-atom hybrid orbital in the A-H bond. This is in agreement with Bent's rule since the hydrogen bond formation leads to the increase of the electropositive character of H-atom. There are other consequences of the complexation as for example the increase of the polarization of A–H proton donating bond.

The changes of parameters are also observed for the complexes analyzed here. For the $Cl^- \cdots HCCH$ complex (see Fig. 1), the increase in s-character of the carbon hybrid orbital in the C–H proton donating bond occurs since percentage s-character is equal to 50.7 for the complex while this value for the corresponding monomer amounts 48.1. The increase of the polarization of this C–H bond (% of electron density on carbon) is observed from 61.2 % in the monomer to 65.3 % in the complex. The similar changes are observed for the $Cl^- \cdots ClCCH$ halogen bonded system (Fig. 1), i.e. the increase in the s-character and the increase of the polarization of the C–Cl bond in the complex (42.6 and 49.8 %, respectively) in comparison with the ClCCH monomer (41.1 and 47.9 %, respectively). The changes of the s-character and the polarization are greater for the $F^- \cdots ClCCH$ complex than for the $Cl^- \cdots ClCCH$ (see Fig. 1) since the halogen bond interaction in the former moiety is stronger than in the latter one (see Table 2).

The similar trends observed for the halogen bond and for the hydrogen bond confirm the idea of the same mechanisms ruling numerous Lewis acid—Lewis base interactions. In general, the Lewis acid—Lewis base interaction leads to the electron charge transfer from the Lewis base sub-unit into the Lewis acid [57]. Hence, the Lewis base is slightly positive and the Lewis acid slightly negative if the isolated monomers are neutral. However, there is the further electron charge transfer within the Lewis acid sub-unit. This is observed here for the complexes of HCCH and HCCCl molecules where the C–H and C–Cl bonds are in

contacts with the Lewis base (Cl^- or F^-). The H-atom and Cl-atom positive charges for these bonds are greater in complexes than in the corresponding monomers (see Fig. 1). Hence, the H and Cl-atoms may be treated as more electropositive in complexes than they are in the isolated monomers. This leads, according to Bent's rule, to the increase in s-character of the C-atom hybrid orbital. The increase of the positive charge of the H-atom in A–H⋯B hydrogen bond is often treated as one of its more important characteristics [19, 58]. It was found recently [24] that for numerous non-covalent interactions, this is the characteristic of the Lewis acid center being in contact with the Lewis base.

Figure 2 presents results for the complexes linked by F–H⋯π hydrogen bonds. The s-character in the F-hybrid orbital and the polarization of the H–F bond for monomer are equal to 20.7 and 77.6 %, respectively. The increase of the latter values is observed in the HCCH⋯HF and HCCCl⋯HF complexes, s-character increases to 23.4 and 23.1, respectively, while the polarization to 78.6 and 78.4 %. The positive charge of the H-atom in the HF molecule increases from +0.547 au in the monomer to +0.556 au and +0.554 au in the complexes.

Figures 3, 4, 5 show the s-characters, the bond polarizations and the atomic charges in the triads, respectively. One can observe here the same tendencies which were observed before for the dyads: the increases of the s-character, the polarization and the positive charge of the atom (H or Cl-atom) being in contact with the Lewis base if these parameters are compared with the corresponding parameters of monomers. However, such changes in the triads are greater than those observed in the dyads. These findings support the earlier observations based on the geometrical and energetic parameters of the favorable (positive) cooperativity in the triads. For example, there is the increase of the positive charge of the Cl-atom from +0.143 au in the ClCCH monomer (Fig. 1) to +0.262 au in the $F^- \cdots ClCCH \cdots HF$ complex (Fig. 5). This charge for the $F^- \cdots ClCCH$ dyad is equal to +0.232 au (Fig. 1). The s-character and the polarization of the C–Cl bond amount 43.8 and 51.7 % in the triad, more than in the corresponding dyad, respectively. The similar comparison may be performed for the HF monomer, the HCCCl⋯HF dyad and the $F^- \cdots ClCCH \cdots HF$ complex where one can observe the following increase in s-character of F-atom hybrid orbital: 20.7, 23.1 and 24.9. There are the following values of the polarization of the H–F bond for the latter series: 77.6, 78.4 and 79.6 %. One can also see that there is the greater electron charge transfer from F^- ion for the $F^- \cdots ClCCH \cdots HF$ triad than for the corresponding $F^- \cdots ClCCH$ dyad since the F-ion charge for the triad is equal to -0.900 au while for the dyad, it amounts -0.920 au.

The mechanism of the electron charge redistribution for the $F^- \cdots ClCCH \cdots HF$ complex may be described in the
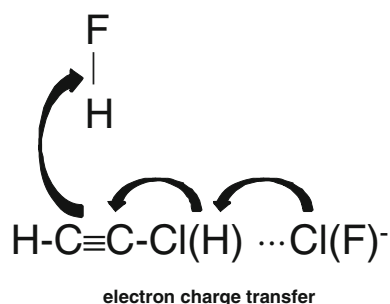
Scheme 2 The electron charge transfer for the triad

following way. There is the electron charge transfer from the F$^-$ Lewis base into the ClCCH unit acting here as the Lewis acid. The ClCCH molecule acts simultaneously as the Lewis base interacting with the HF Lewis acid. Hence, there is the further electron charge transfer from the ClCCH into the HF molecule. These processes of the electron charge transfer are common for all triads considered here (see Scheme 2).

4 Summary

The cooperativity effects analyzed for the Cl$^-$···HCCH··· HF, Cl$^-$···ClCCH···HF and F$^-$···ClCCH···HF triads show that all hydrogen bonds and halogen bonds which exist in the corresponding dyads are enhanced in the triads. In other words, the favorable cooperativity effects are observed for the systems analyzed here. The latter observation is based on the geometrical and energetic parameters.

The cooperativity is also analyzed in terms of Bent's rule. The A–X···B halogen bond and A–H···B hydrogen bond formation leads to the increase of the electropositive character of the X and H-atoms since their positive charges increase. This is connected with the increase of the polarization of the A–X (A–H) bond and the increase in s-character of the A-hybrid orbital. This is important that such changes being the effect of the electron charge redistribution are common for both interactions: hydrogen and halogen bonds (Scheme 1). The similar processes and changes of parameters as the result of complexation are observed for dyads and for triads. These processes are even enhanced and the changes are greater in the triads in comparison with the dyads. This means that the cooperativity between hydrogen and halogen bond is ruled by the same mechanisms as the cooperativity between two hydrogen bonds.

References

1. Jeffrey GA, Saenger W (1991) Hydrogen bonding in biological structures. Springer, Berlin
2. Jeffrey GA (1997) An introduction to hydrogen bonding. Oxford University Press, New York
3. Desiraju GR, Steiner T (1999) The weak hydrogen bond in structural chemistry and biology. Oxford University Press, New York
4. Grabowski SJ (2006) Hydrogen bonding—new insights. Vol 3 of the series: challenges and advances in computational chemistry and physics. Leszczynski J (ed) Springer, Dordrecht
5. Alkorta I, Rozas I, Elguero J (1998) Chem Soc Rev 27:163–170
6. Rozas I, Alkorta I, Elguero J (1997) J Phys Chem A 101:4236–4244
7. Cotton FA, Matonic JH, Murillo CA (1998) J Am Chem Soc 120:6047–6052
8. Grabowski SJ, Sokalski WA, Leszczynski J (2006) Chem Phys Lett 422:334–339
9. Peris E, Lee JC Jr, Rambo J, Eisenstein O, Crabtree RH (1995) J Am Chem Soc 117:3485–3491
10. Wessel J, Lee JC Jr, Peris E, Yap GPA, Fortin JB, Ricci JS, Sini G, Albinati A, Koetzle TF, Eisenstein O, Rheingold AL, Crabtree RH (1995) Angew Chem Int Ed Engl 34:2507–2509
11. Crabtree RH, Siegbahn PEM, Eisenstein O, Rheingold AL, Koetzle TF (1996) Acc Chem Res 29:348–354
12. Bakhmutow VI (2008) Dihydrogen bonds. John Wiley, New Jersey
13. Metrangolo P, Resnati G (2001) Chem Eur J 7:2511–2519
14. Formigué M, Batail P (2004) Chem Rev 104:5379–5418
15. Zordan F, Brammer L, Sherwood P (2005) J Am Chem Soc 127:5979–5989
16. Clark T, Hennemann M, Murray J, Politzer P (2007) J Mol Model 13:291–296
17. Murray JS, Riley KE, Politzer P, Clark T (2010) Aust J Chem 63:1598–1607
18. Politzer P, Murray JS, Clark T. Phys Chem Chem Phys 12:7748–7757
19. Weinhold F, Landis C (2005) Valency and bonding, a natural bond orbital donor—acceptor perspective. Cambridge University Press, Cambridge
20. Reed AE, Curtiss LA, Weinhold F (1988) Chem Rev 88:899–926
21. Weinhold F (1997) J Mol Struct (Theochem) 398–399:181–197
22. Alabugin IV, Manoharan M, Peabody S, Weinhold F (2003) J Am Chem Soc 125:5973–5987
23. Bent HA (1961) Chem Rev 61:275–311
24. Grabowski SJ (2011) J Phys Chem A 115:12340–12347
25. Grabowski SJ (2012) J Phys Chem A 116:1838–1845
26. Garau C, Quiñonero D, Frontera A, Ballester P, Costa A, Deyà PM (2005) J Phys Chem 109:9341–9345
27. Frontera A, Quiñonero D, Garau C, Costa A, Ballester P, Deyà PM (2006) J Phys Chem A 110:9307–9309
28. Quiñonero D, Frontera A, Garau C, Ballester P, Costa A, Deyà PM (2006) Chem Phys Chem 7:2487–2491
29. Frontera A, Quiñonero D, Costa A, Ballester P, Deyà PM (2007) New J Chem 31:556–560
30. Hunt SW, Higgins KJ, Craddock MB, Brauer CS, Leopold KR (2003) J Am Chem Soc 125:13850–13860
31. Parra RD, Bulusu S, Zeng XC (2005) J Chem Phys 122:184325
32. Kobko N, Dannenberg JJ (2003) J Phys Chem A 107:6688–6697

33. Parra RD, Ohlssen J (2008) J Phys Chem A 112:3492–3997
34. Araújo RCMU, Soares VM, Oliveira BG, Lopes KC, Ventura E, DoMonte SA, Santana OL, Carvalho AB, Ramos MN (2006) Int J Quantum Chem 106:2714–2722
35. Ziółkowski M, Grabowski SJ, Leszczynski J (2006) J Phys Chem A 110:6514–6521
36. Zhao Q, Feng D, Hao J (2011) J Mol Mod 17:2817–2823
37. Solimannejad M, Malekani M, Alkorta I (2011) Mol Phys 109:1641–1648
38. Solimannejad M, Malekani M (2010) J Phys Chem A 114:12106–12111
39. Grabowski SJ, Bilewicz E (2006) Chem Phys Lett 427:51–55
40. Alkorta I, Blanco F, Deyà PM, Elguero J, Estarellas C, Frontera A, Quiñonero D (2010) Theor Chem Acc 126:1–14
41. Frisch MJ, Trucks GW, Schlegel HB et al (2009) Gaussian 09, Revision A.02, Inc., Wallingford
42. Read WG, Flygare WH (1982) J Chem Phys 76:2238–2246
43. Allen FH, Davies JE, Galloy JE, Johnson JJ, Kennard O, Macrave CF, Mitchel EM, Smith JM, Watson DG (1991) J Chem Inf Comput Sci 31:187–204
44. Ghassemzadeh M, Harms K, Dehnicke K (1996) Chem Ber 129:115–120
45. Yamamoto HM, Maeda R, Yamaura JI, Kato R (2001) J Mater Chem 11:1034–1041
46. Grabowski SJ (2006) Annu Rep Prog Chem Sect C 102:131–165
47. Boys SF, Bernardi F (1979) Mol Phys 19:553–566
48. Salvador P, Szczęśniak MM (2003) J Chem Phys 118:537–549
49. Glendening ED, Streitwieser A Jr (1994) J Chem Phys 100:2900–2909
50. Schenter GK, Glendening ED (1996) J Phys Chem 100:17152–17156
51. Glendening ED (1996) J Am Chem Soc 118:2473–2482
52. Glendening ED (2005) J Phys Chem A 109:11936–11940
53. NBO 5.0. Glendening ED, Badenhoop JK, Reed AE, Carpenter JE, Bohmann JA, Morales CM, Weinhold F (2001). Theoretical Chemistry Institute, University of Wisconsin, Madison
54. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su SJ, Windus TL, Dupuis M, Montgomery JA (1993) J Comput Chem 14:1347–1363
55. Sobczyk L, Grabowski SJ, Krygowski TM (2005) Chem Rev 105:3513–3560
56. Grabowski SJ (2011) Chem Rev 11:2597–2625
57. Lipkowski P, Grabowski SJ, Leszczynski J (2006) J Phys Chem A 110:10296–10302
58. Popelier P (2000) Atoms in molecules. An introduction. Prentice Hall, Pearson Education Limited, Harlow UK

Isotope effects on the dynamics properties and reaction mechanism in the Cl(²P) + NH₃ reaction: a QCT and QM study

Manuel Monge-Palacios · Cipriano Rangel · Joaquin Espinosa-García · Hong Fu · Minghui Yang

Abstract Isotope effects on dynamics properties and reaction mechanism in the title reaction, which evolve through deep wells in the entry and exit channels, were analysed using both quasi-classical trajectory and reduced dimensional quantum mechanical calculations, for collision energies in the range 3.0–10.0 kcal mol^{-1}, on an analytical potential energy surface (PES-2010) recently developed by our group. The analysis of different dynamics properties (reaction probability, product energy partitioning, and product rovibrational distributions) shows the reaction behaviour of the two reactions, Cl + NH₃ and Cl + ND₃, to be similar: direct mechanism, with a small percentage of indirect mechanisms, that is, trapping complex mediated. We find that the only dynamics property dependent on the isotope effect is the product scattering distribution. In particular, while the perdeuterated reaction favours backward scattering, the perprotio analogue favours forward scattering. This behaviour is related to the smaller maximum impact parameter for the perdeuterated reaction.

Manuel Monge-Palacios and Hong Fu contributed equally to this work.

M. Monge-Palacios · C. Rangel · J. Espinosa-García (✉)
Departamento de Química Física, Universidad de Extremadura, 06071 Badajoz, Spain
e-mail: joaquin@unex.es

H. Fu · M. Yang
State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences, Wuhan 430071, People's Republic of China

Keywords Cl + NH₃ reaction · Isotope effect · Reaction dynamics · Reaction mechanism

1 Introduction

The chemistry of the reaction of ammonia with chlorine atoms is very complex, with many intermediate fast reactions being involved [1]. Perhaps unsurprisingly therefore, the Cl(²P) + NH₃(v) → HCl + NH₂ hydrogen abstraction reaction has been little studied, either experimentally or theoretically. Only two experimental kinetics measurements have been reported [2, 3], and we know of only four theoretical studies [3–6]. The reaction is endothermic and very slow, but the most interesting characteristic is the presence of deep wells in the entry and exit channels which can influence the reaction dynamics.

In a very recent work [7], we constructed an analytical potential energy surface, PES-2010, for this reaction and its isotope analogues. This surface is based exclusively on high-level ab initio calculations. We found that the ab initio information used in the fit is well reproduced by the new PES-2010 surface, especially the barrier height and the depth of the wells, and that this surface reproduces the experimental forward rate constants in the common temperature range. However, to the best of our knowledge, no dynamics information about this reactive system is available, either experimental or theoretical, and only recently has our group begun a series of dynamics studies using quasi-classical (QCT) and quantum–mechanical (QM) calculations to analyse the role that these complexes in the entry and exit channels play in the dynamics and the atomic-level mechanisms [8, 9]. Due to the presence of the reactant well, we found different mechanisms of reaction

depending on the collision energy: indirect at low (<3 kcal mol^{-1}) and direct at high (>5 kcal mol^{-1}) collision energies [8].

To go deeper into the understanding of this reaction and the role of the isotopic substitution, in the present work, we describe a combined QCT/QM dynamics study of the deuterated analogue reaction, $Cl(^2P) + ND_3(v = 0) \rightarrow DCl + ND_2$, using the same analytical PES-2010 surface. By comparing the results with the perprotio system, we shall analyse the influence of the complexes in the entry and exit channels on the dynamics and the atomic-level mechanism. Unfortunately, as in the $Cl + NH_3$ reaction, neither experimental nor theoretical dynamics information is available for comparison.

The article is structured as follows. Section 2 briefly outlines the potential energy surface, and the energy and computational details of the quasi-classical trajectory (QCT) and reduced dimensionality quantum mechanical (QM) methods for the dynamics study of the deuterated reaction. Section 3 presents the QCT and QM dynamics results, together with a comparison with the perprotio reaction. This latter is re-analysed based on the new findings of the reduced dimensionality QM calculations which called into question the previous conclusions about the mechanism at very low energies [8]. Finally, the conclusions are presented in Sect. 4.

2 Theory and computational details

2.1 Potential energy surface

All calculations were performed on a potential energy surface (PES-2010) recently constructed by our group for the $Cl(^2P) + NH_3(v) \rightarrow HCl + NH_2$ gas-phase polyatomic reaction and its isotopic analogues [7]. The surface is symmetric with respect to permutation of the ammonia hydrogen atoms, a feature especially interesting for dynamics calculations. The functional form was developed in the cited work and therefore will not be repeated here. Basically, it consists of three LEP-type stretch terms (str), augmented by valence (val) bending terms, and a series of switching functions allowing the smooth change from pyramidal NH_3 to the planar NH_2 product in the hydrogen abstraction reaction. It depends on 32 adjustable parameters. As input data in the fitting process, we used exclusively very high-level ab initio calculations [CCSD(T) = FULL/aug-cc-pVTZ]. If one neglects the spin–orbit coupling, then electronic ^2P ground state of the chlorine atom is triply degenerate. Interaction with the NH_3 molecule removes this degeneracy, giving rise in general to three orthogonal states. In the Cs symmetry, the states are as follows: $1^2A'$, $2^2A'$, and $2^2A''$. Nevertheless, only the

one with lowest energy ($1^2A'$) adiabatically correlates with the $HCl(X^1\sum^+) + NH_2$ (X^2) products. Thus, we shall focus attention on the ground state ($1^2A'$).

The PES-2010 has a reaction energy of 9.2 kcal mol^{-1}, and the transition state is located 'late' on the reaction path with a barrier height of 6.6 kcal mol^{-1}. In addition, it presents two deep wells—one in the entry valley, stabilised 7.2 kcal mol^{-1} with respect to the reactants, and another in the exit channel, stabilised 7.3 kcal mol^{-1} with respect to the products. This surface has been tested [7], finding good agreement with the available experimental data. Figure 1 shows the potential energy and the adiabatic changes (i.e., zero-point energy included) along the reaction paths for the $Cl(^2P) + ND_3$ and $Cl(^2P) + NH_3$ for comparison. While the reactant wells present similar stabilities, the adiabatic barriers, the products wells, and the endothermicity are 1.3 kcal mol^{-1} higher for the deuterated reaction. Note however that with a higher barrier height, the deuterated reaction presents a deeper reactant well. Obviously, these differences could have consequences on the kinetics and dynamics.

2.2 Available energy and reaction threshold

Formally, the total energy (zero-point energy included) available to reaction is

$$E_{\text{avail}} = E_{\text{int,react}} + E_{\text{coll}} \qquad (1)$$

where $E_{\text{int,react}}$ and E_{coll} are, respectively, the internal energy of the reactants and the collision energy. For the endothermic $Cl(^2P) + ND_3(v = 0)/NH_3(v = 0)$ reactions, the minimum energy to open the deuterium/hydrogen abstraction reaction does not depend on the adiabatic barrier, 3.5/2.2 kcal mol^{-1}, but on the adiabatic reaction
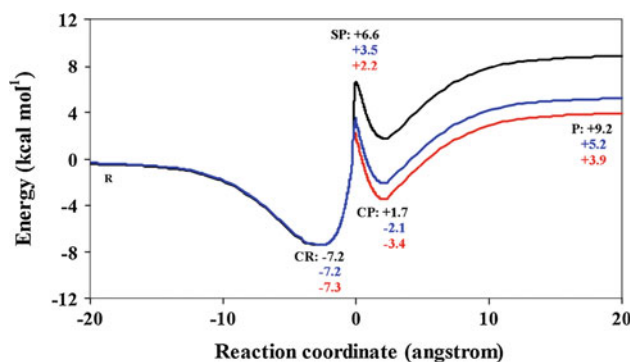
Fig. 1 Plots of the non-relativistic potential energy (*black line*) and adiabatic (i.e., zero-point energy included) profiles along the reaction path for the $Cl(^2P) + ND_3$ (*blue line*) and $Cl(^2P) + NH_3$ (*red line*) reactions using the PES-2010 surface. For each stationary point, the first entry corresponds to the potential energy along the path, the second to the adiabatic energy for the $Cl + ND_3$ reaction, and the third to the adiabatic energy for the $Cl + NH_3$ reaction. All values with respect to the reactants, level zero

energy, 5.2/3.9 kcal mol^{-1}, respectively, which is the reaction threshold in each reaction. Classically, one must expect that only encounters with E_{avail} above-the-reaction threshold, 5.2 kcal mol^{-1} for the perdeuterated reaction and 3.9 kcal mol^{-1} for the perprotio reaction, will result in products. However, in this particular reaction, these data merit more attention, as was extensively analysed in our previous paper [8]. Here, therefore, we present just a brief synthesis.

First, the experimental enthalpy of reaction is unknown and must be obtained from the corresponding enthalpies of formation, $\Delta H_f(0K)$. However, while the ΔH_f's of all the compounds are well established [10], the proposed experimental values of the NH_2/ND_2-free radical are a subject of serious controversy, with an uncertainty of ± 1.0 kcal mol^{-1} in accordance with the latest experimental value [11]. Second, the reactant NH_3/ND_3 rotational energy can be calculated theoretically from the equipartition energy, $(3/2)RT$, or sampled from a thermal distribution, as in the present paper in which a temperature of 300 K is used. Obviously, this rotational energy is available for reaction. At this temperature, $(3/2)RT = 0.9$ kcal mol^{-1}.

Therefore, given the reactant rotational energy and the experimental uncertainties in the enthalpy of reaction, it is very difficult to propose a minimum collision energy with which to surpass the reaction threshold. A reasonable theoretical approach is to suppose a lower limit for the collision energy. In our previous study on the $Cl + NH_3$ reaction, we use a value ≥ 2.5 kcal mol^{-1}, and in the present study for the $Cl + ND_3$ reaction, we propose a value ≥ 4.0 kcal mol^{-1}. Obviously, one expects that the experimental detection of products at the lower limit, if the interaction finally results in products, will present serious problems since the products' signal will be blurred by the experimental noise. Thus, these reactions represent a clear challenge experimentally at low energies.

2.3 Quasi-classical trajectory calculations

Quasi-classical trajectory calculations [12–14] were carried out using the VENUS96 code [15], customised to incorporate our analytical PESs. The accuracy of the trajectories was checked by the conservation of total energy and total angular momentum. The integration step was 0.1 fs, with an initial separation between the Cl atom and the ammonia centre of mass of 10.0 Å. Moreover, additional tests were performed at larger distances, 12.0 and 14.0 Å, finding no differences in final results. The trajectories were finished when the Cl–N distance was greater than 11.0 Å. As in the $Cl + NH_3$ reaction, at low collision energies, this large distance represents very long times of flight as compared with the high-energy regime. For direct comparison with the perprotio reaction, the vibrational energy corresponds

to the ND_3 in its ground vibrational state, and the rotational energies were obtained by thermal sampling at 300 K from a Boltzmann distribution. The reactant collision energies considered ranged from 3.0 to 10.0 kcal mol^{-1}.

For each energy, the maximum value of the impact parameter, b_{max}, was determined by first calculating batches of 10,000 trajectories at fixed values of the impact parameter, b, systematically increasing the value of b until no reactive trajectories were obtained. This value varied depending on the collision energy, from 3.0 Å at 3.0 kcal mol^{-1} to 2.7 Å at 10.0 kcal mol^{-1}. Note that for the $Cl + NH_3$ reaction, the values in this range vary from 3.9 to 3.2 Å.

Then, batches of 200,000 trajectories were calculated for each collision energy, with the impact parameter, b, sampled from

$$b = b_{max} R^{1/2} \tag{2}$$

where R is a random number in the interval [0,1]. Moreover, to allow comparison with the QM calculations which were performed at $J = 0$, selected calculations on the PES-2010 surface were repeated with an impact parameter $b = 0$. In total, about 2×10^6 trajectories were run.

The reaction probability, $P_r = N_r/N_T$, is the ratio of the number of reactive trajectories to the total number of trajectories, while the reaction cross section is defined as

$$\sigma_r \pi b_{max}^2 (N_r/N_T) \tag{3}$$

Because of the large number of trajectories run, the sampling with respect to the impact parameter and rovibrational states is good enough to allow one to expect that the results do not depend on the averaging method. Therefore, Eq. (3) can be applied despite its simplicity.

It is well known that one of the difficulties with quasi-classical simulations is related to the question of how to handle the zero-point energy (ZPE) problem. Various strategies have been proposed to correct this quantum–mechanical effect, but no really satisfactory alternatives have emerged (see, for instance, Refs. [16, 17] and references therein). Here, we employed a pragmatic solution, the so-called passive method, which consists of running the trajectories with no quantum constraint to subsequently analyse the trajectories and discard those that are not allowed in a quantum mechanical world, even knowing that this method perturbs the statistics and can therefore lead to uncertainties in the dynamics study [18]. Recently, our group has analysed the ZPE problem in QCT calculations using full-dimensional QM results on the same PES as a target for the $H + NH_3$ reaction [19]. So, in Eq.(3), the correct determination of the number of reactive trajectories, N_r, and the total number of trajectories, N_T, has to be done having in mind the ZPE problem. So, we have three possible values of calculating N_r: counting all the reactive

trajectories (N_r^{all}), counting only those that lead to both products with vibrational energy above their ZPE, which is named histogram binning with double ZPE constraint, HB-DZPE, and counting only the trajectories that lead to the newly formed bond, HCl, with vibrational energy above its ZPE, following Thrular's [20] and Schatz's [21] suggestions, who suggested that the ZPE constraint should be applied just to the newly formed bond. We call this method histogram binning with ZPE-HCl constraint, HB-ZPE-HCl.

However, this way of removing trajectories from the N_r count without taking into account, the behaviour of the total ensemble of trajectories can lend erroneous results, because as mentioned above, it modifies the statistics [16, 18, 22]. So, here we count the total number of trajectories, N_T, in two ways. First, as it has been usual in the QCT literature, we consider the total number of trajectories ran in the calculation, N_T^{all}. In the second approach, N_T in Eq. (3) is replaced by the total number of trajectories minus the number of reactive trajectories whose final vibrational energy is below the ZPE of the two products or the H$_2$ product (depending on whether we use the HB-DZPE or HB-ZPE-H2 criterion), and minus the number of nonreactive trajectories whose final vibrational energy is below the ZPE of the NH$_3$ reactant. This number we call, N_T^*. Therefore, we have six counting methods for the calculations of the reaction probability, N_r/N_T, and therefore of the reaction cross section: N_r^{all}/N_T^{all}, $N_r^{HB-DZPE}/N_T^{all}$, $N_r^{HB-ZPE-HCl}/N_T^{all}$, N_r^{all}/N_T^*; $N_r^{HB-DZPE}/N_T^*$, $N_r^{HB-ZPE-HCl}/N_T^*$.

Taking the full-dimensional QM calculations as a target we found [19] that the best agreement was obtained with the $N_r^{HB-ZPE-HCl}/N_T^*$ approach. In the previous paper [8] on the Cl + NH$_3$, we used this criterion, and given that in the present paper, we study isotope effects on the Cl + NH$_3$ reaction, we use the same approach. So, possible errors in the counting methods can cancel out.

2.4 Quantum mechanical calculations

A seven-dimensional quantum dynamics model for the reaction X + YCZ$_2$ was employed in this work. The details of the method have been described elsewhere [23] and will only be briefly outlined here. For the Cl + ND3 reaction, Jacobi coordinates are employed to describe the motion of the system (Fig. 2). R is the vector from the centre of mass of ND$_3$ to the Cl atom; r is the vector from the centre of mass of ND$_2$ to the third D atom; q is the bond length of ND and is fixed at its equilibrium values; χ is the bond angle between the two ND bonds; and the vector s is the symmetry axis of the NH$_2$ group. The bending angle between vectors R and r is defined to be θ_1; φ_1 is the azimuth angle of the rotation of ND$_3$ around the vector r; θ_2 is the bending angle between vectors r and s; and φ_2 is

the azimuth angle of the rotation of ND$_2$ around vector s. The atomic masses are denoted as m_{Cl}, m_N, and m_D.

The model Hamiltonian for the Cl + ND$_3$ system is

$$\hat{H} = -\frac{1}{2\mu_R}\frac{\partial^2}{\partial R^2} - \frac{1}{2\mu_r}\frac{\partial^2}{\partial r^2} + \frac{(\hat{J}_{tot} - \hat{J})^2}{2\mu_R R^2} + \frac{\hat{l}^2}{2\mu_r r^2} + \hat{K}_{ND_2}^{vib} + \hat{K}_{ND_2}^{rot} + V(R, r, \chi, \theta_1, \phi_1, \theta_2, \phi_2) \quad (4)$$

where μ_R is the reduced mass of the Cl + ND$_3$ system and μ_r is the reduced mass of D + ND$_2$. The first two terms are the kinetic energy operators for R and r, respectively; $\hat{J}_{tot}$ is the total angular momentum operator of the system; $\hat{J}$ is the rotational angular momentum operator of ND$_3$; and $\hat{l}$ is the orbital angular momentum operator of atom Y with respect to ND$_2$. The last term $V(R, r, \chi, \theta_1, \phi_1, \theta_2, \phi_2)$ is the potential energy.

By fixing the ND bond lengths at their equilibrium values, the vibrational kinetic energy operator of the ND$_2$ bending vibration is

$$\hat{K}_{ND_2}^{vib} = -\frac{1}{\mu_1 q^2}\frac{1}{\sin\chi}\frac{\partial}{\partial\chi}\sin\chi\frac{\partial}{\partial\chi} + \frac{1}{\mu_2 q^2}\left(\frac{\cos\chi}{\sin\chi}\frac{\partial}{\partial\chi}\sin\chi\frac{\partial}{\partial\chi} + \sin\chi\frac{\partial}{\partial\chi}\right) \quad (5)$$

where $\mu_1 = m_N$ and $\mu_2 = m_N m_D/(m_N + m_D)$ are the reduced masses. The rotation of the ND$_2$ group is approximated to be decoupled from its vibrational motion and described by an asymmetric rigid rotor Hamiltonian, with the rotational constants $\{B_X, B_Y, B_Z\}$ varying with the bending angle χ.

$$\hat{K}_{ND_2}^{rot} = \frac{1}{2}(B_X + B_Y)\hat{j}^2 + \left[B_Z - \frac{1}{2}(B_X + B_Y)\right]\hat{j}_z^2 + \frac{1}{2}(B_X - B_Y)(\hat{j}_+^2 + \hat{j}_-^2) \quad (6)$$

where $\hat{J}$ is the angular momentum operator of the ND$_2$ group, $\hat{J}_Z$ is the z-component of $\hat{J}$, and $\hat{J}_+$ and $\hat{J}_-$ are the shift operators.
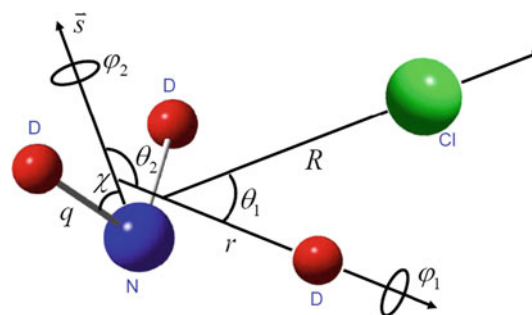

Fig. 2 Jacobi coordinates of the X + YCZ$_2$ model for the Cl + ND$_3$ reaction

Because of the double well in this reaction, we used an extremely large basis set in the work. In the L-shape technique used for coordinates R and r, a total of 480 basis sine functions ranging from 2.5 to 23.0 Bohr were used for the R basis set expansion with 410 nodes in the interaction region. The range of r was set to be from 1.0 to 21.0 Bohr to include the well in the product side, and 5 and 140 basis functions were used in the asymptotic and interaction regions, respectively. For the ND_2 bending mode, 1 PODVR (potential-optimised discrete variable representation) node was used. The size of the rotational basis function set is controlled by the parameters, $J_{max} = 48$, $l_{max} = 30$, $j_{max} = 18$, and $k_{max} = 18$. After considering parity and C_{2v} symmetry, the size of the set of rotational basis functions was 54, 979, and the number of nodes for the integration of the rotational basis set was 654, 493. The size of the total basis function set was 3.175×10^9 and the number of nodes was 3.78×10^{10}. The dividing surface was defined at $r = 3.5$ Bohr. Due to the deep well in the entry channel and the heavy mass of the Cl atom, the entire propagation time was 25,000 a.u. with a step size of 10 a.u. for the reaction to converge from ND_3 initially in its vibrational ground state. In sum, the time-dependent quantum mechanical calculations provide total reaction probabilities as a function of the translational energy for $J_{tot} = 0$.

Before ending this section, we believe it is interesting to note that, compared with other polyatomic reactions, the Cl + NH3 reaction is very complex due to its two deep van der Waals complexes on the reactant and on the product sides. The QCT approach, although classical in nature, considers all the reaction coordinates and produces very useful information.

Also, the computational cost makes it difficult to carry out a full-dimensional QM calculation for this heavy system and complex reaction, although we actually have developed the full-dimensional method. The reduced dimensional QM model used in the present paper is reasonable since the values of the spectator bond length (N–H) and bond angle (H–N–H) are almost same in the geometries of the reactant and transition state. The reduced QM calculation is thus able to explore quantum effects in this reaction at an affordable computational cost and is an useful complement to the full-dimensional QCT calculations.

3 Results and discussion

3.1 Reaction probability: influence of the reactant well and mechanism of reaction

Figure 3 shows the reaction probability as a function of the collision energy for the Cl + ND_3 reaction, together with
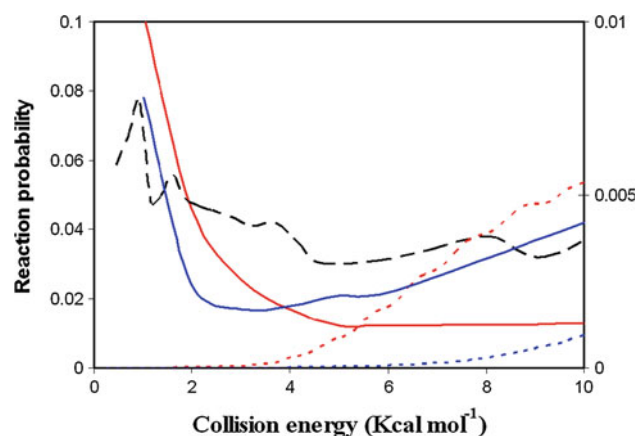

Fig. 3 Reaction probability as a function of the collision energy (kcal mol^{-1}) using the PES-2010 surface for the Cl + ND_3 reaction (*blue lines*) and Cl + NH_3 reaction (*red lines*). *Solid lines*: QCT calculations (b = 0); *dashed lines*: QM calculations ($J = 0$). We also include the wrong 7D-QM calculations for the Cl + NH_3 reaction for completeness (*black-dashed line*). Note that for a clearer discussion, the scales in the *right* and *left axis* are different: Only the QCT (*red-solid line*) and the old 7D-QM (*black-dashed line*) results are represented on the *left axis*

the Cl + NH_3 reaction for comparison, using the QCT and QM calculations on the same PES-2010 surface. With this plot, we shall analyse the dynamics method, the influence of the reactant well, and the mechanism of reaction.

First, it is important to note that we have found a mistake in the previous 7D-QM calculations for the Cl + NH_3 reaction [8], and the peak in the reaction probability at low energies is wrong in the reduced dimensionality quantum dynamics calculations. In particular, the parameter to define the dividing surface was wrongly selected. In that paper, the parameter was wrongly set to be $r = 2.2$ Bohr and therefore in the entry channel and before of the saddle point ($r = 2.5$ Bohr), where r is defined as the distance between the third atom and the centre of mass of NH_2. Thus, some wavepackets did not pass through the saddle point were wrongly considered. In the present paper, all calculations set the flux surface to be $r = 3.5$ Bohr which is in the product channel.

The new 7D-QM calculations for the Cl + NH_3 reaction show a very different behaviour, where the reaction probability increases smoothly with collision energy from the reaction threshold. Thus, the new QM calculations show the typical behaviour of a reaction with barrier, and neither resonances nor another quantum effects are observed below the reaction barrier. Therefore, the increase in reaction probability at low energies obtained with the QCT calculations in that work [8], due to repeated collisions between the Cl and NH_3 reactants in the entry valley, was really an artefact of the QCT method based on the ZPE violation, where all trajectories appear with the vibrational energy in the NH_2 product below its ZPE, which is quantum

mechanically forbidden. The well in the entry channel yields an unphysical vibrational → translational energy transfer, which artificially favours the use of this energy to promote the reaction, and the classical above-the-threshold reaction occurs even when the collision energy is below the quantum–mechanical threshold to reaction.

Second, due to the higher threshold for the perdeuterated reaction with respect to the perprotio ones (Fig. 1), the reactivity is smaller at all collision energies. This result agrees with the kinetic isotope effects found in our previous kinetics study [7] using the same surface, which ranged from 9.2 at 300 K to 1.4 at 2000 K. Thus, the deuteration diminishes the reactivity because the threshold is higher (~ 1.3 kcal mol^{-1}) due to the different ZPEs.

Third, for the $Cl + ND_3$ reaction, the QCT probability presents slight increases in reactivity with collision energy from the threshold of reaction, with slight increases also at lower energies. Is this behaviour at low energies a real effect or an artefact of the QCT calculations as in the case of the $Cl + NH_3$ reaction? The 7D-QM calculations also give a very small reactivity, indicating that few wavepackets pass though the reaction threshold, although there is no enhancement of reactivity at low collision energies. Therefore, the QCT reactivity at low energies is an artefact of the method (as in the $Cl + NH_3$ reaction), and the QM results are more realistic.

To look more deeply into the QCT results at low energies, we performed a detailed analysis of the trajectories. At low energies, we observed all reactive trajectories to appear with the vibrational energy in the ND_2 product below its ZPE, an effect known as ZPE violation. Thus, part of the leaked vibrational energy is used to promote the reaction, and the classical above-the-threshold reaction occurs even when the collision energy is below the quantum–mechanical threshold to reaction (the zero-point energy corrected threshold). In other words, although QCT assumes classical motion along the reaction path, there is a possibility for the reaction to occur by lowering the vibrational energy below the ZPE, even though there is not enough energy to surpass the threshold. In the analysis of trajectories in the low-energy regime, we observed that orthogonal motions on the reaction path (vibrational modes of the ND_3–ND_2 evolution) are transformed into translational energy along the path, and that this $V \rightarrow T$ transfer was used to overcome the threshold. Therefore, the ND_2 product appears with vibrational energy below its ZPE, which is quantum mechanically forbidden. This is the reason for the significant values of the reaction probability at low energies observed in Fig. 3, both for the $Cl + ND_3$ and for the $Cl + NH_3$ reactions. The larger enhancement for the $Cl + NH_3$ reaction is due to lower collision energies, permitting a larger number of collisions between the Cl and NH_3 reactants in the entry valley, which favours the

unphysical $V \rightarrow T$ energy transfer. Therefore, the reactivity at low energies is an artefact of the QCT calculations, which is confirmed by the QM results.

For comparison purposes, we examined the influence of another two counting methods—the VEQMT [24] and the VEQMTc [25] approaches. Using the VEQMT approach, we obtained no reactive trajectory in the energy range considered, 3–10 kcal mol^{-1}, confirming that this approach is too restrictive, as has previously been noted [25]. In this sense, the VEQMT approach is similar to our DZPE counting method. When the less stringent VEQMTc approach is used, it should be taken into account that the zero-point energy of the product complex is 0.2 kcal mol^{-1} higher than the ZPE of the products, so that no reactive trajectory should be expected with this counting method. Indeed, an energy of 10.0 kcal mol^{-1} is required to obtain only 2 % of the reactive trajectories. For this system, therefore, the VEQMT and VEQMTc approaches are both too restrictive. Another non-active scheme [24] would be to require the internal energy of each product molecule to be larger than or equal to its ZPE (internal energy quantum mechanical threshold (IEQMT) scheme), which is a priori less demanding than VEQMT. In fact, we obtained no reactive trajectories in the energy range 3–5 kcal mol^{-1}, while at 10.0 kcal mol^{-1}, we obtain only 3 % of reactive trajectories. Therefore, for the present reaction, the IEQMT scheme follows being too restrictive. These results show that there is no unambiguous way to fix the leakage problem of classical mechanics, so that the only solution is via quantum mechanics [25].

Finally, we also examined the influence of the impact parameter. All the QCT calculations were re-run with b_{max} (instead of $b = 0$). The variation of the reaction probability with the collision energy was similar in shape to the $b = 0$ case and so is not shown here. Therefore, the behaviour with energy and the possible mechanisms are independent of the choice of impact parameter. After this comparison with QM calculations, in the rest of this communication, the value used for the QCT calculations will be b_{max}.

3.2 Nature of the trajectories and atomic-level reaction mechanisms: comparison with the $Cl + NH_3$ reaction

In order to determine the nature of the trajectories and the atomic-level mechanism, we performed an analysis of the individual reactive trajectories over the wide collision energy range considered, comparing them with the $Cl + NH_3$ system previously studied [8]. We made a distinction between direct and indirect (i.e., trapping complex mediated) mechanisms. The former involves a very brief encounter between the reactants and is the most likely in the high-energy regime. In the latter, the influence of the

Table 1 Atomic-level mechanism as a function of the collision energy for $Cl(^2P) + ND_3$ and $Cl(^2P) + NH_3$ reactions using the PES-2010 surface

Collision energy[a]	Direct trajectories (%)[b]		Indirect trajectories (%)[c]	
	$Cl + ND_3$	$Cl + NH_3$	$Cl + ND_3$	$Cl + NH_3$
3.0	–	85	–	15
4.0	90	100	10	0
5.0	100	100	0	0
10.0	100	100	0	0

[a] In kcal mol^{-1}

[b] Percentages calculated with respect to the total of reactive trajectories

[c] Total percentage of indirect trajectories (calculated with respect to the total of reactive trajectories), that is, trajectories that enter the reactant and/or product complex regions

wells in the entry and/or exit channels is important, and indeed is the dominant factor in the low-energy regime. Table 1 summarises the influence of the energy regime on the reaction mechanism, presenting the percentages of the reactive trajectories which occur via direct and indirect mechanisms for each reaction. The data at collision energies of 3.0 and 4.0 kcal mol^{-1} for NH_3 and ND_3 reactions correspond, respectively, to the limit of the respective low-energy regime threshold.

To gain insight into the reaction mechanism, Fig. 4 shows typical cases of the evolution of the broken bond N–D (dashed line) and formed bond D–Cl (solid line) distances as a function of time. In the direct or impulse-type mechanism (Fig. 4a), $E_{coll} > 4.0$ kcal mol^{-1}, the Cl collides with ND_3 and forms the DCl product which immediately recoils away. In the indirect mechanism (Fig. 4b–d),

10 % of reactive trajectories at $E_{coll} = 4.0$ kcal mol^{-1}, the situation is more complicated since the trajectories pass through the regions of the complexes. Two basic types of trajectories were found—those that enter the reactant complex region (Fig. 4b), and those that enter the product complex region (Fig. 4c). In the former case, the Cl and ND_3 reactants undergo several collisions in the entry channel, with different energy transfers between T, R, and V. The reactive system does not fly apart following these collisions, and after a certain time, the trajectory suddenly 'finds' the way towards the products. The complex in the entry channel thus allows repeated encounters between the reactants before the formation of the products, explaining the reaction probability at low energies. In the latter case (Fig. 4c), once formed, the DCl and ND_2 products do not fly apart. Instead, they undergo several collisions in the exit channel, usually associated with rotation of the DCl product. After a certain time (whose magnitude depends on the trajectory), the products separate. Obviously, a combination of these two kinds of trajectory (Fig. 4b,c) is also possible, as was indeed found in the visualisation of individual reactive trajectories (Fig. 4d). At low collision energies, these indirect reactions (Fig. 4b–d) represent about 10 % of the trajectories, while at high collision energies, they represent practically 0 %. Classically, these nearly trapped trajectories, similar to the motion of a spring or to the popular 'yo-yo' toy, are a straightforward consequence of the presence of the reactant well. This behaviour of the direct and indirect trajectories is similar to that observed for the $Cl + NH_3$ reaction [8], although at higher collision energies in the perdeuterated case (see Table 1).

However, considering the arguments given in Sect. 3.1 about the possible artefact of the QCT calculations at low energies, confirmed with the reduced dimensionality QM

Fig. 4 Representative plots of different reaction mechanisms, shown as the evolution of the N–D (*dashed line*) and D–Cl (*solid line*) distances as a function of time (ps), using the PES-2010 surface: **a** direct trajectory; **b** indirect trajectory in the entry channel; **c** indirect trajectory in the exit channel; and **d** indirect trajectory in the entry and exit channels
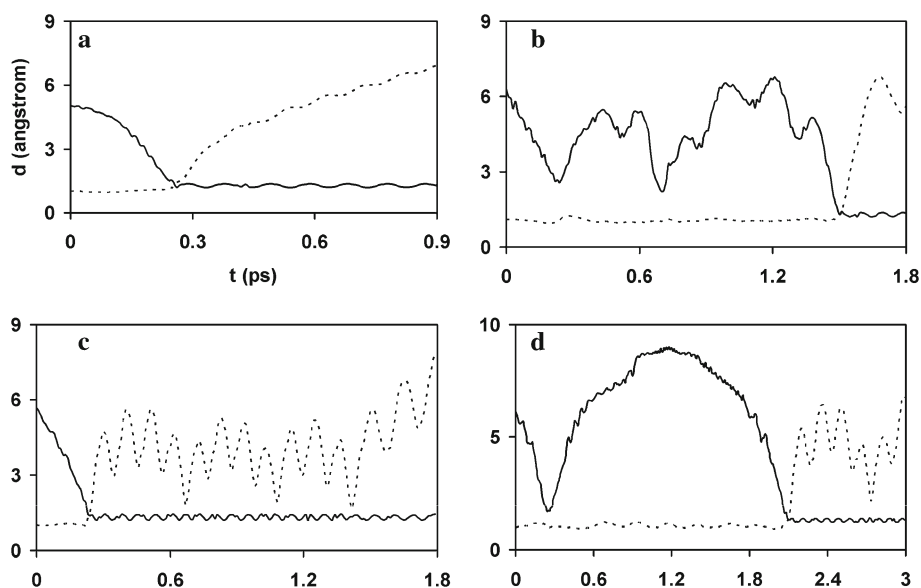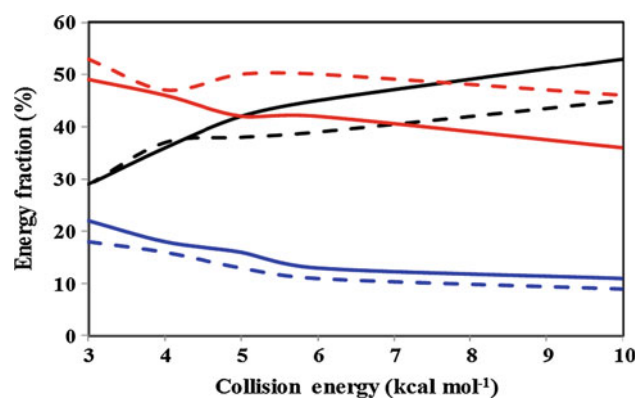
Fig. 5 QCT results using the PES-2010 surface for the Cl + ND$_3$ reaction (*solid lines*) and Cl + NH$_3$ reaction (*dashed lines*) for the average translation energy, f_{trans}, (*black line*), and internal energy of the products, f_{int}(ND$_2$/NH$_2$), (*blue line*) and f_{int}(DCl/HCl), (*red line*), in percentages, as a function of the collision energy

Fig. 6 QCT vibrationally resolved DCl/HCl($v = 0$) rotational distributions using the PES-2010 surface for the Cl + ND$_3$ reaction (*blue lines*) and Cl + NH$_3$ reaction (*red lines*), for two collision energies—low and high: 4.0 and 10.0 kcal mol^{-1} (*upper* and *lower panels*, respectively)

calculations, one must be very careful about the importance of the indirect mechanism in these reactions.

3.3 Product energy partitioning

Although the product energy partitioning is an experimental property that is easily measurable and could represent a test of our theoretical results (PES and dynamic methods), unfortunately neither are data on this property currently available. Figure 5 shows, in percentages, the average translational energy, f_{trans}, and internal energy of the products, $f_{int(ND2)}$ and $f_{int(DCl)}$, as functions of the collision energy, together with the QCT results for the Cl + NH$_3$ reaction for comparison. The values and behaviour of the two reactions are similar—the product internal energies decrease with collision energy, while the translational energy increases by the same amount. This similarity in behaviour indicates that the partitioning of the energy in the products is fairly insensitive to the isotope effect and to the depth of the reactant well.

3.4 Product rovibrational distributions

At all the collision energies analysed, 3.0–10.0 kcal mol^{-1}, the ND$_2$ and DCl products appeared essentially in their respective vibrational ground state, $v = 0$.

The QCT vibrationally resolved DCl ($v = 0$) rotational distributions for the Cl + ND$_3$ reaction are plotted in Fig. 6 for two collision energies –low and high: 4.0 and 10.0 kcal mol^{-1}—together with the results for the Cl + NH$_3$ reaction for comparison. The two reactions present similar shapes, with the diatomic product appearing with a broad and hot rotational distribution. Note, however, that the QCT methods tend to give rotational distributions that are hotter and broader than experiment and the results
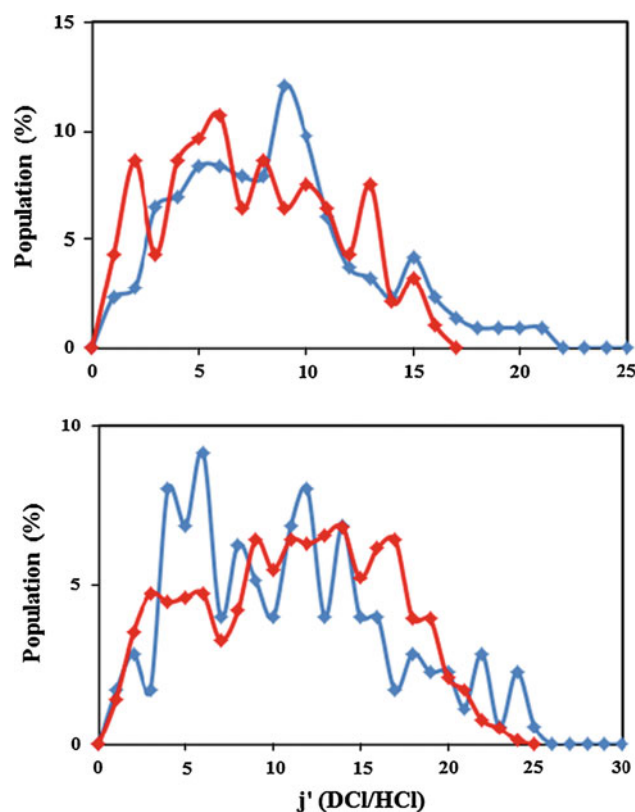
of quantum–mechanical methods [26–31], an effect which is an artefact of these classical methods. Given the linear transition state for this reaction, which generates little torque during an impulsive release, a priori one might have expected a colder rotational distribution. Thus, the hot rotational distribution for the two reactions could be related to the existence of deep wells in the entry and exit channels. The attractive interaction in the reactant well provides more torque and therefore more rotational excitation, while the attractive interaction in the product well may provide a more pronounced torque on the two separating products, and a greater randomization of the rotational energy, thereby explaining the hot and broad rotational distribution. Unfortunately, there are neither experimental nor other theoretical results available for comparison.

3.5 Product angular distribution

The product angular distribution is doubtless one of the most sensitive dynamics properties with which to test the quality of the surface and to examine the role of intermediate complexes in the reaction. To analyse the isotope effect, we consider two collision energies: the threshold of reaction (4.0 kcal mol^{-1}

for the perdeuterated reaction, and 2.5 kcal mol^{-1} for the perprotio ones), and an energy clearly above this threshold, 10.0 kcal mol^{-1}, in both cases. The QCT angular scattering distribution of the diatomic product (DCl or HCl) with respect to the incident chlorine atom (obtained as the differential cross section, DCS, which was fitted using the Legendre moment method [32]) is plotted in Fig. 7.

In the threshold of reaction (Fig. 7, upper panel), while the Cl + NH$_3$ reaction showed a forward–backward shape, the Cl + ND$_3$ reaction presented a more isotropic behaviour. This is clearly related to the indirect trajectories: 38 versus 10 %. The indirect mechanism is characterised by a forward–backward distribution, as is typical of reactions occurring via a deep minimum. Thus, due to the smaller collision energy for the Cl + NH$_3$ reaction (and consequently the greater importance of the indirect mechanism), the effect of the entry valley is more pronounced.

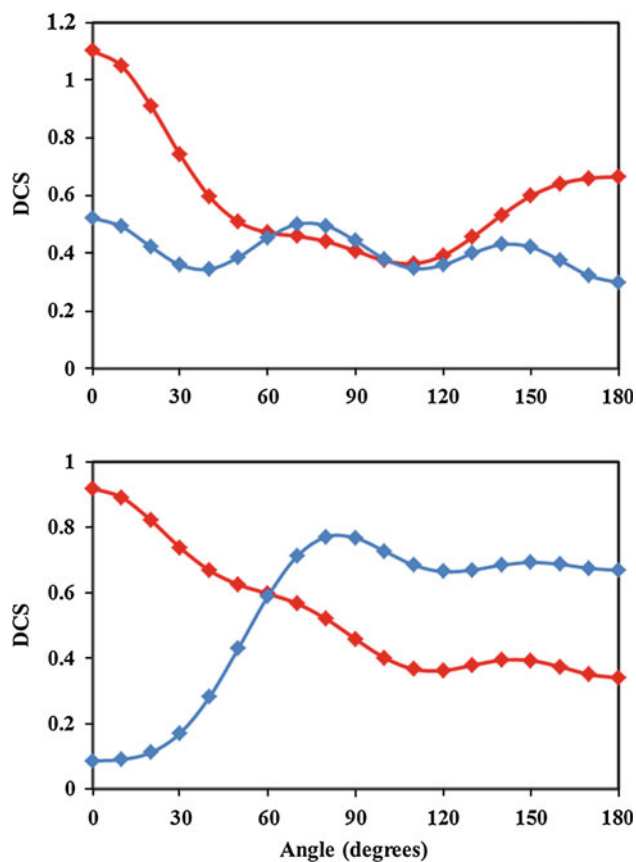At 10.0 kcal mol^{-1}, the collision energy is enough to surpass the reaction threshold, the direct trajectories

predominate (100 % for both reactions), and the behaviour is very different (Fig. 7, lower panel). While the Cl + NH$_3$ reaction showed a sideways-forward distribution, associated with a stripping mechanism with high impact parameters, the Cl + ND$_3$ reaction shows a sideways-backward scattering, associated with lower impact parameters. Indeed, at 10.0 kcal mol^{-1}, the b$_{max}$ values are 3.2 and 2.7 Å, respectively, for the NH$_3$ and ND$_3$ systems. To analyse more clearly the influence of the impact parameter, Fig. 8 shows the opacity function (reaction probability vs. impact parameter) in the two energy regimes analysed, threshold (upper panel) and high (lower panel). It is known that the contribution of smaller impact parameters leads to backward-scattered products. In both energy regimes, the Cl + ND$_3$ reaction presents a larger contribution of small impact parameters, favouring backward scattering.

In summary, the product scattering distribution is the only dynamics property analysed which is strongly

Fig. 7 QCT results using the PES-2010 surface for the DCl product angular distribution (with respect to the incident Cl) for the Cl + ND$_3$ → DCl + ND$_2$ reaction (*blue lines*) and for the HCl product angular distribution for the Cl + NH$_3$ → HCl + NH$_2$ reaction (*red lines*). *Upper panel* low collision energy—4.0 kcal mol^{-1} for the Cl + ND$_3$ reaction and 2.5 kcal mol^{-1} for the Cl + NH$_3$ reaction. *Lower panel* high collision energy—10.0 kcal mol^{-1} for both reactions

Fig. 8 Opacity function (reaction probability vs. impact parameter) using the PES-2010 surface for the Cl + ND$_3$ reaction (*blue lines*) and Cl + NH$_3$ reaction (*red lines*). *Upper panel* low collision energy—4.0 kcal mol^{-1} for the Cl + ND$_3$ reaction and 2.5 kcal mol^{-1} for the Cl + NH$_3$ reaction. *Lower panel* high collision energy— 10.0 kcal mol^{-1} for both reactions. In each case, the distributions are normalised so that the area under the common levels is the same

dependent on the isotope effect, and it seems to indicate that there are different mechanisms for the two isotopes. Clearly more theoretical and new experimental studies will be needed to understand these differences.

4 Conclusions

We have described an extensive state-to-state dynamics study performed on an analytical potential energy surface (PES-2010) recently developed for the gas-phase abstraction reaction $Cl(^2P) + ND_3 \rightarrow DCl + ND_2$, using reduced dimensionality quantum mechanical (7D-QM) and quasi-classical trajectory (QCT) calculations. In the QCT study, the zero-point energy problem was explicitly considered to avoid leakage along the trajectories, and the collision energy ranged from 3.0 to 10.0 kcal mol^{-1}. The dynamics results were compared with those for the perprotio reaction in order to analyse the isotope effect.

The QCT product energy partitioning showed that the translational energy increases with collision energy, while the internal energy of the products, DCl and ND_2, decreases by the same amount. At all the collision energies studied, the DCl and ND_2 products appeared practically in their respective vibrational ground states, while the $DCl(v = 0)$ rotational distribution was broad and hot. Given that this reaction presents a linear transition state geometry, the hot distribution could be related to the existence of wells in the entry and exit channels. The attractive interaction in these wells could provide a more pronounced torque and greater randomization of the rotational energy. These two dynamics properties—product energy partitioning and rovibrational distributions—showed similar behaviour to the case of the $Cl + NH_3$ reaction, indicating that they are insensitive to the isotope effect.

With respect to the variation of the reaction probability with the collision energy, the $Cl + ND_3$ and $Cl + NH_3$ reactions present similar behaviour, with the probability increasing smoothly from the reaction threshold, which is the typical behaviour for reactions with a barrier. Hence, the effect of the entry valley is negligible for these reactions. We find that the product scattering distribution is the only dynamics property strongly dependent on the isotope effect, and it seems to indicate that there are different mechanisms for the two isotopes. The smaller maximum impact parameter for the perdeuterated reaction favours backward scattering, as compared to forward scattering for the perprotio analogue. In sum, the deuteration (through the ZPE effect) changes the shape of the adiabatic energy profile, possibly modifying the reaction mechanism.

Unfortunately, no experimental dynamics information is available for this polyatomic system, whether perprotio or perdeuterated. We therefore trust that this and similar theoretical studies might encourage future experimental research into this complex reaction and its mechanism, especially at low energies.

References

1. Miller JA, Bowman CT (1989) Prog Energy Combust Sci 15:287
2. Westenberg AA, DeHaas N (1977) J Chem Phys 67:2388
3. Gao Y, Alecu IM, Hsieh PC, Morgan BP, Marshall P, Krasnoperov LN (2006) J Phys Chem A 110:6844
4. Kondo S, Tokuhashi K, Kaise M (2000) J Hazard Mat A 79:77
5. Xu ZF, Lin MC (2007) J Phys Chem A 111:584
6. Monge-Palacios M, Espinosa-Garcia J (2010) J Phys Chem A 114:4418
7. Monge-Palacios M, Rangel C, Corchado JC, Espinosa-Garcia J (2012) Int J Quantum Chem 112:1887
8. Monge-Palacios M, Yang M, Espinosa-Garcia J (2012) Phys Chem Chem Phys 14:4824
9. Monge-Palacios M, Corchado JC, Espinosa-Garcia J (2012) Phys Chem Chem Phys 14:7497
10. Chase MW, Davis CA, Downey JR, Frurip DJ, McDonald RA, Syverud AN (1985) JANAF Thermochemical tables, J Phys Chem Reference Data vol 14, Suppl 1
11. Song Y, Qian XM, Lan KC, Ng CY, Liu J, Chen W (2001) J Chem Phys 115:2582
12. Porter RN, Raff LM (1976) In: Miller WH (ed) Dynamics of molecular collisions, part B, Ed. Plenum Press, New York
13. Truhlar DG, Muckerman JT (1979) In: Bernstein RB (ed) Atommolecules collision theory. Plenum Press, New York
14. Raff LM, Thompson DL (1985) In: Baer M (ed) Theory of chemical reaction dynamics, vol 3. CRC Press, Boca Raton
15. Hase WL, Duchovic RJ, Hu X, Komornicki A, Lim KF, Lu Dh, Peslherbe GH, Swamy KN, van de Linde SR, Varandas AJC, Wang H, Wolf RJ (1996) VENUS96: a general chemical dynamics computer program. QCPE Bull. 16:43
16. Varandas AJC (1994) Chem Phys Lett 225:18
17. Duchovic RJ, Parker MA (2005) J Phys Chem A 109:5883
18. Guo Y, Thomson DL, Sewell TD (1996) J Chem Phys 104:576
19. Corchado JC, Espinosa-Garcia J, Yang M (2011) J Chem Phys 135:014303
20. Gray JN, Garrett BC, Truhlar DG (1979) J Chem Phys 70:5921
21. Kudla K, Schatz GC (1993) Chem Phys 175:71
22. Bonnet L (2008) J Chem Phys 128:044109
23. Yang M, Corchado JC (2007) J Chem Phys 126:214312
24. Varandas AJC (1993) J Chem Phys 99:1076
25. Varandas AJC (2001) Chem Phys Lett 340:62
26. Gerrity DP, Valentini JJ (1984) J Chem Phys 81:1298
27. Kliner DAV, Rinnen KD, Zare RN (1990) Chem Phys Lett 166:107
28. Dean BD, Ayers JD, Fernandez-Alonso F, Zare RN (2001) J Phys Chem A 105:2228
29. Dean BD, Ayers JD, Fernandez-Alonso F, Zare RN (2002) J Chem Phys 116:6634
30. Pomerantz AE, Ausfelder F, Zare RN, Althorpe SC, Aoiz FJ, Bañares L, Castillo JF (2004) J Chem Phys 120:3244
31. Xiao T, Bowman J, Duff JW, Braunstein M, Ramachandran B (2005) J Chem Phys 122:014301
32. Truhlar DG, Blais NC (1977) J Chem Phys 67:1532

Manipulating the singlet–triplet transition in ion strings by nonresonant dynamic Stark effect

Patricia Vindel-Zandbergen · Mirjam Falge ·
Bo Y. Chang · Volker Engel · Ignacio R. Sola

Abstract Using strong laser pulses, we show that it is possible to control the spin state in a model system based on a two-electron extension with spin couplings of the Shin–Metiu Hamiltonian, truncated to account for the lowest electronic energy states. We consider two different models depending on the number of electronic states included in the calculation. The initial electronic state determines when the spin state is stable or not in the absence of an external field. In the latter case, by nonresonant dynamic Stark effect, we show that it is possible to avoid spin transitions with strong fields, using different pulse frequencies. This effective spin locking requires minimizing absorption to excited singlets as well as decoupling the singlet and triplet electronic states. In the first case, we show that it is possible to force the spin transition by a combination of two pulses, a chirped pulse and a transform limited pulse, where the time-delay must be chosen to maximize spin switching on a different electronic state. Our results show that forcing the spin switching is a more difficult goal than avoiding it and that this goal becomes highly restricted when many electronic pathways or multi-photon processes are available.

P. Vindel-Zandbergen · I. R. Sola (✉)
Departamento de Quimica Fisica, Universidad Complutense,
28040 Madrid, Spain
e-mail: ignacio@tchiko.quim.ucm.es

M. Falge · V. Engel
Institut für Physikalische und Theoretische Chemie und Röntgen Research Center for Complex Material Systems, Universität Würzburg, Campus Nord, Emil-Fischer-Str. 42, 97074 Würzburg, Germany

B. Y. Chang
School of Chemistry (BK21), Seoul National University,
Seoul 151-747, Republic of Korea

1 Introduction

Laser control of quantum processes is an active arena particularly in the application of femtosecond laser pulses to quantum-state excitation and unimolecular reactions [1–3]. Successful laser control experiments have been reported in an increasing variety of physical systems, including complex chemical and biological processes [1, 2].

Addressing the dynamics of complex systems, the mechanism of the optical control is often understood from the spectral features of the pulses [4] (albeit with important caveats [5]), implying correlations between the pulse frequencies and the Hamiltonian resonances, with relative phases of the spectrum adding important dynamical information concerning the cross-talk of the resonances [6]. However, this picture is no longer valid when strong fields are used. Nonresonant effects may then completely shift or distort the Hamiltonian spectrum, which is no longer independent of the pulse spectrum [7–13]. It is possible to base the control mechanism solely on the effects of the nonresonant dynamic Stark effect (NRDSE) [14–18]. This strategy is particularly useful when the aim of the control problem is to "disconnect" an undesired transition [16, 17]. The NRDSE is behind many interesting control scenarios involving molecular alignment [19–26], the control of photodissociation reactions [14, 15, 27–29], or the control

of singlet–triplet transitions [16–18], which is also the motivation of this work.

The spin–orbit coupling is a relativistic effect that mainly affects the Hamiltonian of systems with heavy nuclei [30]. In molecules, the spin–orbit transition is one of the main sources that causes the breakdown of the Born–Oppenheimer approximation in the dynamics of excited states, inducing intersystem crossing (ISC) and altering the spectroscopy and photodissociation of molecules and the relaxation rates in complex biological systems [31, 32]. Although in not heavy atoms, the mass is often invoked as a reason to neglect the ISC processes, simulations of the dynamics of polyatomic molecules with light atoms have shown that the effect of ISC must be taken into account in many occasions [33–35]. In solids, the spin coupling between different states can be used to create states of mixed multiplicity and to prepare optical spin switches [36–39]. These are important in spintronics [40] and may have potential application as quantum information storage or quantum information processing devices [41].

Rather than using a realistic Hamiltonian for a particular system, in this work we use the very generic Shin–Metiu (SM) Hamiltonian that models charge transfer processes in some matrix environments [42–47]. The original SM model is conveniently extended to include two electron processes with singlet–triplet couplings [48, 49]. This extended Shin–Metiu (ESM) model is particularly interesting because one can treat electron and nuclear processes at the same level, without invoking the Born–Oppenheimer approximation. Recently, we have shown that in the dynamics of the ESM with strong fields, ionization is important and it is in fact the dominant process [49]. In this work, however, we neglect ionization and use the adiabatic states (electronic potential curves) obtained from the electronic Hamiltonian of the ESM in the Born–Oppenheimer approximation to simulate the dynamics of the nuclear coordinates under the influence of laser pulses and singlet–triplet couplings. For this reason, in Sect. 2 we study the ESM and give the different electronic potentials and couplings. We also show the spin-coupling dynamics of the system starting from different initial states.

The goal of the study is to survey the extent to which optical control of the singlet–triplet transfer is possible, either by adiabatically "freezing" the spin populations by means of a laser, such that the rate of spin transfer is substantially reduced, or by forcing the spin transition when this is negligible. The control of these processes is based either on a pump–dump pulse sequence or on the NRDSE scheme, whose principles are briefly reviewed for the system of study in Sect. 3. Preliminary results have shown that the NRDSE scheme can induce spin freezing, although ionization takes over the system in <50 fs. In this work, we study in detail the processes of both spin locking

and spin switching. Two models of different complexity (involving different number of electronic states and couplings) are introduced. In Sect. 4, we analyze spin locking for both models, outlining the different roles of the pulse parameters, and in Sect. 5, we study the more difficult process of spin-state switching. Finally, Sect. 6 is the summary and conclusions.

2 Hamiltonian model and field-free dynamics

The electronic potentials and dipole couplings are obtained from the Born–Oppenheimer limit of an extended version of the Shin–Metiu model (ESM), including spin couplings. The ESM consists of an ion (coordinate R) and two electrons (coordinates x, y) confined to move in a single dimension, interacting with each other and two additional fixed ions through screened Coulomb interactions. Because of exchange symmetry, the spatial wave function is either symmetric ($\psi^S(x, y, R, t)$, singlet (S)) or anti-symmetric ($\psi^T(x, y, R, t)$, triplet (T)) with respect to electron exchange. Details are given in ref. [49].

The spin uncoupled Born–Oppenheimer potential energy curves and electronic wave functions are obtained by imaginary time propagation of an electronic wave function of singlet/triplet symmetry $\psi^M(x, y, t; R)$ ($M = S, T$) at fixed nuclear position R

$$\psi^M(x, y, t; R) = e^{-H_{el}t}\psi^M(x, y, t = 0; R) \tag{1}$$

with the electronic Hamiltonian operator

$$H_{el}(x, y, R) = -\frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{1}{2}\frac{\partial^2}{\partial y^2} + V(x, y, R) \tag{2}$$

for different values of R. In the absence of spin coupling in the electronic Hamiltonian, the wave function symmetry is conserved. For long times, t, $\psi^M(x, y, t; R)$ converges to the ground state $\varphi_1^M(x, y, R)$ of the respective symmetry and the norm of the wave function decreases as $N(t) = e^{-2V_1(R)t}$. After the ground state is determined, the next higher eigenstate $\varphi_2^M(x, y, R)$ is obtained by another imaginary time propagation starting from an initial state where the ground state is projected out:

$$\begin{aligned}\tilde{\psi}^M(x, y, t = 0; R) &= \psi^M(x, y, t \\ &= 0; R) - \langle\psi^M(x, y, t = 0; R)|\varphi_1^M(x, y, R)\rangle \\ &\quad \cdot \varphi_1^M(x, y, R).\end{aligned} \tag{3}$$

Higher eigenstates are calculated successively using the same scheme. In this way, the basis of Born–Oppenheimer electronic states $\{\varphi_j^M(x, y, R)\}$ and their respective potential curves $V_j^M(R)$ are obtained. Dipole and spin couplings are evaluated with respect to these electronic basis as well.

To monitor spin transitions, we introduce in a heuristic way a spin-coupling term of the form

$$J(x, y) = \lambda(x - y), \qquad (4)$$

where the coupling strength-parameter is chosen as $\lambda = 1.028 \times 10^{-3}$ eV/Å. For simplicity, we neglect the possible dependence of the coupling on the nuclear coordinate R. $J(x, y)$ creates new potential energy terms that couple (and mix) the Born–Oppenheimer electronic potentials. They are obtained from

$$V_{S_n, T_m}(R) = \int dx \, dy \, \varphi_n^S(x, y, R) J(x, y) \varphi_m^T(x, y, R). \qquad (5)$$

where φ_n^S and φ_m^T denote the electronic eigenfunctions of singlet and triplet symmetry, respectively. To monitor laser-driven dynamics, we calculate the transition dipole moments

$$\mu_{M_n, M_m}(R) = \int dx \, dy \, \varphi_n^M(x, y, R)(-x - y)\varphi_m^M(x, y, R). \qquad (6)$$

where $M = S$ or T.

In Fig. 1, we show the first electronic singlet and triplet potentials obtained from the ESM Hamiltonian for a particular choice of parameters (screening parameters $R_c = R_e = R_f = 1.5$ Å and charge numbers $Z_1 = Z_2 = Z = 1$ for the ESM Hamiltonian, see ref. [49]), neglecting the spin-coupling terms. The general feature of these potentials is that the singlet and triplet states come on pairs, showing either a double-well structure or a single equilibrium geometry at $R = 0$. Because the two electrons tend to be at opposite sides of the central ion (that is, each facing a different end-ion), the exchange symmetry of the electronic wave function does not lead to substantial energy differences between the singlet and the triplet wave functions. Thus, the singlet–triplet transfer due to the V_{S_j, T_j} terms (for $j = 1, 3$, etc.) is expected to be large. However, V_2^S and V_2^T (V_5^S and V_5^T as well, for higher energies) have a different structure near the equilibrium geometry and the spin-coupling mechanism is not efficient for any reasonable value of λ.

In addition to the spin coupling, nonadiabatic coupling terms could have been included. The energy separation between the singlet curves (and between the triplet curves as well) makes internal conversion processes very unlikely, as shown in the numerical results obtained solving the full vibronic Hamiltonian [49]. Thus, we neglect their contribution in the present work.

In Fig. 2, we show the population dynamics starting either from V_1^S or from V_2^S in the absence of any external field. We solve the time-dependent Schrödinger equation for the nuclear motion in two electronic states (S_j and T_j, for j either 1 or 2),
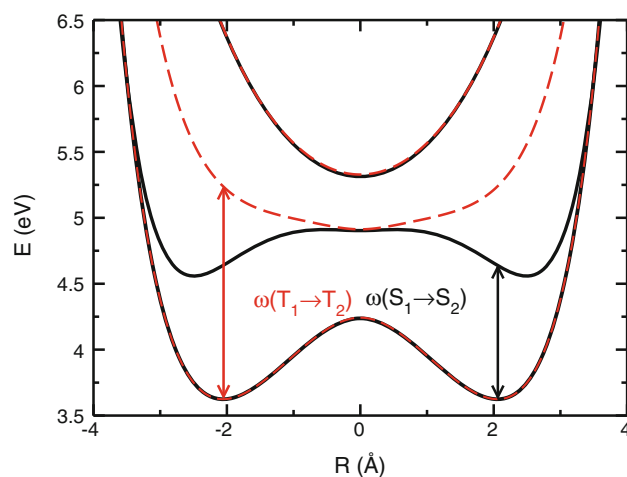
Fig. 1 Potential *curves* for the lowest three electronic states with singlet ($M = S$, *solid lines*) and triplet ($M = T$, *dashed lines*) symmetries. Notice that T_1 and T_3 practically coincide with S_1 and S_3, while S_2 and T_2 have different equilibrium geometries. Also shown in the *figure* are the frequencies for the Franck–Condon resonant transition between $S_1 \rightarrow S_2$ and $T_1 \rightarrow T_2$

$$i \frac{\partial}{\partial t} \begin{pmatrix} \psi_j^S \\ \psi_j^T \end{pmatrix} = -\frac{1}{2m} \frac{d^2}{dR^2} \begin{pmatrix} \psi_j^S \\ \psi_j^T \end{pmatrix} + \begin{pmatrix} V_j^S(R) & V_{S_j, T_j}(R) \\ V_{S_j, T_j}(R) & V_j^T(R) \end{pmatrix} \begin{pmatrix} \psi_j^S \\ \psi_j^T \end{pmatrix} \qquad (7)$$

using the split-operator with fast Fourier transform technique. In Eq. (7), m is the mass of a proton.

For the choice of the coupling parameter λ, we observe full singlet–triplet switching in a period $T \sim 120$ fs starting in V_1^S, and practically no triplet contamination starting from V_2^S (the maximum population in T_2 is 3×10^{-5}). The goal of this work is to design optical processes to avoid the spin

Fig. 2 Singlet triplet population transfer in the absence of an external field. In *solid lines*, we show the population of S_1 and T_1 when the dynamics starts in S_1. In *dashed line,* we show the population of S_2 when the dynamics starts from this excited state. In this case, the population in T_2 remains practically zero and it is not shown

dynamics from V_1^S (that is, to induce *spin locking*) and to force it starting from V_2^S (that is, to induce *spin switching*). To that end, we will use the NRDSE or a pump–dump scheme.

3 The NRDSE scheme

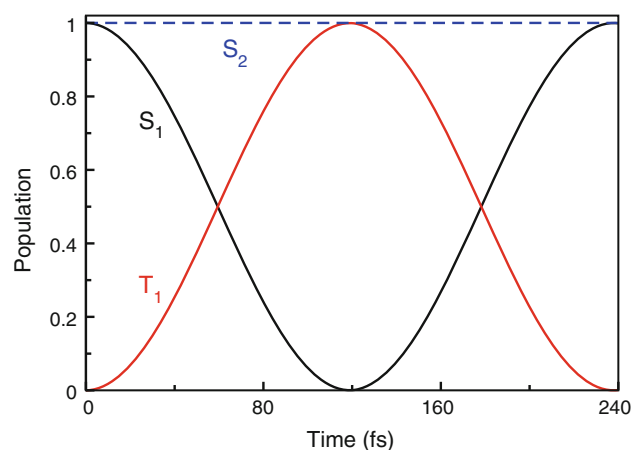In order to understand the physical mechanism underlying the control of the spin coupling via the NRDSE, it is enough to consider the simplest polarizable two-level system with one singlet (S) and one triplet (T) state, coupled via V_{ST}, whose energy can be Stark shifted by a nonresonant electromagnetic field with envelope $\mathcal{E}(t)$ [50]. Suppose that for a given field frequency, the dynamic polarizabilities for the singlet and triplet states are $\alpha_j(\omega)$ ($j = S, T$, we will often omit the frequency dependence in the following). Then choosing the field-free singlet energy as the zero point of energy, the effective Hamiltonian for the system is

$$H = \begin{pmatrix} -\alpha_S \mathcal{E}(t)^2/2 & V_{ST} \\ V_{ST} & \Delta(0) - \alpha_T \mathcal{E}(t)^2/2 \end{pmatrix} \quad (8)$$

where $\Delta(0)$ is the energy difference between the states in the absence of the field. One can choose an instantaneous zero point of energy that includes the Stark-shift contribution to the singlet energy, making the Hamiltonian

$$H = \begin{pmatrix} 0 & V_{ST} \\ V_{ST} & \Delta(\mathcal{E}) \end{pmatrix} \quad (9)$$

where $\Delta(\mathcal{E}) = \Delta(0) - (\alpha_T - \alpha_S)\mathcal{E}^2(t)/2$. The Rabi formula can then be used to predict the spin-coupling dynamics. Suppose the system is initially in the singlet state, assuming for simplicity a constant envelope $\mathcal{E}$, the population in the triplet will be

$$P_T = \left(\frac{V_{ST}}{\Omega_e}\right)^2 \sin^2 \Omega_e t \quad (10)$$

where $\Omega_e = \sqrt{V_{ST}^2 + \Delta(\mathcal{E})^2}$. For instance, if $\Delta(0) \ll |V_{ST}|$ (as in the ESM model, starting in V_1^S where $\Delta(0) = 0$), substantial spin transfer will occur at $\tau_{ST} = \pi / 2V_{ST}$. This spin-mixing dynamics can be halted while a laser pulse acts if $\Delta(\mathcal{E}) \gg V_{ST}$. Conversely, if $\Delta(0) \gg V_{ST}$ (as in the ESM starting in V_2^S) full population transfer can be induced by making $\Delta(\mathcal{E}) = 0$ and timing the pulse duration exactly as $\tau_e = \pi/2\Omega_e$. Thus, a laser-controlled spin switch can be created whenever the dynamic polarizability of the singlet state is different from that of the triplet state, $\alpha_S \neq \alpha_T$. Moreover, the dynamic polarizabilities depend on the pulse frequency ω as well. Additionally, the two-level theory can be used to propose alternative (more robust) quantum control scenarios [50].

The extension of the effective 2-level Hamiltonian to the nuclear Hamiltonian in the Born–Oppenheimer

approximation is straightforward. Neglecting population absorption one can obtain [51]

$$H = -\frac{1}{2m}\frac{d^2}{dR^2}I$$
$$+ \begin{pmatrix} V_1^S(R) - \alpha_S(R,\omega)\mathcal{E}(t)^2/2 & V_{S_1,T_1}(R) \\ V_{S_1,T_1}(R) & V_1^T(R) - \alpha_T(R,\omega)\mathcal{E}(t)^2/2 \end{pmatrix} \quad (11)$$

Since $V_{S_1,T_1}(R)$ is almost constant (and small) around the equilibrium geometries of the V_1^S potential and $V_1^S(R) \approx V_1^T(R)$, in the absence of the field, the population dynamics is practically that of a two-level system, where the singlet–triplet population transfer does not lead to wave-packet dynamics. This prediction explains the excellent agreement between the population dynamics in Fig. 2 and the result of applying the Rabi formula. For strong fields, however, Eq. (11) does not give quantitative results. Firstly, because it is difficult to evaluate the dynamic polarizabilities $\alpha_j(R, \omega)$, and secondly, because multi-photon absorption to excited states usually cannot be neglected. Therefore, in the following results, we solve the time-dependent Schrödinger equation including different sets of electronic states.

Within the Born–Oppenheimer approximation, the nuclear dynamics Hamiltonian in the presence of the control field $\epsilon_c(t)$, is

$$H^{eff} = -\frac{1}{2m}\frac{d^2}{dR^2}I$$
$$+ \begin{pmatrix} V_1^S & -\mu_{S_1,S_2}\epsilon_c(t) & V_{S_1,T_1} & V_{S_1,T_2} \\ -\mu_{S_1,S_2}\epsilon_c(t) & V_2^S & V_{S_2,T_1} & V_{S_2,T_2} \\ V_{T_1,S_1} & V_{T_1,S_2} & V_1^T & -\mu_{T_1,T_2}\epsilon_c(t) \\ V_{T_2,S_1} & V_{T_2,S_2} & -\mu_{T_1,T_2}\epsilon_c(t) & V_2^T \end{pmatrix} \quad (12)$$

where I is the 4×4 unit matrix. In Eq. (12), we have restricted the Born–Oppenheimer expansion to the minimal symmetric set (MSS) of electronic states that can represent the dynamics of the system under the influence of an external field that controls the singlet–triplet transition. In the numerical calculations, we use "square" laser pulses with fast ramps of $\sin^2(\pi(t - t_0)/\tau_s)$ form. The switching on/off time is set to $\tau_s = 25$ fs.

The extension of the model to include a more complete set of electronic states is straightforward. In addition to the MSS, we have considered a model including the 6 lowest singlet and triplet states (LSS). These 12 Born–Oppenheimer states span the energy range of any 3-photon process starting in S_1 for all laser frequencies used in this work. It should be noted though that the ESM model favors fast multi-photon excitation from the initial state to the ionization continuum [49]. The ionization process is not considered in this work. However, as pointed out in [49], the ionization process does not affect considerably the spin

dynamics and some observables such as the averaged spin angular momentum of the system, hint that the controlled process is still acting on the small population surviving ionization.

4 Spin-state locking

4.1 Dynamics using the MSS model

The results with the MSS model can be used as proof-of-principles for the optical control of the singlet–triplet transition. The initial state is the ground vibrational eigenstate of S_1, which is an even function with maximum probability on both potential wells. As an example, in Fig. 3, we show the population dynamics in the four electronic states using a single control pulse ϵ_c with different parameters that give best results over different frequency intervals. The overall duration of the pulse is set to 200 fs to display more clearly the effect of spin locking.

We have found robust solutions in the space of parameters for a broad range of peak intensities and frequencies. Although strong pulses are needed to overcome the small transition dipole, the results are not very sensitive to the peak intensity. In general, for larger peak amplitude ϵ_0 the
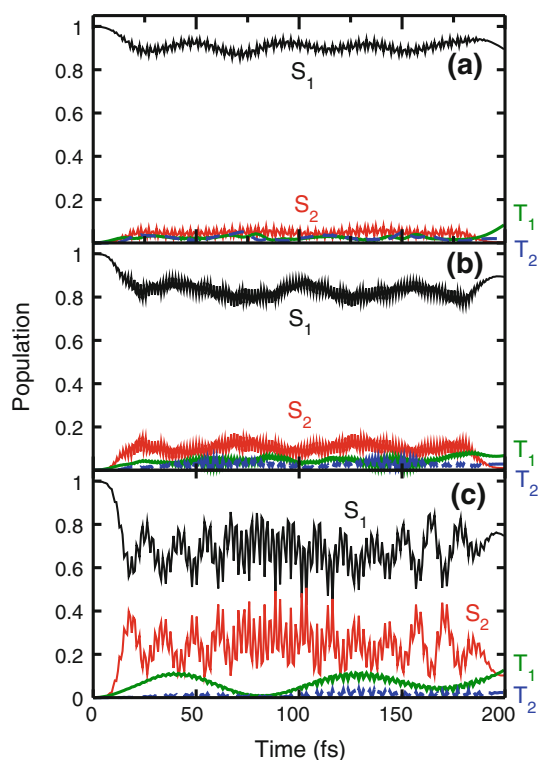

Fig. 3 Electronic states population dynamics in the MSS model with different pulse frequencies and intensities: **a** with $\hbar\omega = 1.63$ eV and $I_0 = 10.1$ TW/cm², **b** with $\hbar\omega = 1.41$ eV and $I_0 = 10.1$ TW/cm² and **c** with $\hbar\omega = 0.87$ eV and $I_0 = 8.95$ TW/cm²

population is more evenly spread between the singlet potentials with smaller excitation of the triplet states.

The frequency must be chosen detuned from the $S_1 \rightarrow S_2$ resonance ($\hbar\omega(S_1 \rightarrow S_2) \approx 1$ eV), defined in the initial Franck–Condon region, to avoid absorption to a different state. In the example of Fig. 3a, $\hbar\omega = 1.68$ eV is slightly above the resonance between T_1 and T_2 ($\hbar\omega(T_1 \rightarrow T_2) \approx 1.6$ eV) and thus clearly blue shifted with respect to $\omega(S_1 \rightarrow S_2)$. Therefore, we expect a larger positive Stark-shift on V_1^T than in V_1^S, allowing the effective decoupling of these potentials with $\epsilon_0 = 0.017$ a.u. or larger (implying a peak intensity I_0 of 10.1 TW/cm²).

Other frequency intervals can be chosen: If ω is between $\omega(S_1 \rightarrow S_2)$ and $\omega(T_1 \rightarrow T_2)$ the dynamic Stark-shift should be positive in V_2^S but negative in V_1^T. Weaker fields could then decouple the electronic states. However, it is more difficult to avoid populating the V_2^S potential. For the results in Fig. 3b, we use $\hbar\omega = 1.39$ eV and the same peak intensity ($I_0 = 10.1$ TW/cm²).

With $\omega < (\omega(S_1 \rightarrow S_2), \omega(T_1 \rightarrow T_2))$ both V_2^S and V_1^T experience a negative Stark-shift (larger in V_1^S) that give similar results as those shown in case (a). However, multiphoton excitation to S_2 becomes resonant and as a result, more population is excited to S_2. The different shapes of the potentials induce nuclear motion that shows up in some Rabi oscillations between the electronic populations in S_1 and S_2. In Fig. 3c, for $\hbar\omega = 0.84$ eV, we have slightly reduced the pulse intensity to $I_0 = 8.95$ TW/cm².

4.2 Dynamics in the LSS model

When considering a larger set of electronic states, the control of the process becomes more complex. On one hand, the different electronic states normally *add* to the polarizabilities, increasing the effect of the Stark effect: weaker pulses could in principle be used. On the other hand, one has to care about many possible resonances ($\omega(S_1 \rightarrow S_j), \omega(S_j \rightarrow S_k)$) that lead to absorption and multi-photon ladder climbing of electronic states. Given the intensities required to decouple the singlet–triplet transition, the second effect is clearly more dominant than the first one, making the spin-state locking quite more challenging in the LSS model than in the MSS model.

In Fig. 4, we analyze the time-averaged populations in the triplets and excited singlets (all except S_1) as a function of the pulse frequency and intensity

$$\langle P_T \rangle = \frac{1}{\tau_l} \sum_{j=1}^{6} \int_0^{\tau_l} \left\langle \psi_j^T(t) \middle| \psi_j^T(t) \right\rangle dt$$

$$\langle P_{ES} \rangle = \frac{1}{\tau_l} \sum_{j>1}^{6} \int_0^{\tau_l} \left\langle \psi_j^S(t) \middle| \psi_j^S(t) \right\rangle dt$$

(13)

where $\tau_l = 200$ fs is the pulse duration. The goal of decoupling the $S_1 \rightarrow T_1$ transition involves minimizing $\langle P_T \rangle$ (avoiding the singlet–triplet transition) and $\langle P_{ES} \rangle$ (avoiding the absorption to other electronic states). Whereas with stronger pulses, $\langle P_T \rangle$ is small for all the range of frequencies explored in this work, $\langle P_{ES} \rangle$ is particularly large at lower frequencies, obviously around $\omega(S_1 \rightarrow S_2)$, but also very much until $\omega > \omega(T_1 \rightarrow T_2)$. Disregarding vibrational excitation, the LSS model includes all singlet states that can be accessed by three photons at $\hbar\omega \sim 1.7$ eV.

In Fig. 5, we show the population dynamics for three different cases. For the three frequency windows $\omega < \omega(S_1 \rightarrow S_2)$ (smaller), $\omega(S_1 \rightarrow S_2) \leq \omega \leq \omega(T_1 \rightarrow T_2)$ (intermediate) and $\omega > \omega(T_1 \rightarrow T_2)$ (larger), we have chosen ω and ϵ_0 such that both $\langle P_T \rangle$ and $\langle P_{ES} \rangle$ are minimized. This gives $\hbar\omega = 0.87$ eV, $I_0 = 7.9$ TW/cm^2; $\hbar\omega = 1.41$ eV, $I_0 = 5.0$ TW/cm^2; and $\hbar\omega = 1.63$ eV, $I_0 = 10.1$ TW/cm^2, respectively.

In general, for all frequencies below the $S_1 \rightarrow S_2$ resonance at the pulse intensities needed for spin locking, there is always substantial absorption to excited singlet states. For the best results shown in Fig. 5, the population in the singlets is mostly steady at 0.9 (although there is some slow decay into the triplets) but the population is widely spread

between S_1, S_2 and S_3. The ladder climbing excitation mechanism is very efficient in the ESM model.

As discussed in Sect. 3, the NRDSE mechanism is more efficient using intermediate frequencies. Thus, spin locking can be achieved with less intense pulses. To minimize absorption to S_3, the frequency must be smaller than 1.42 eV. However, 3-photon excitation to S_5 cannot be avoided. In the best results shown in Fig. 5, 90 % of the population is kept on the singlet states, but there is a clear beating between population in S_1 and S_5. Within the constraints of the LSS model, best results are obtained for larger frequencies. Here, part of the population goes to S_6.

In Fig. 6, we show the evolution of the nuclear probability density of all the singlet states for smaller and intermediate frequencies (the results for larger frequencies are very similar to those for intermediate frequencies). In the first case, the nuclear wave packet spreads and becomes mostly delocalized (although some coherent vibrational motion can be observed). This is mainly because the geometry of S_3 (with a single minima) is very different from that of S_1 and S_2. Surprisingly, very little singlet–triplet transfer is observed in the population of the different excited singlet states. When the laser is turned off only a small fraction of the population undergoes singlet–triplet conversion.
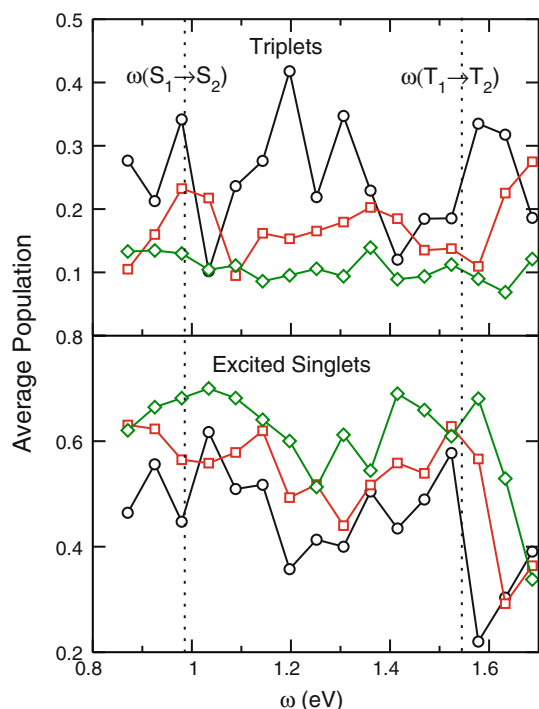
Fig. 4 Time-averaged population in all the triplet and all excited singlet states as a function of pulse frequency and intensity, for peak intensities (in TW/cm^2) of 5.0 (*circles*), 7.9 (*squares*) and 10.1 (*diamonds*). The vertical lines represent the frequencies at which the $S_1 \rightarrow S_2$ and $T_1 \rightarrow T_2$ Franck–Condon transitions are resonant at the equilibrium geometry
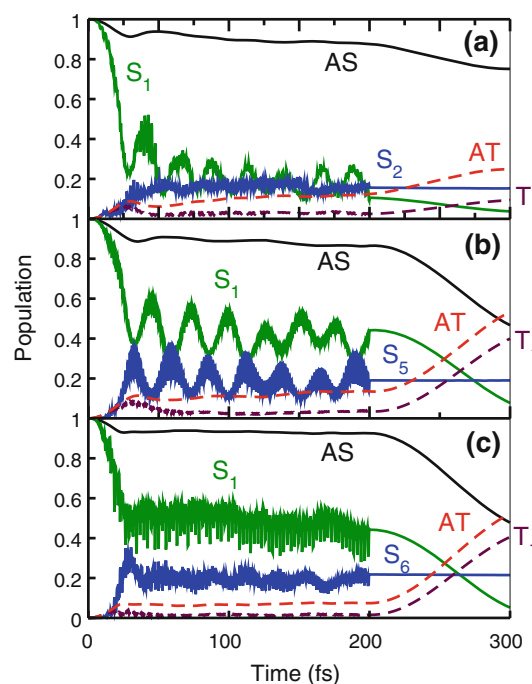
Fig. 5 Electronic states population dynamics in the LSS model with best pulse parameters in the three frequencies ranges. **a** with $\hbar\omega = 0.87$ eV $< \hbar\omega(S_1 \rightarrow S_2)$, **b** with $\hbar\omega(S_1 \rightarrow S_2) \leq \hbar\omega = 1.41$ eV $\leq \hbar\omega(T_1 \rightarrow T_2)$ and **c** with $\hbar\omega = 1.63$ eV $> \hbar\omega(T_1 \rightarrow T_2)$. Other parameters are given in the text. *AS* stands for the population of all singlet states and *AT* for the population of all the triplet states

For intermediate and larger frequencies, the nuclear wave-packet dynamics is still mostly confined to the bottom of the potential energy in V_1^S and much singlet–triplet conversion is observed when the laser is turned off.

5 Spin-state switching dynamics

As discussed in Sect. 2, almost all electronic singlet states practically overlap with the triplet states, so that the spin-transfer dynamics is very efficient. For the lowest energy potentials, only in S_2, there is no singlet–triplet transition. Following the NRDSE recipe (Sect. 3), in order to transfer the population from S_2 to T_2, one first needs to force a near-degeneracy of the two electronic states by Stark shifting the states using a strong nonresonant pulse. The electronic structure of the singlets around S_2 is quite different from that of the triplets around T_2 (S_2 is closer to S_1 while T_2 is closer to T_3), so that in principle the dynamic polarizabilities can be quite different and the NRDSE can be applied. However, the energy difference between the states at the equilibrium configuration of S_2 is approximately 1 eV, while the transition electronic dipoles are relatively small, making it impossible to compensate the energy difference by Stark-shift using any reasonably strong laser pulse. Thus, the singlet–triplet transfer cannot be indirectly induced by a single control pulse.

Instead of applying the NRDSE, in this work, we design an alternative dump–pump strategy that requires two pulses controlling the time-delay between them. The idea is sketched in Fig. 7. First, we apply a dump pulse that moves the population from S_2 to S_1. Then, we wait for the efficient singlet–triplet transfer between S_1 and T_1. Finally, a pump pulse is applied to transfer the population from T_1 to T_2. In the following, we present the results obtained within the MSS model.
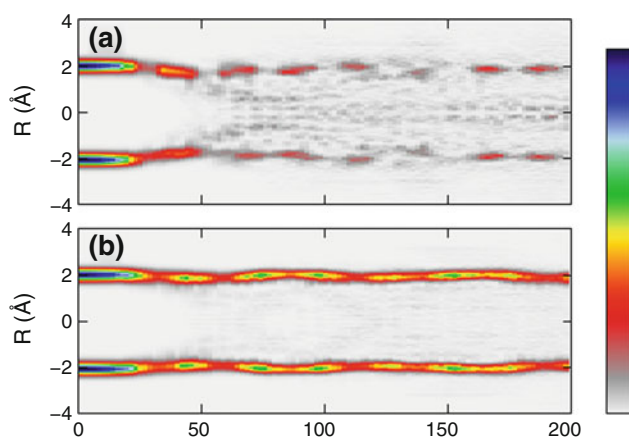
A typical result with approximately optimized pulse parameters is shown in Fig. 8. Here, we have chosen a negatively first chirped pulse with intensity $I_0 = 5.9$ TW/cm^2, duration $\tau_l = 90$ fs and frequency $\omega(t) = \omega_0 + \beta (t - t_0)/2$, where $\hbar\omega_0 = 1.15$ eV, $\hbar\beta = -0.0136$ eV/fs and t_0 is the center of the pulse. The transformed-limited second pulse has $I_0 = 0.9$ TW/cm^2, $\tau_l = 40$ fs and $\hbar\omega_0 = 1.52$ eV. The parameters of the first pulse are optimized to achieve maximal population transfer (approx 90 %) in a reasonably short time. Since the equilibrium geometry of S_2 and S_1 is very similar, it is difficult to disentangle the optical Rabi flopping between S_2 and S_1 from the spin Rabi flopping between S_1 and T_1. To avoid the $S_1 \rightarrow S_2$ back-transition, one needs to use chirped pulses. We chose a negative chirp, but it is possible to use positively chirped pulses as well. Since the $S_2 \rightarrow S_1$ Franck–Condon transition is at lower energies than the $S_2 \rightarrow S_3$ transition, at first look the choice of a positive chirp would seem a better option to minimize absorption into highly excited singlet states. For the MSS model, however, the results are slightly better for the negatively chirped pulses.

On the other hand, the equilibrium geometry of T_1 is very different from that of T_2. The time-delay between the pulses, $\tau_d = 110$ fs, is chosen to facilitate maximal singlet–triplet conversion between S_1 and T_1. For the second pulse, one can use a transformed-limited pulse because the wave packet in T_2 moves away from the Franck–Condon window and naturally deactivates the probability of stimulated emission. The parameters of the second pulse are chosen to maximize the yield of the $T_1 \rightarrow T_2$ transition in a very short time.

In Fig. 9, we show a contour plot of the time evolution of the nuclear probability densities of singlet and triplet spin character, separately. During the $S_2 \rightarrow S_1$ stimulated

Fig. 6 Nuclear probability density for all the singlet states in the LSS model with **a** $\hbar\omega = 0.87$ eV and $I_0 = 7.9$ TW/cm^2 and **b** $\hbar\omega = 1.41$ eV and $I_0 = 5.0$ TW/cm^2
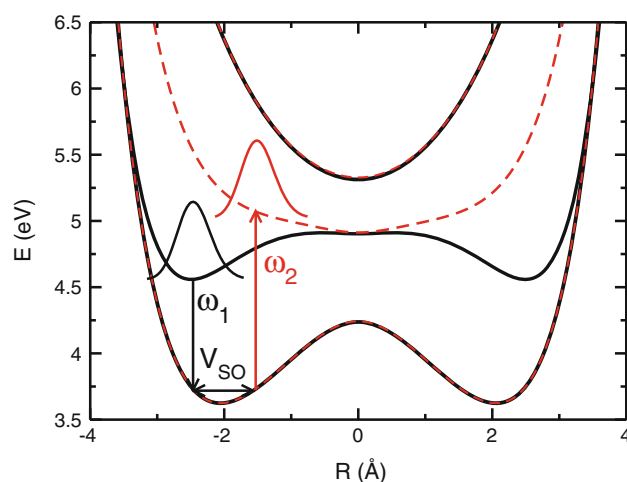

Fig. 7 Scheme of the dump–pump strategy used to induce spin switching between S_2 and T_2. We show the potential energy *curves* and the wave packets after the different stages of the process
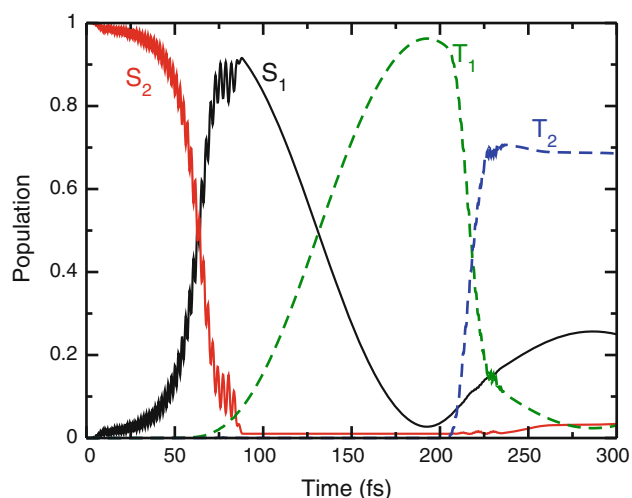
Fig. 8 Electronic states population dynamics in the MSS model applying the dump–pump scheme to induce singlet–triplet transfer from S_2 to T_2. The pulse parameters are given in the text

emission and the $S_1 \rightarrow T_1$ spin transfer, the singlet nuclear wave packet remains basically in the same region (around $R = \pm 2.5$ Å). The triplet nuclear wave packet appears at later states and moves toward $R = 0$, where the minima of T_2 is located.

Unfortunately, it is not possible to achieve similar results with the LSS model, regardless of the sign of chirp. The main problem is that in the Franck–Condon region the transition dipole between S_1 and S_2 is much smaller (a factor of 100 times smaller) than the transition dipole between S_2 and S_3 or between S_1 and S_5 or S_6. As a consequence, the dump pulse must be very strong and the transition competes with two- and three-photon absorption to other excited states, similarly at what happened when using low frequencies to lock the population in S_1. For the optimal parameters used above, a population of only 0.06
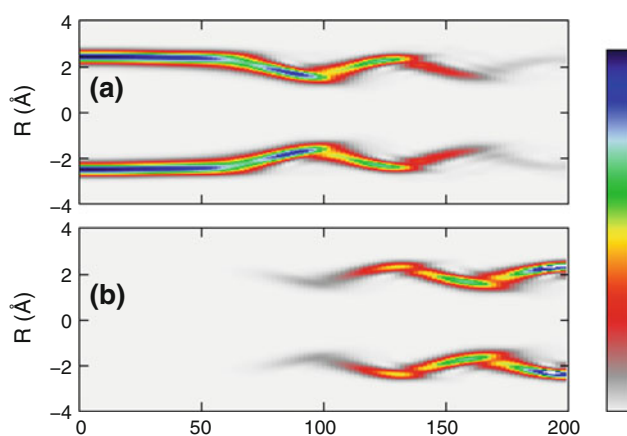

Fig. 9 Nuclear probability density in the spin-switching process for **a** all the singlet and **b** all the triplet states. The laser parameters are given in the text

arrives to T_2 at the end of the process, while singlet–triplet transitions between several excited states account for a total 22 % of population in the triplet states.

6 Summary and conclusions

In this work, we analyzed in detail the role of different laser parameters in the control of the singlet–triplet transition by means of the NRDSE. Following previous work [49], we used the ESM model to obtain the electronic potentials, dipole couplings and singlet–triplet couplings. Depending on the number of potentials included in the calculation, we defined two approximate models: the minimal MSS model and the larger LSS model. Multiphoton ionization and internal conversion were not taken into account. The time-dependent Schrödinger equation for the nuclear motion was then solved for these models starting from different electronic states in order to force spin locking when the laser-free dynamics implied full singlet-triplet conversion, or spin switching, when the laser-free dynamics conserved the spin state.

Certainly, the 3-D 5-particle collinear ESM Hamiltonian implies strong restrictions on the motion of the electrons and ion which show in the potentials and couplings. In particular, the singlet and triplet electronic states tend to have very similar energies, facilitating fast ISC processes. On the other hand, the model favors multiphoton ladder-type ionization, making strong-field control rather difficult. Still, we believe that the ESM is sufficiently flexible to allow the interplay of very different processes which can be analyzed to great detail, making the model an excellent numerical "laboratory" to test different control scenarios.

In this work, we focused on finding appropriate laser strategies and tuning the laser parameters in conditions where there is strong competence between different nonlinear processes, albeit disregarding ionization. Although appropriate laser parameters for spin switching were difficult to find even in the MSS model and were not found in the LSS model, spin locking was shown to be possible for different frequency windows. The smallness of the dipole couplings in the Franck–Condon region required the use of very strong pulses, which drove population to excited singlet states and thus opened new routes for spin transitions from the excited singlet to excited triplet states. Forcing a balance between two goals, minimizing the population in the triplet states and the population in the excited singlet states, was found to be a good strategy to identify control pulse frequencies and intensities that would maximally decouple the initial singlet state from the triplet states while at the same time they minimize the disturbance on the system.

Although previous work [49] showed that multi-photon ionization is the fastest process in the ESM model under the pulse intensities used in this work, the results also showed that ionization occurred mainly in the Franck–Condon window and did not affect the spin dynamics. In particular, the NRDSE was shown to maintain the control on the spin populations of the remaining (nonionized) parts of the system. Given that the full vibronic dynamics requires considerable computation time, the strategy used in this work also paves the way to identifying the best control scenarios when ionization is taken into account. We expect that this control strategy will be effective when the required laser intensity is smaller, for instance in weaker spin-coupling conditions.

References

1. Rice SA, Zhao M (2000) Optical control of molecular dynamics. Wiley-Interscience, New York
2. Brixner T, Pfeifer T, Gerber G, Wollenhaupt M, Baumert T (2005) Chap 9: Optimal control of atomic, molecular and electron dynamics with tailored femtosecond laser pulses. In: Hannaford P (ed) Femtosecond laser spectroscopy. Springer, New York, pp 225–266
3. Brumer P, Shapiro M (2012) Quantum control of molecular processes. Wiley-VCH, New York
4. Silberberg Y (2009) Quantum coherent control for nonlinear spectroscopy and microscopy. Annu Rev Phys Chem 60:277–292
5. Geremia GM, Zhu W, Rabitz H (2000) Incorporating physical implementation concerns into closed loop quantum control experiments. J Chem Phys 113:10841–10848
6. Brixner T, Kiefer B, Gerber G (2001) Problem complexity in femtosecond quantum control. Chem Phys 267:241–246
7. Bartels R, Backus S, Zeek E, Misoguti L, Vdovin G, Christov IP, Murnane MM, Kapteyn HC (2000) Shaped-pulse optimization of coherent emission of high-harmonic soft X-rays. Nature 406:164–166
8. Levis RJ, Menkir GM, Rabitz H (2001) Selective bond dissociation and rearrangement with optimally tailored strong-field laser pulses. Science 292:709–713
9. Yuan JM, George TF (1978) Semiclassical theory of unimolecular dissociation induced by a laser field. J Chem Phys 68:3040–3052
10. Bandrauk AD, Sink ML (1981) Photodissociation in intense laser fields: predissociation analogy. J Chem Phys 74:1110–1117
11. Allendorf SW, Szoke A (1991) High-intensity multiphoton ionization of H_2. Phys Rev A 44:518–534
12. Guisti-Suzor A, Mies FH, DiMauro LF, Charron E, Yang B (1995) Dynamics of H_2^+ in intense laser fields. J Phys B 28:309–339
13. Frasinski LJ, Posthumus JH, Plumridge J, Codling K, Taday PF, Langley AJ (1999) Manipulation of bond hardening in H_2^+ by chirping of intense femtosecond laser pulses. Phys Rev Lett 83:3625–3628
14. Sussman BJ, Ivanov MY, Stolow A (2005) Nonperturbative quantum control via the nonresonant dynamic Stark effect. Phys Rev A 71:051401(R)
15. Sussman BJ, Townsend D, Ivanov MY, Stolow A (2006) Dynamic Stark control of photochemical processes. Science 314:278–281
16. González-Vázquez J, Sola IR, Santamaria J, Malinovsky VS (2006) Quantum control of spinorbit coupling by dynamic Stark-shifts induced by laser fields. Chem Phys Lett 431:231–235
17. González-Vázquez J, Sola IR, Santamaria J, Malinovsky VS (2006) Optical control of the singlet–triplet transition in Rb2. J Chem Phys 125:124315
18. González-Vázquez J, Sola IR, Santamaria J, Malinovsky VS (2007) Vibrationally state-selective spinorbit transfer with strong nonresonant pulses. J Phys Chem A 111:2670–2678
19. Friedrich B, Herschbach D (1995) alignment and trapping of molecules in intense laser fields. Phys Rev Lett 74:4623–4626
20. Seideman T (1995) Rotational excitation and molecular alignment in intense laser fields. J Chem Phys 103:7887–7896
21. Ortigoso J, Rodriguez M, Gupta M, Friedrich B (1999) Time evolution of pendular states created by the interaction of molecular polarizability with a pulsed nonresonant laser field. J Chem Phys 110:3870–3875
22. Stapelfeldt H, Seideman T (2003) Colloquium: aligning molecules with strong laser pulses. Rev Mod Phys 75:543–557
23. Leibscher M, Averbukh IS, Rabitz H (2003) Molecular alignment by trains of short laser pulses. Phys Rev Lett 90:2130011–2130014
24. Suzuki T, Sugawara Y, Minemoto S, Sakai H (2008) Optimal control of nonadiabatic alignment of rotationally cold N_2 molecules with the feedback of degree of alignment. Phys Rev Lett 100:033603
25. Bisgaard CZ, Clarkin OJ, Wu GR, Lee AMD, Gessner O, Hayden CC, Stolow A (2009) Time-resolved molecular frame dynamics of fixed-in-space CS2 molecules. Science 323:1464–1468
26. Sussman BJ, Underwood JG, Lausten R, Ivanov MY, Stolow A (2006) Quantum control via the dynamic Stark effect: application to switched rotational wave packets and molecular axis alignment. Phys Rev A 73:053403
27. Chang BY, Choi H, Shin S, Lee S, Sola IR (2009) Ultrafast photodissociation assisted by strong nonresonant Stark effect: the straddling control pulse. J Mod Opt 56:811–821
28. Chang BY, Shin S, Santamaria J, Sola IR (2009) Bond breaking in light-induced potentials. J Chem Phys 130:124320
29. Chang BY, Shin S, Sola IR (2009) Further aspects on the control of photodissociation in light-induced potentials. J Chem Phys 131:204314
30. McQuarrie DA, Simon JD (1997) Physical chemistry. University Science Books, Sausalito
31. Hare PM, Crespo-Hernández CE, Kohler B (2007) Internal conversion to the electronic ground state occurs via two distinct pathways for pyrimidine bases in aqueous solution. Proc Natl Acad Sci USA 104:435–440
32. Kwo WM, Ma C, Phillips DL (2008) A doorway state leads to photostability or triplet photodamage in thymine DNA. J Am Chem Soc 130:5131–5139
33. Etinski M, Fleig T, Marian CM (2009) Intersystem crossing and characterization of dark states in the pyrimidine nucleobases uracil, thymine, and 1-methylthymine. J Phys Chem A 113:11809–11816
34. Parker DSN, Minns RS, Penfold TJ, Worth GA, Fielding HH (2009) Ultrafast dynamics of the S1 excited state of benzene. Chem Phys Lett 469:43

35. Richter M, Marquetand P, González-Vázquez J, Sola I, González L (2012) Femtosecond intersystem crossing in the DNA nucleobase cytosine. J Phys Chem Lett 3:3090
36. Gómez-Abal R, Hübner W (2002) Simple model for laser-induced electron dynamics. Phys Rev B 65:195114
37. Gómez-Abal R, Hübner W (2003) The role of spinorbit coupling in optically induced ultrafast magnetic dynamics. J Phys Condens Mater 15:S709–S722
38. Gómez-Abal R, Ney O, Satitkovitchai K, Hübner W (2004) All-optical subpicosecond magnetic switching in NiO(001). Phys Rev Lett 92:227402
39. Satitkovitchai K, Pavlyukh Y, Hübner W (2005) Ab initio study of spin–orbit coupling effects on the low-lying excited states of NiO. Phys Rev B 72:045116
40. Zutic I, Das Sarma S (2004) Spintronics: fundamentals and applications. Rev Mod Phys 76:323–410
41. Bouwmeester D, Ekert A, Zeilinger A (2000) The physics of quantum information. Springer, New York
42. Shin S, Metiu H (1995) Nonadiabatic effects on the charge transfer rate constant: a numerical study of a simple model system. J Chem Phys 102:9285–9295
43. Shin S, Metiu H (1996) Multiple time scale quantum wavepacket propagation:electro nnuclear dynamics. J Phys Chem 100:7867–7872
44. Erdmann M, Marquetand P, Engel V (2003) Combined electronic and nuclear dynamics in a simple model system. J Chem Phys 119:672–679
45. Erdmann M, Engel V (2004) Combined electronic and nuclear dynamics in a simple model system II spectroscopic transitions. J Chem Phys 120:158
46. Erdmann M, Baumann S, Gräfe S, Engel V (2004) Electronic predissociation: a model study. Eur Phys J D 30:327–333
47. Gräfe S, Engel V (2006) Local control theory applied to coupled electronic and nuclear motion. Chem Phys 329:118
48. Erdmann M, Gross EKU, Engel V (2004) Time-dependent electron localization functions for coupled nuclear-electronic motion. J Chem Phys 121:9666–9670
49. Falge M, Engel V, Lein M, Vindel-Zandbergen P, Chang BY, Sola IR (2012) Quantum wave-packet dynamics in spin-coupled vibronic states. J Phys Chem A 116:11427. doi:10.1021/jp306566x
50. Sola IR, González-Vázquez J, Malinovsky VS (2006) Optical control of a spin switch in the weak spin–orbit coupling limit. Phys Rev A 74:043418
51. Sussman BJ (2011) Five ways to the nonresonant dynamic Stark effect. Am J Phys 79:477–484

Exohedral interaction in cationic lithium metallofullerenes

Maitreyi Robledo · Fernando Martín ·
Manuel Alcamí · Sergio Díaz-Tendero

Abstract We present a density functional theory study on the exohedral interaction between a singly positively charged lithium atom and fullerenes: Li^+-C_n, $n = 70, 60$ and 50. We have found that the interaction is of electrostatic nature: the cation polarizes the π electronic cloud of the fullerene in an ion-induced dipole attraction. We show that these systems present a shallow potential energy surface where the local minima correspond to the interaction of the cation on top of one pentagonal or one hexagonal face of the fullerene and transition states connect them through a movement of the cation over a C−C bond. The type of interaction and the shape of the potential energy surface give rise to the so-called planetary systems, where the alkali cation is revolving around the carbon cage in orbits. The studied systems present several pathways that are more likely than others to behave as potentially favorable orbits.

M. Robledo · F. Martín · M. Alcamí · S. Díaz-Tendero (✉)
Departamento de Química, Módulo 13, Universidad Autónoma de Madrid, 28049 Madrid, Spain
e-mail: sergio.diaztendero@uam.es

F. Martín
Instituto Madrileño de Estudios Avanzados en Nanociencias (IMDEA-Nanociencia) Cantoblanco, 28049 Madrid, Spain

1 Introduction

The discovery of the C_{60} fullerene [1] meant the landmark of several works devoted to develop possible applications of these structures in the field of nanotechnology (see e.g. [2]). In the particular case of interaction with metal atoms, it has been shown that $C_{60}M_x$ species present superconductivity properties [3, 4] and it is believed that metal-doped carbon structures can have promising applications as hydrogen storage nanodevices [5]. Thus, the understanding of the interaction of metal atoms with fullerenes is of prime importance. In this work, we evaluate the exohedral interaction between Li^+ and different fullerene structures. Our aim is to establish which are the most stable positions and to have a complete overview of the potential energy surface of an alkaline cation approaching a fullerene.

Beyond the most studied fullerene, C_{60}, there are other carbon cages that imply stable structures. Indeed, mass-spectrometric studies based on carbon ions produced in hydrogen–carbon flames detected that the most prominent peaks result at $m/q = 50$, 60, and 70 [6], showing that these are the most favored sizes to be formed. C_{70} and C_{50}, together with a large amount of C_{60}, were also observed in laser-induced evaporation from graphite in the pioneer studies of Kroto et al. [1]. We have therefore considered in our work the most stable isomers for these three fullerenes. The most stable isomer of C_{60} and C_{70} follows the so-called isolated pentagon rule (IPR) [2, 7], which states that none of the 12 pentagonal faces touches each other, hence minimizing the strain introduced by adjacent pentagons (AP). For C_{70} and C_{60} fullerenes, only one isomer follows the IPR. The IPR isomer of C_{60} has icosahedral symmetry (I_h), and in the case of C_{70}, the IPR isomer has D_{5h} symmetry. In fullerenes smaller than C_{60}, this rule cannot be applied, and the stability of the different isomers is given

by the pentagon adjacency penalty rule (PAPR) [7–13], which predicts that, among all the possible non-IPR isomers, the most stable one is that with the minimum number of APs. In the case of the C_{50}, only one isomer (with D_{5h} symmetry) presents 5APs; the rest of the isomers have 6APs or more. Nevertheless, it has been found that one of the C_{50} isomers with 6APs (and D_3 symmetry) is degenerated in energy with the D_{5h} isomer with 5APs [14–16]. This unusual situation was explained in terms of what it was called spherical aromaticity [17]. The C_{50} fullerene has the exact number of π electrons to form an electronic closed shell $(2l + 1)^2$; the isomer with D_3 symmetry and 6APs gains in stability because it presents the closer structure to a sphere. This additional stabilization energy is enough to compensate the fact of showing one more pentagon–pentagon adjacency in its structure [16].

Among the previous studies on metal-doped fullerenes, in the pioneering work of Martin et al. [18], the synthesis of alkali metal doped C_{60} was achieved for the first time producing carbon–metal complexes of the form $(C_{60})_n M_x$ ($n = 1, 2, 3; 1 \leq x < 150$; M = Li, Na and K). Some theoretical studies have focussed on the comprehension of C_{60} fullerenes coated with lithium, the metal of interest in this work. One important result of these studies is that in LiC_{60} [19] and LiC_{60}^+ [20] the isomers with the metal attached on the hexagonal and pentagonal faces are nearly degenerate. The dynamics of a Li^+ cation on a C_{60} surface have been also investigated following a Molecular Orbital–Molecular Dynamics (MO–MD) method [21, 22]. Differences in the stability between the hexagonal and the pentagonal isomers depending on the density functional theory (DFT) method employed were found. The diffusion of the Li^+ cation on a C_{60} was also addressed performing quasiclassical trajectories [23]. In both dynamical studies, a ring-hopping behavior due to the low-energy barriers between adjacent rings was obtained. Theoretical studies have been also focused on evaluating the electronic properties of $Na_n C_{60}$ and $Li_n C_{60}$ ($n \leq 12$) [19]. Semiempirical calculations were employed to evaluate the interaction between Li^+ and C_{60}, C_{60}^-, and C_{60}^{2-} [24] and the characteristics of $Li_n^{n+} C_{60}$ endohedral and exohedral complexes ($n = 1, 2, 3, 4$) [25]. Furthermore, the structure of $Li_{12}C_{60}$ [5] and $Li_6 C_{60}$ [26] and their potential use as molecules storing hydrogen have been also addressed. In contrast to all the studies aimed at evaluating the properties of metal doped C_{60} compounds, much less works have been reported for metal-coated species involving other fullerene sizes. Studies concerning C_{70} have been limited to (1) alkali metal intercalation compounds formed by single crystals of electric-field-doped fullerenes (KC_{70} and $K_4 C_{70}$), which present superconductivity properties [27]; (2) other intercalation alkaline-earth and rare-earth metal doped C_{70} compounds [$Ba_x C_{70}$ and $Sr_x C_{70}$ ($x = 3, 4, 6, 9$), and $Eu_y C_{70}$ ($y = 3,9$)] studied by X-ray measurements and yielding

stable phases [28]; and (3) the geometry and electronic structure of $C_{70} Na_n^-$ ($n = 0$–15) clusters studied with photoelectron spectroscopy [29]. To the best of our knowledge, none previous works have been reported on metal coated C_{50}. However, this fullerene has been synthesized in the form of $C_{50}Cl_{10}$, that is, saturated with halogen atoms and giving rise to the first isolated fullerene smaller than C_{60} [30]. In the obtained compound, each of the 10 chlorine atoms forms covalent bonds with carbon atoms constituting the central belt of the $C_{50}(D_{5h})$ isomer.

On the other hand, the interaction of Li^+ ions with clusters of phosphorus (P_n) was found to give rise to stable species [31]. Indeed, the term planetary system was first proposed for the $[LiP_4]^+$ complex, where the lithium cation is revolving around the highly symmetric neutral cluster and present thermally accessible paths ("orbits") [31]. The ferrocene-lithium cation complex was also shown to present a planetary system in one of its isomers [32, 33].

In this work, we present a theoretical study of the exohedral interaction of Li^+ with $C_{60}(I_h)$, C_{70} (D_{5h}), C_{50} (D_3), and C_{50} (D_{5h}) fullerenes. Our aim is not only to establish which are the most stable structures, but to make a complete exploration of the potential energy surface (PES) in order to determine which is the energy required for Li^+ to move among the different fullerene sites. The stability and high symmetry of the studied fullerenes made them promising structures for finding planetary systems when Li^+ interacts with their external surface: exohedral metallofullerenes. We thus evaluate the most favorable paths, that is, the thermally accessible paths leading to orbits around the carbon cage. This may be of interest for potential applications of these compounds in nanotechnology. For instance, regarding the H_2 storage capacity, the studies up to now have considered $C_{60}Li_6$ [26] or $C_{60} Li_{12}$ [5] as stable molecules to which up to 60 H_2 attach. A complete study of the PES, as the one presented here, will shed some light to know whether this static view should be changed by a more dynamical view of the problem in which Li atoms and H_2 move around the fullerene cage.

2 Theoretical methods

Geometries and interaction energies of all exohedral metallofullerenes have been computed in the framework of the DFT. In particular, we have employed the B3LYP functional, which combines the three parameters' non-local hybrid exchange potential proposed by Becke—B3 [34] with the non-local gradient corrected correlation functional of Lee, Yang, and Parr—LYP [35]. We have used the 6-31G(d) Pople basis set, with polarization functions in the lithium and carbon atoms. This method and basis set has shown to give accurate results, comparable to those

obtained with perturbational theory (MP2) in fullerenes [36]. We have also evaluated the harmonic vibrational frequencies at the same B3LYP/6-31G(d) level of theory to confirm that the calculated geometry effectively corresponds to a minimum or a saddle point (first order transition state—TS) in the potential energy surface (PES) and to compute the zero-point energy (ZPE) correction. All

calculations were carried out with the Gaussian 09 program package [37].

Adsorption energies E_{ad} are given by:

$$E_{ad} = E(C_n) + E(Li^+) - E([LiC_n]^+) \quad (1)$$

where $E(C_n), E(Li^+)$ and $E([LiC_n]^+)$ are the absolute electronic energies after optimization (including the ZPE

Fig. 1 Structure of the minima obtained with the cation attached to one hexagonal (*h*) or pentagonal (*p*) face for each $[LiC_n]^+$ ($n = 50, 60,$ and 70) complex
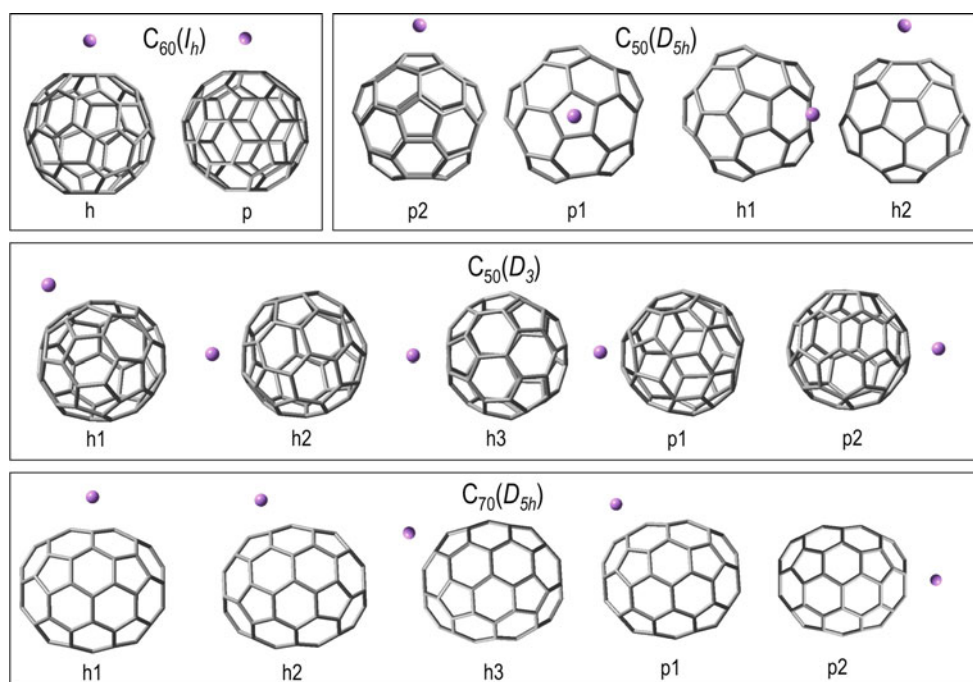
Fig. 2 Schlegel projection of the $[LiC_n]^+$ complexes with the stationary points found as follows: *circles* correspond to minima and *triangles* to transition states
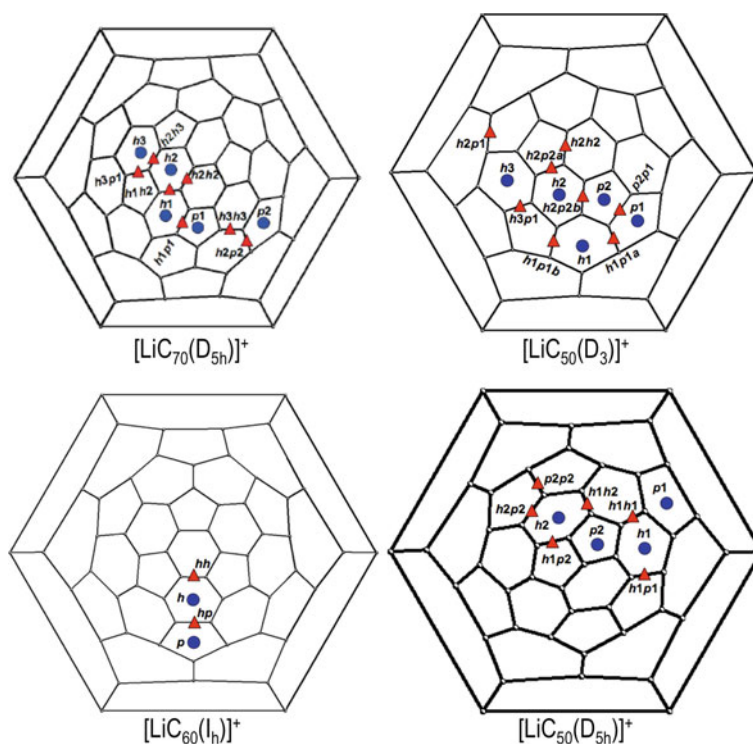
correction) of the fullerene, the cation and the complex, respectively.

We have first optimized the geometry of each neutral isolated fullerene. Starting with the obtained structures, we have then computed the whole system (i.e. $[LiC_n]^+$; $n = 50, 60$ and 70). To this, we have considered several initial guess geometries with the metal atom interacting at different positions of the external surface of the carbon cage. All the geometry optimizations were carried out without any symmetry constraints.

3 Results and discussion

We have explored the potential energy surface (PES) of the exohedral interaction of Li^+ with $C_{60}(I_h)$, $C_{70}(D_{5h})$, $C_{50}(D_3)$, and $C_{50}(D_{5h})$. For each fullerene, we have found several stationary points as shown in Figs. 1 and 2. For all the cases studied, the minima (Fig. 1) correspond to the Li^+ cation placed on top of a pentagonal or hexagonal face. In the transition states connecting them, the Li^+ cation is on the C–C bond shared by both faces (Fig. 2). We have

Fig. 3 Molecular electrostatic potential maps for $C_{50}(D_3)$, $C_{50}(D_{5h})$ and $C_{70}(D_{5h})$. The *numbers* on the surfaces represent their potential countour isovalues (in a.u.) in the selected plane
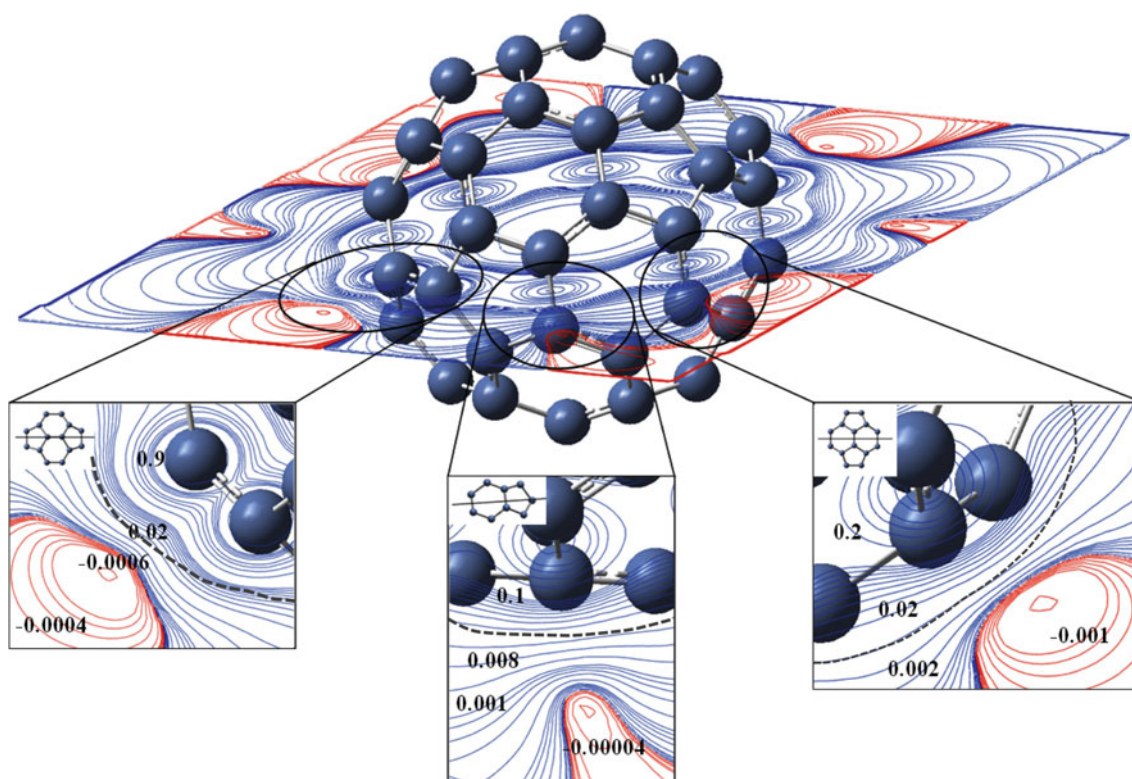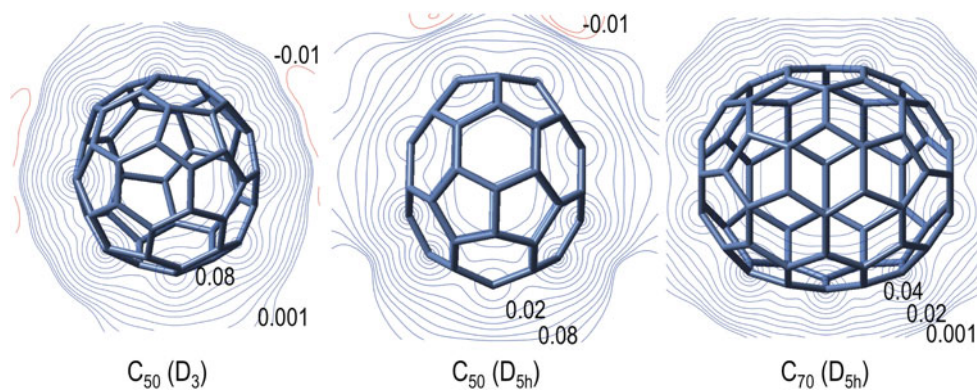
Fig. 4 Molecular electrostatic potential map for the $C_{60}(I_h)$ fullerene, indicating three cutoffs of different parts of the plane. The path presumably followed by the Li^+ cation (position of the minima and transition states after geometry optimization) is indicated with a *gray dashed line*

disregarded those positions in which the Li^+ is placed on top of a carbon atom, because in all the tries we have considered, either second-order transition states were found or they fell into the other stationary points of interest already mentioned (minima or first order TS). In the most symmetric studied structure, $C_{60}(I_h)$, all the pentagons and hexagons are equivalent and, thus, we have found four stationary points: The cation placed on the hexagonal and pentagonal faces (h and p, respectively), and in the C–C bonds shared between two hexagons (hh) and between one hexagon and one pentagon (hp). On the other hand, in the other three studied fullerenes, $C_{50}(D_3)$, $C_{50}(D_{5h})$, and $C_{70}(D_{5h})$, due to their lower symmetry with respect to the $C_{60}(I_h)$, the carbon cage presents a larger number of non-equivalent positions where Li^+ may be placed: h_i and p_j for the different hexagonal and pentagonal faces and, $h_ih_j$, p_ip_j and h_ip_j for the different C–C bonds connecting them. For instance, $C_{50}(D_{5h})$ shows two different pentagonal faces: $p1$ is an isolated pentagonal ring surrounded by five hexagons: and $p2$ is connected to other $p2$ rings giving rise to five APs in the structure; and two non-equivalent hexagonal faces: $h1$ represent the hexagonal rings surrounding $p1$ and $h2$ is next to the APs formed by $p2$. The bonds shared between these faces give rise to 6 different locations for the cation: $h1h2$, $h1p1$, $h1p2$, $h2p2$, $p2p2$. For the $C_{50}(D_3)$ /$C_{70}(D_{5h})$ fullerenes, we have found five minima and eleven / eight transition states, respectively.

We first analyze the kind of interaction between the Li^+ and the fullerenes. For this, we present in Fig. 3 the molecular electrostatic potential (MEP) maps for $C_{70}(D_{5h})$, $C_{50}(D_3)$, and $C_{50}(D_{5h})$ in a plane that divides the fullerene cage in two halves. A detailed MEP for $C_{60}(I_h)$ is shown in Fig. 4. The contours represent the interaction energy for a positive test charge located at a certain point and it includes the interaction with both the electron density cloud and the nuclei with a non-polarized fullerene. We observe positive values of the potential contours close to the fullerene, and low negative values far from it. These values imply that a positive charge would be repelled when located close to the carbon cage and attracted far from the fullerene surface. In our case, the interaction due to the electrostatic potential is repulsive at the carbon cage—Li distances obtained in the geometry optimization ~ 2.5 Å (approximate path denoted by dashed lines in Fig. 4). Nevertheless, since the Li^+ is located close to the fullerene surface, it strongly polarizes the π electronic cloud, and therefore, the ion-polarizability term becomes dominant against the pure electrostatic effect revealed by the MEP.

A deeper analysis on the type of interaction between the carbon cage and the metal cation has been performed on the most stable isomer of each system using the natural bonding orbital (NBO) approach [39–42] and the Bader's Quantum Theory of Atoms In Molecules— QTAIM

[43, 44]. With the former, we have computed the charge distribution on each atom. QTAIM theory provides us useful information on the basis of electron density properties: molecular graphs showing bond critical points (BCP), the Laplacian of the electron density, $\nabla^2(\rho)$, and total electronic energy density, $H(\rho)$. These magnitudes give complementary information: The contour maps of $\nabla^2(\rho)$ and $H(\rho)$ give a qualitative picture of the kind of interaction. In addition, a BCP where $\nabla^2(\rho) < 0$ is dominated by a local reduction in potential energy (i.e. covalent interaction), while a BCP where $\nabla^2(\rho) > 0$ is dominated by a local excess of kinetic energy (i.e. electrostatic interaction). Finally, with the sign of $H(\rho)$ in the BCP, we can unambiguously predict the character of the bond: $H(\rho) > 0$ ionic bond, $H(\rho) < 0$ covalent bond. The values of $H(\rho)$ at the BCP provide also a quantitative estimation of the interaction strength.

The results obtained for the NBO charges and the AIM molecular graphs with the contour maps of the total electronic energy density $H(\rho)$ are presented in Fig. 5. Our calculations show that the positive charge is mainly localized on the Li atom, as expected. However, we also observe that the carbon atoms close to the cation present a negative charge, therefore involving a local polarizability. If we look at the shape of the $H(\rho)$ maps, we observe how the electron density contributing to the energy appears as spherical contours around the Li^+ nuclei in the manner characteristic of two separated closed shell distributions. This behavior undoubtedly points to a purely electrostatic interaction between the carbon cage and the metal cation. We have obtained positive (and low) values for $\nabla^2(\rho)$ at the BCPs linking Li^+ and C_n confirming the electrostatic (and weak) nature of the interaction. The positive values of $H(\rho)$ observed in the BCPs are a further indication of the ionic interaction.

The results of our exploration of the potential energy surfaces (PES) of the $[LiC_n]^+$ species studied (energetic diagrams) are shown in Fig. 6:

- For $[LiC_{60}]^+$, Fig. 6b, we found degeneracy between the h and p geometries (the energy difference is 0.01 kcal/mol) in agreement with previous works which predicted the p isomer the most stable by only 0.057 eV [20]. The structure where the Li^+ is placed on the C–C bond shared by these two rings (hp) corresponds to a transition state that connects these two minima and is located at 3.1 kcal/mol above the global minimum. The bond shared by two hexagons, hh with double bond character, is the attachment point of the Li^+ in the transition states that connects two h minima and is found at 4.1 kcal/mol above the minimum. Thus, a lithium cation approaching the C_{60} fullerene, will attach to the hexagonal face with an adsorption energy of 38.50 kcal/mol and since the entrance channel is at

$H(\rho)$ NBO

$[LiC_{70}(D_{5h})]^+$ (h1)

$H(\rho) = 0.0035$
$\nabla^2(\rho) = 0.09$
$q_{NBO}(Li) = 0.903$
$q_{NBO}(C_{ring}) = [-0.099, -0.033]$

$[LiC_{60}(I_h)]^+$ (h)

$H(\rho) = 0.0029$
$\nabla^2(\rho) = 0.07$
$q_{NBO}(Li) = 0.903$
$q_{NBO}(C_{ring}) = [-0.074, -0.072]$

$[LiC_{50}(D_3)]^+$ (p1)

$H(\rho) = 0.0039$
$\nabla^2(\rho) = 0.08$
$q_{NBO}(Li) = 0.899$
$q_{NBO}(C_{ring}) = [-0.115, -0.091]$

$[LiC_{50}(D_{5h})]^+$ (p2)

$H(\rho) = 0.0039$
$\nabla^2(\rho) = 0.08$
$q_{NBO}(Li) = 0.907$
$q_{NBO}(C_{ring}) = [-0.110, -0.089]$
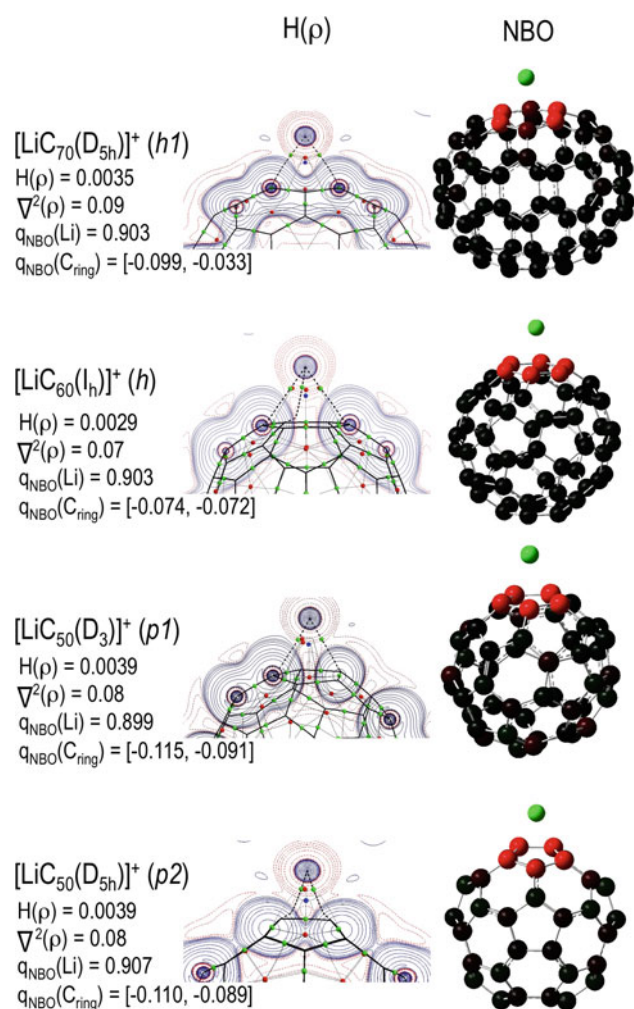
Fig. 5 *Central column* contour maps of the total electronic energy density $H(\rho)$ (using [38]). *Right column* NBO analysis of the most stable isomer for each $[LiC_n]^+$ studied complex: *Red* atoms denote negative charge, *green* atoms positive charge, and *black* atoms neutral. *Left column* Values of $\nabla^2(\rho)$ and $H(\rho)$ at the bond critical points formed between the Li and the carbon cage, the NBO charge on the Li atom, $q_{NBO}(Li)$ and on the carbon atoms of the ring involved in the attachment $q_{NBO}(C_{ring})$ (maximum and minimum)

much higher energy than the transition states, it would be able to move to a close minimum, the pentagonal, or hexagonal face. Thus, performing movements in orbits around the carbon cage.

- For $[LiC_{70}(D_{5h})]^+$, Fig. 6a, our results show that the alkali cation attaches to the carbon cage with an adsorption energy of 41.75 kcal/mol. The most stable geometry is the hexagonal face labeled as h1, which is surrounded by 4 hexagonal and 2 pentagonal rings. This minimum is connected to the pentagonal face p1, which is almost degenerated (0.4 kcal/mol), by a transition state (h1p1) located 4.6 kcal/mol above and with the hexagonal face h2 through h1h2 at 5.4 kcal/mol. Thus, these points involve possible Li^+
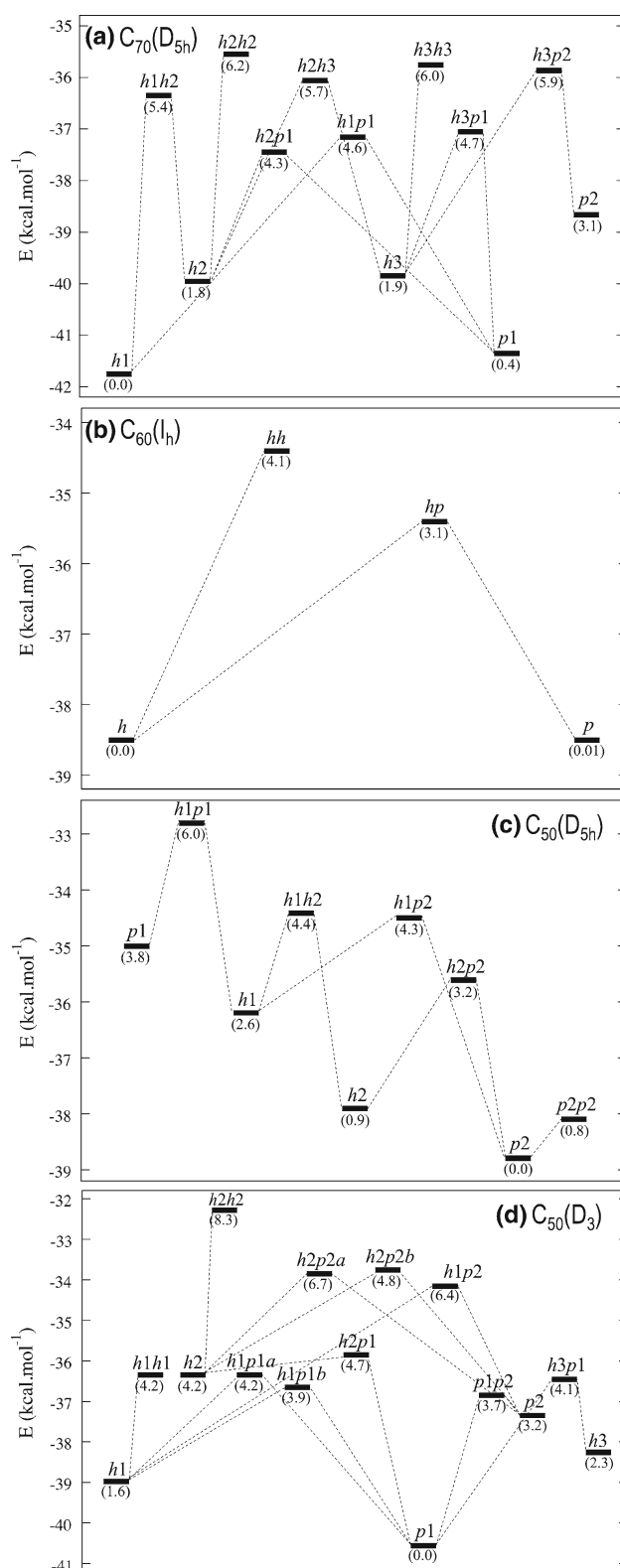
Fig. 6 Energetic diagrams showing the minima at the hexagonal or pentagonal faces and the transition states at the C–C bonds for **a** $C_{70}(D_{5h})$, **b** $C_{60}(I_h)$, **c** $C_{50}(D_{5h})$, and **d** $C_{50}(D_3)$. Energy in the axis is given with respect to the adsorption energy to the most stable minimum. In *brackets*, the relative energy (in kcal/mol) of each structure with respect to the most stable isomer in each system is given

movements in the nearby of the central rings of the $C_{70}(D_{5h})$ fullerene.

- For $[LiC_{50}(D_{5h})]^+$, Fig. 6c, our results show that the most stable isomer is the one labeled as $p2$ with and adsorption energy of 38.79 kcal/mol. It corresponds to the Li^+ attached to one of the pentagonal faces that form the 5APs. This pentagonal face is also surrounded by $h1$ and $h2$. $p1$ is the less stable isomer, corresponding to a pentagon surrounded by hexagons. The movement of Li^+ from $p2$ to the adjacent hexagon faces ($h1$ and $h2$) implies energy barriers of 3.2 and 4.4 kcal/mol, respectively. To reach structure, $p1$ implies however a larger energy barrier (6 kcal/mol). Thus, the most favorable orbits would imply the movement of the Li^+ cation to $h1$, $h2$, and $p2$ positions.

- For $[LiC_{50}(D_3)]^+$, Fig. 6d, we have found the most stable isomer when the Li^+ is attached at the pentagonal face labeled as $p1$, with an adsorption energy of 40.55 kcal/mol, followed by the hexagonal faces $h1$ and $h3$. The lowering in symmetry of this structure leads to a more complicated PES. In general, some transition states are found at relative energies of ~ 4 kcal/mol with respect to the global minima, and these paths are therefore more favorable. As in the previous cases, the cation will perform orbits in those paths.

As for the neutral isolated C_{50} isomers, in the case of $[LiC_{50}]^+$, the D_3 structure is more stable than the D_{5h} one. In particular, the most stable isomer of $[LiC_{50}(D_3)]^+$, $p1$, is 3.46 kcal/mol more stable than the most stable isomer of $[LiC_{50}(D_{5h})]^+$, $p2$.

4 Conclusions

The exohedral interaction of the most stable fullerenes (C_{60}, C_{70} and C_{50}) with Li^+ has been studied by using DFT-B3LYP. Our calculations have shown that in the cationic exohedral metallofullerenes, the bond between the cation and the external face of the fullerene has an almost pure electrostatic origin: charge-induced dipole interaction. The exploration of the potential energy surfaces has shown that several degenerated minima (within c.a. 4 kcal/mol) are possible, all of them correspond to the attachment of the cation on one of the faces of the fullerene (hexagonal or pentagonal). The transition states connecting them are structures where the Li^+ moves on the top of the $C-C$ bond shared by these faces with energy barriers of the order of 4–6 kcal/mol. All this implies that a Li^+ cation approaching a neutral fullerene will be attached to any of these minima, with an adsorption energy of ~ 40 kcal/mol, finding afterward an extremely shallow potential energy

surface in which it will be trapped. However, since some of the barriers between the minima are higher in energy, metal translations may be restricted to certain paths, expecting that the cation will not be able to move at will, but following orbits leading to a planetary system.

References

1. Kroto HW, Heath JR, O'Brien SC, Curl RF, Smalley RE (1985) Nature 318:162
2. Kadish K, Ruoff R (2000) Fullerenes: chemistry, physics, and technology. Wiley, New York
3. Haddon RC, Hebard AF, Rosseinsky MJ, Murphy DW, Duclos SJ, Lyons KB, Miller B, Rosamilia JM, Fleming RM, Kortan AR, Glarum SH, Makhija AV, Muller AJ, Eick RH, Zahurak SM, Tycko R, Dabbagh G, Thiel FA (1991) Nature 350:320
4. Rosseinsky MJ, Ramirez AP, Glarum SH, Murphy DW, Haddon RC, Hebard AF, Palstra TTM, Kortan AR, Zahurak SM, Makhija AV (1991) Phys Rev Lett 66:2830
5. Sun Q, Jena P, Wang Q, Marquez M (2006) J Am Chem Soc 128(30):9741
6. Gerhardt P, Lffler S, Homann K (1987) Chem Phys Lett 137(4):306
7. Kroto HW (1987) Nature 329:529
8. Fowler PW (1996) Contemp Phys 37:235
9. Fowler PW, Heine T, Mitchell D, Orlandi G, Schmidt R, Seifert G, Zerbetto F (1996) J Chem Soc Faraday Trans 92:2203
10. Schmalz TG, Seitz WA, Klein DJ, Hite GE (1988) J Am Chem Soc 110:1113
11. Albertazzi E, Domene C, Fowler PW, Heine T, Seifert G, Alsenow CV, Zerbetto F (1999) Chem Chem Phys 1:2913
12. Ayuela A, Fowler PW, Mitchell D, Schmidt R, Seifert G, Zerbetto F (1996) J Phys Chem 100:15634
13. Campbell EEB, Fowler PW, Mitchell D, Zerbetto F (1996) Chem Phys Lett 250:544
14. Lu X, Chen Z, Thiel W, Schleyer P, Huang R, Zheng L (2004) J Am Chem Soc 126:14871
15. Zhechkov L, Heine T, Seifert G (2004) J Phys Chem A 108:11733
16. Diaz-Tendero S, Alcami M, Martin F (2005) Chem Phys Lett 407:153
17. Buhl M, Hirsch A (2001) Chem Rev 101:1153
18. Martin TP, Malinowski N, Zimmermann U, Naher U, Schaber H (1993) J Chem Phys 99:4210
19. Rabilloud F (2010) J Phys Chem A 114:7241
20. Hamamoto N, Jitsukawa J, Satoko C (2002) Eur Phys J D 19:211
21. Tachikawa H (2007) J Phys Chem C 111(35):13087
22. Tachikawa H (2011) J Phys Chem C 115(42):20406
23. Bernshtein V, Oref I (2001) Phys Rev A 63:043201
24. Santos J, Bulhoes L, Longo E, Varela J (1995) J Mol Struct Theochem 335:149

25. Zhao Y, Pan X, Zhou D, Su Z, Wang R (2003) Synth Met 227:135–136
26. Wang Q, Jena P (2012) J Phys Chem Lett 3(9):1084
27. Schon JH, Kloc C, Siegrist T, Steigerwald M, Svensson C, Batlogg B (2001) Nature 413(6858):831
28. Takenobu T, Iwasa Y, Ito T, Mitani T (2001) AIP Conf Proc 590(1):385
29. Palpant B, Otake A, Hayakawa F, Negishi Y, Lee GH, Nakajima A, Kaya K (1999) Phys Rev B 60:4509
30. Xie SY, Gao F, Lu X, Huang RB, Wang CR, Zhang X, Liu ML, Deng SL, Zheng LS (2004) Science 304(5671):699
31. Abboud JLM, Alkorta I, Davalos JZ, Gal JF, Herreros M, Maria PC, Mo O, Molina MT, Notario R, Yañez M (2000) J Am Chem Soc 122(18):4451
32. Irigoras A, Mercero JM, Silanes I, Ugalde JM (2001) J Am Chem Soc 123(21):5040
33. Rodriguez-Otero J, Cabaleiro-Lago EM, na∼Gallego AP, Montero-Campillo MM (2009) Tetrahedron 65(11):2368
34. Becke AD (1993) J Chem Phys 98:5648
35. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785
36. Haser M, Almof J, Scuseria GE (1991) Chem Phys Lett 181(6):497
37. 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 Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian∼09 Revision B.01 (2010). Gaussian Inc., Wallingford CT
38. Keith TA AIMAll 11.10.16 (2011) http://aim.tkgristmill.com/. TK Gristmill Software
39. Foster JP, Weinhold F (1980) J Am Chem Soc 102:7211
40. Reed AE, Weinhold F (1983) J Chem Phys 78:4066
41. Reed AE, Curtiss LA, Weinhold F (1988) Chem Rev 88(6):899
42. Glendening ED, Reed AE, Carpenter JE, Weinhold F NBO Version 3.1. http://www.gaussian.com/g_tech/g_ur/m_citation.htm
43. Bader RFW (1991) Chem Rev 91(5):893
44. Bader RFW (1994) Atoms in molecules: a quantum theory. Oxford University Press, Oxford

Comparison of pure and hybrid DFT functionals for geometry optimization and calculation of redox potentials for iron nitrosyl complexes with "μ-SCN" bridging ligands

**Nina Emel'yanova · Nataliya Sanina ·
Alexander Krivenko · Roman Manzhos ·
Konstantin Bozhenko · Sergey Aldoshin**

Abstract The geometry and electronic structure of neutral molecules and mono-anions of iron nitrosyl complexes with "μ-SCN" bridging ligands $[(\mu\text{-}SC_7H_4SN)Fe(NO)_2]_2$ (I) and $[(\mu\text{-}SC_2H_3N_4)Fe(NO)_2]_2$ (II) were studied using the density functional theory: pure functionals BP86, OLYP, OPBE, TPSS, PW91 and hybrid functionals B3LYP, B1B95, B3PW91 along with 6-311++G**//6-31G*. For geometry optimization, the pure functionals are more appropriate than the hybrid functionals. For mono-anions, pure functionals favor the doublet over the quartet state by 8.0–13.9 kcal/mol, whereas hybrid functionals yield quartet spin state of these anions to be more stable. The redox potentials of complexes I and II have been computed by DFT and compared with experimental results obtained from cyclic voltammetry method. The calculated redox potentials for I and II with pure functionals (especially BP86) are close to the experimental values, whereas those obtained by using hybrid functionals significantly deviate from the experiment.

Keywords DFT · Pure and hybrid functionals · Iron nitrosyl complexes · Redox potentials

1 Introduction

The nonheme iron nitrosyl complexes are of great interest due to their biological and pharmacological activities.

N. Emel'yanova (✉) · N. Sanina · A. Krivenko · R. Manzhos ·
K. Bozhenko · S. Aldoshin
Institute of Problems of Chemical Physics of RAS,
142432 Chernogolovka, Russian Federation
e-mail: n_emel@mail.ru

Along with nitrosothiols, they are assumed to be a potential candidate for storage and transport of nitrogen monoxide, which plays an important role in many physiological processes [1–5]. Once produced, NO can react with thiols and metal centers in proteins to regulate a variety of physiological processes, including controlling blood pressure, preventing platelet aggregation, modulating vasodilation, smoothing muscle proliferation, and acting as a biological messenger. In recent years, iron nitrosyl complexes have been widely investigated by DFT methods, but the main part of publications are devoted to Roussin's red and black salts and Roussin's red esters [6–12]. The theoretical prediction of the redox potentials for iron nitrosyl complexes using quantum chemical calculations [12–14] has been of particular interest because redox potentials of iron nitrosyl complexes are important parameters controlling the redox reactions. In our previous work [12], the geometry and electronic structure of mono-anions of Roussin's red esters $[Fe_2(\mu\text{-}RS)_2(NO)_4]$ (R = Me, Et, i-Pr, t-Bu) together with their redox potentials have been computed by the DFT methods. In this study, we try to use our calculations for more complicated systems—iron nitrosyl complexes with "μ-SSCN" bridging ligands—$[(\mu\text{-}SC_7H_4SN)Fe(NO)_2]_2$ (I) and $[(\mu\text{-}SC_2H_3N_4)Fe(NO)_2]_2$ (II)—in dichloromethane solution. These complexes differ from Roussin's red esters due to the presence of large Fe–Fe distances, with the value of intramolecular exchange interaction being inconsiderable; therefore, the complexes are paramagnetic at ambient temperature, and DFT calculations of these systems with complicated electronic structure present a considerable problem.

In order to compare the calculated values of redox potentials with the experimental ones, these complexes were studied by the cyclic voltammetry (CVA) method.

2 Computation details

The program package Gaussian03 [15] was used for all calculations. The pure functionals BP86, OLYP, OPBE, TPSS, PW91 and hybrid functionals B3LYP, B1B95, B3PW91 were used with 6-311++G**//6-31G* basis set. The frequency calculations demonstrated the absence of imaginary frequencies, proving that the optimized structures are true minima. Polarizable continuum model (PCM) was applied to predict the Gibbs free energies in dichloromethane using the extended basis 6-311++G**.

3 Molecular geometries

The structure of $[(\mu\text{-}SC_7H_4SN)Fe(NO)_2]_2$ (I) and $[(\mu\text{-}SC_2H_3N_4)Fe(NO)_2]_2$ (II) complexes obtained from our calculations is shown in Fig. 1. Table 1 presents a comparison between the optimized geometry of the $[(\mu\text{-}C_7H_4SN)Fe(NO)_2]_2$ (I) determined by using different functionals and experimental data [16], and Table 2 presents a comparison between the optimized geometry of the $[(\mu\text{-}SC_2H_3N_4)Fe(NO)_2]_2$ determined by using different functionals and experimental data [17]. As shown in the tables, the best functionals for predicting molecular structure were TPSS for complex I and OPBE for complex II (least error in Tables 1, 2). It should be noted that all chosen functionals except B1B95 overestimate the N–O bond length for complex I. For complex II, it was not possible to get the same great difference in the NO bond lengths as in the experiment. The Fe–N distances are better represented by using pure functionals, especially for complex II; however, they all underestimate this distance and all chosen hybrid functionals significantly overestimate it. For the Fe–N (ligand) bond length, pure functionals (except OLYP) describe it better. For the Fe–S bond, the OPBE and TPSS functionals show the best agreement with the experiment. As a whole, the theoretical values of geometrical parameters obtained with all the pure functionals are closer to the experiment than those found with hybrid functionals.

Fig. 1 The structure of $[(\mu\text{-}SC_7H_4SN)Fe(NO)_2]_2$ (I) and $[(\mu\text{-}SC_2H_3N_4)Fe(NO)_2]_2$ (II) complexes
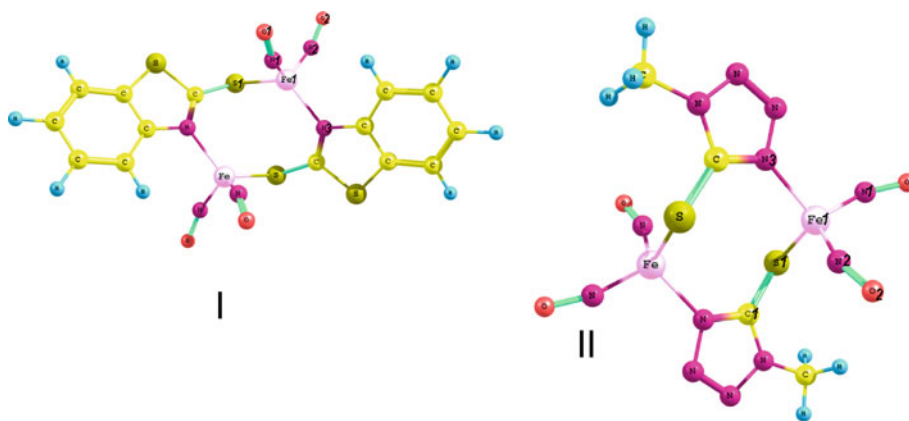
Table 1 Selected optimized geometry parameters (bond lengths, Å, angle, deg) for $[(\mu\text{-}C_7H_4SN)Fe(NO)_2]_2$ and comparison with experiment [16]

	Exp	BP86	OLYP	OPBE	TPSS	PW91	B3LYP	B1B95	B3PW91
Fe1-N1	1.684	1.654	1.664	1.646	1.652	1.648	1.743	1.736	1.728
Fe1-N2	1.689	1.655	1.667	1.644	1.658	1.654	1.744	1.746	1.731
Fe1-N3	2.046	2.048	2.120	2.081	2.049	2.044	2.104	2.078	2.084
Fe1-S1	2.290	2.269	2.309	2.275	2.279	2.267	2.343	2.331	2.324
N1-O1	1.174	1.188	1.183	1.177	1.185	1.185	1.181	1.172	1.176
N2-O2	1.162	1.187	1.182	1.176	1.185	1.185	1.179	1.172	1.175
N1-Fe1-N2	113.35	111.89	109.77	109.64	112.51	112.11	113.11	114.92	112.92
O1-N1-Fe1	167.37	163.2	159.29	159.89	165.1	164.51	160.1	165.47	160.91
O2-N2-Fe1	161.21	161.1	158.34	159.27	161.42	160.36	159.7	160.80	160.30
Δr		0.021	0.027	0.03	0.018	0.022	0.042	0.032	0.029
Max		0.034	0.074	0.045	0.032	0.036	0.059	0.057	0.044
Min		0.002	0.009	0.014	0.003	0.002	0.007	0.002	0.002
Δ0		1.19	4.84	4.38	0.86	1.65	3.00	1.29	2.60

The similar results have been obtained for Roussin's red esters [12]. This supports the view that the hybrid functionals usually perform better for main group chemistry and pure functionals perform better for transition metal chemistry [18].

4 Geometry and electronic structure of anions

The geometry optimization of mono-anions has been performed for two various spin states—doublet (d) and quartet (q). Unfortunately, no crystallographic data were reported for anions of iron nitrosyl complexes with "μ-SSCN" bridging ligands, and we did not have an opportunity to compare our results with the experiment. For all functionals, the mono-anions and the neutral molecules have the same structure. The exception is the mono-anion structure of complex I obtained by using B3LYP in doublet spin state; it will be discussed later. Table 3 presents the optimized geometry of $[(\mu\text{-}C_7H_4SN)Fe(NO)_2]_2^-$, and Table 4 presents the optimized geometry of $[(\mu\text{-}SC_2H_3N_4)Fe(NO)_2]_2^-$ determined by using different functionals. As shown in the tables, the key geometry parameters differ for different spin states. All the structures obtained from pure functionals have pseudo-centrosymmetric binuclear structure like the neutral molecules, whereas hybrid functionals give asymmetrical structure for doublet spin states for the most cases. This is associated with the spin density distribution in mono-anions obtained from different functionals (Table 5). It was already pointed out [19–21] that the spin density distributions for the iron nitrosyl complexes are strongly dependent on the choice of the exchange–correlation functional. As shown in Table 5, the spin density distribution for all the chosen pure

functionals is similar and it is symmetric. This concerns both the doublet and the quartet spin states. A completely different picture is obtained for all the chosen hybrid functionals. They give higher values of the spin density on the iron and NO-groups compared to pure functionals for the quartet spin states, and the spin density distribution for these states is almost symmetric. For the doublet spin states, from using B1B95 and B3PW91 functionals, the asymmetrical spin density distribution has been obtained, so this is reflected in optimized geometry of mono-anions.

Table 6 presents an energy difference between the doublet and the quartet state for all the chosen functionals. For mono-anion I, all the pure functionals favor the doublet over the quartet state by 8.0-13.9 kcal/mol. For mono-anion II, this value is slightly higher (14.4–17.4 kcal/mol) for BP86, OLYP, TPSS, and PW91 functionals. As noted previously [22], the behavior of the OPBE functional is somewhat ambiguous. In our case, it gives very large spin-state splitting (47.0 kcal/mol). All the chosen hybrid functionals favor quartet state for mono-anion II, although B3PW91 yields virtually the same total energy for both the low-spin state and the high-spin state (0.7 kcal/mol). For complex I, one can see the difference between the chosen hybrid functionals. B1B95 and B3PW91 predict the quartet spin state more stable, but values of spin-state splitting differ greatly (30.2 (B1B95) and 4.8 (B3PW91) kcal/mol). As noted above, the structure of the mono-anion of complex I obtained by using B3LYP in the doublet spin state is of practical interest. As shown in Table 6, B3LYP yields the doublet ground state favored by about 7.8 kcal/mol. It is seen from Table 5 that symmetrical spin density distribution has been obtained for this spin state for the mono-anion of complex I. But in this case, the optimized

Table 2 Selected optimized geometry parameters (bond lengths, Å, angle, deg) for $[(\mu\text{-}SC_2H_3N_4)Fe(NO)_2]_2$ and comparison with experiment [17]

	Exp	BP86	OLYP	OPBE	TPSS	PW91	B3LYP	B1B95	B3PW91
Fe1-N1	1.661	1.655	1.667	1.646	1.657	1.654	1.736	1.723	1.736
Fe1-N2	1.645	1.651	1.663	1.643	1.653	1.650	1.736	1.722	1.735
Fe1-N3	2.011	2.002	2.061	2.031	2.006	1.999	2.059	2.044	2.046
Fe1-S1	2.300	2.289	2.330	2.293	2.297	2.284	2.364	2.343	2.345
N1-O1	1.129	1.184	1.179	1.174	1.182	1.181	1.176	1.172	1.169
N2-O2	1.178	1.187	1.182	1.176	1.185	1.184	1.179	1.175	1.172
S1-C1	1.709	1.736	1.729	1.716	1.734	1.732	1.734	1.725	1.726
N1-Fe1-N2	114.3	111.6	109.2	109.4	112.3	111.8	113.6	113.6	115.8
O1-N1-Fe1	164.6	162.0	159.0	159.9	162.8	162.2	161.8	162.3	164.0
O2-N2-Fe1	167.5	163.3	160.2	160.8	163.7	163.3	162.1	162.6	164.5
Δr		0.018	0.025	0.014	0.015	0.017	0.050	0.040	0.043
Max		0.055	0.050	0.045	0.053	0.052	0.091	0.077	0.090
Min		0.006	0.004	0.002	0.003	0.005	0.001	0.003	0.004
$\Delta 0$		3.2	6.0	5.4	2.5	3.0	3.0	2.6	1.7

Table 3 Selected optimized geometry parameters (bond lengths, Å, angle, deg) for [(μ-C$_7$H$_4$SN)Fe(NO)$_2$]$_2$ (*d*-doublet state, *q*-quartet state)

	BP86		OLYP		OPBE		TPSS		PW91		B3LYP	B1B95		B3PW91	
	d	q	d	q	d	q	d	q	d	q	q	d	q	d	q
Fe–N	1.635	1.672	1.636	1.680	1.620	1.665	1.644	1.674	1.640	1.667	1.740	1.600	1.739	1.559	1.749
	1.642	1.667	1.648		1.625		1.634		1.632	1.670	1.762	1.622	1.773	1.610	1.727
												1.761		1.772	
												1.735		1.752	
Fe–N (L)	2.037	2.105	2.115	2.182	2.072	2.140	2.046	2.107	2.031	2.100	2.133	2.054	2.119	2.010	2.113
	2.030		2.085		2.045		2.015					2.023		2.136	
Fe–S	2.296	2.280	2.333	2.330	2.295	2.288	2.291	2.292	2.288	2.277	2.400	2.344	2.382	2.414	2.375
			2.348		2.281		2.306					2.339		2.456	
N–O	1.203	1.208	1.192	1.200	1.190	1.193	1.196	1.201	1.201	1.200	1.193	1.182	1.194	1.190	1.198
	1.197	1.203	1.196		1.186		1.201	1.207	1.194	1.205	1.202	1.190	1.185	1.190	1.189
												1.192		1.185	
												1.179		1.183	
N–Fe–N	109.0	112.7	107.6	110.7	107.2	110.7	109.8	113.3	109.1	112.7	113.8	117.4	115.5	109.1	113.7
												107.5		103.9	
Fe–N–O	157.9	152.6	155.3	152.1	164.1	152.8	167.2	157.5	167.2	156.5	150.8	148.8	153.0	154.5	151.2
	166.4	156.0	164.4	153.3	156.1	154.8	158.1	152.7	157.4	152.1	159.4	143.2	161.9	154.7	159.3
												156.9		158.8	
												168.6		169.1	

Table 4 Selected optimized geometry parameters (bond lengths, Å, angle, deg) for [(μ-SC$_2$H$_3$N$_4$)Fe(NO)$_2$]$_2$$^-$ (*d*-doublet state, *q*-quartet state)

	BP86		OLYP		OPBE		TPSS		PW91		B3LYP		B1B95		B3PW91	
	d	q	d	q	d	q	d	q	d	q	d	q	d	q	d	q
Fe–N	1.638	1.667	1.640	1.679	1.622	1.659	1.637	1.668	1.636	1.665	1.610	1.738	1.743	1.742	1.601	1.725
											1.613	1.741	1.748	1.746	1.604	
											1.741			1.749	1.727	
											1.748				1.733	
Fe–N(L)	1.970	2.057	2.026	2.125	1.994	2.094	1.971	2.062	1.968	2.052	2.033	2.051	2.016	2.134	2.004	2.041
												2.132	2.026	2.017	2.022	2.117
Fe–S	2.347	2.298	2.403	2.348	2.345	2.308	2.349	2.309	2.342	2.294	2.467	2.365	2.329	2.504	2.420	2.345
											2.351	2.475	2.411	2.333	2.332	2.433
N–O	1.200	1.211	1.194	1.204	1.189	1.197	1.199	1.203	1.197	1.201	1.195	1.189	1.179	1.176	1.192	1.187
	1.201	1.205	1.195	1.201		1.194		1.209	1.198	1.208	1.203	1.192	1.187	1.179	1.198	1.185
											1.185	1.202	1.193	1.199	1.181	1.197
											1.187	1.215			1.181	1.210
N–Fe–N	111.1	113.1	108.5	110.6	108.5	110.4	112.1	113.6	111.1	113.4	112.1	112.2	113.8	113.3	111.9	112.0
											111.6	116.7	112.9	121.2	111.2	116.3
Fe–N–O	163.2	156.4	160.0	153.1	160.4	154.2	164.2	157.7	163.2	156.3	163.8	156.5	162.0	161.4	164.1	157.5
	164.3	155.3	161.9	152.7	162.4	153.8	165.0	156.1	164.3	155.4	166.9	156.8	160.7	162.5	167.4	157.1
											157.6	158.4	165.0	162.5	158.5	158.8
											160.4	158.9	169.0	165.5	161.3	159.2

geometry of mono-anion differs from the neutral molecule. Figure 2 presents the optimized geometry of the obtained structure. One can see the breaking of two bonds, Fe–S and Fe–N, and the new μ—NO bridge between two iron atoms is formed. So we have obtained incorrect geometry for the doublet spin state from B3LYP. It is common knowledge [23] that for the calculation of spin-state splitting, standard pure functionals systematically overstabilize low-spin

Table 5 Spin density on atoms of $Fe(NO)_2$—sites in a 6-31G* basis set for selected functionals

Complex	Spin state		BP86	OLYP	OPBE	TPSS	PW91	B3LYP	B1B95	B3PW91
I	d	Fe	0.75	1.19	1.03	0.91	0.77	3.35	−1.61	3.43
			0.75	1.13	1.03	0.86	0.77	3.12	0.01	−0.03
		NO	−0.17	−0.36	−0.29	−0.22	−0.18	−1.32	−1.18	−1.29
			−0.16	−0.38	−0.31	−0.24	−016	−1.36	−1.39	−1.28
								−1.42	−0.18	0.14
								−1.48 (m)	0.17	−0.09
	q	Fe	1.59	2.04	1.93	1.74	1.61	3.22	3.37	3.20
			1.59	2.04	1.93	1.74	1.61	3.22	3.37	3.20
		NO	−0.12	−0.35	−0.30	−0.19	−0.14	−0.43	−1.42	−1.29
			−0.13	−0.34	−0.32	−0.20	−0.16	−0.40	−0.52	−0.52
II	d	Fe	0.78	1.18	1.06	0.92	0.80	3.31	3.45	3.29
			0.78	1.18	1.06	0.92	0.80	0.01	0.01	0.00
		NO	−0.17	−0.37	−0.31	−0.25	−0.18	−1.24	−1.33	−1.23
			−0.18	−0.38	−0.31	−0.26	−0.18	−1.24	−1.31	−1.23
								0	−0.01	0.01
								0	0.01	0.01
	q	Fe	1.63	2.07	1.97	1.78	1.66	3.26	3.47	3.23
			1.63	2.07	1.97	1.78	1.66	3.13	3.29	3.12
		NO	−0.11	−0.33	−0.29	−0.18	−0.13	−1.08	−1.30	−1.09
			−0.17	−0.37	−0.31	−0.23	−0.19	−1.13	−1.32	−1.06
								−1.30	−1.49	−1.28
								−0.12	−0.12	−0.13

Table 6 Energy differences between quartet and doublet state (kcal/mol) in a 6-31G* basis set for selected functionals

Complex	Spin state	BP86	OLYP	OPBE	TPSS	PW91	B3LYP	B1B95	B3PW91
I	d	0	0	0	0	0	0	30.2	4.8
	q	13.9	8.0	10.1	10.4	13.4	7.8	0	0
II	d	0	0	0	0	0	2.7	9.6	0.7
	q	17.4	15.4	47.0	14.4	16.1	0	0	0

states, while hybrid functionals overstabilize high-spin states due to the inclusion of a portion of Hartree–Fock exchange. For iron nitrosyl complexes, we have the similar situation. For example, it is common knowledge from the experimental data that Roussin's red esters are diamagnetic and EPR-silent due to the antiferromagnetic coupling between the dinuclear irons [8]. So the ground state is singlet. Our calculations of $[Fe_2(\mu\text{-EtS})_2(NO)_4]$ [12] show that the pure functionals predict singlet state to be more stable by 16.6 kcal/mol for BP86, by 12.4 kcal/mol for BLYP, by 12.8 kcal/mol for TPSS, and by 16.5 kcal/mol for PBE. Hence, the hybrid functionals grossly overestimate the triplet state (by 39.1 kcal/mol for B3LYP). However, for mono-anions of iron nitrosyl complexes with "μ-SSCN" bridging ligands, it seems impossible to find the right doublet states with hybrid functionals for the mono-

anions of iron nitrosyl complexes with "μ-SSCN" bridging ligands.

5 Calculation of the redox potentials

Redox potentials for iron nitrosyl complexes with "μ-SSCN" bridging ligands in the dichloromethane solution have been computed by the DFT. As noted previously [24–36], the DFT quite accurately predicts the redox potentials. It has been shown to reproduce experimental results satisfactorily on organic compounds [32] and also on metal complexes [36]. In each specific case, the question is the optimal choice of approach (functional/basis). It has been previously shown that the pure functionals are more appropriate for transition metal compounds, while the hybrid
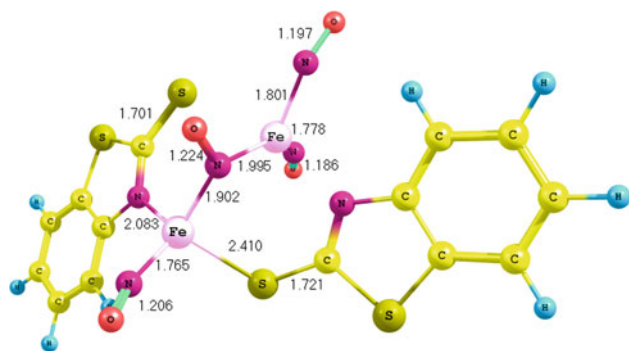
Fig. 2 Optimized geometry $[(\mu\text{-}C_7H_4SN)Fe(NO)_2]_2^-$ (doublet spin state, B3LYP/6-31G*)

functionals act satisfactorily after the introduction of correction coefficients [25].

The procedure for calculating the redox potentials of nitrosyl complexes in solution is based on the Born–Haber thermodynamic cycle, which relates the electron attachment in the condensed phase to the corresponding process in the gas phase through the free energies of solvation for reactants and products. Taking into account expression $E^0 = -\Delta G^0/nF$, the value for the reduction potential for considered iron nitrosyl complexes versus SHE can be calculated as follows:

first reduction potentials for I and II, the calculated values are more negative, a similar trend has been noticed previously [8, 38, 39]. For complex I, the most appropriate potentials are obtained by using BP86 and PW91, and the difference between calculated and experimental redox potentials does not exceed 0.1 V. For TPSS, our calculations of the redox potentials by Eq. (1) give values which agree with the experimental ones with deviation 0.282 V, and this result is quite satisfactory too. Wang et al. [8], who have calculated redox potentials for four Roussin's red esters with alkyl and aril substituents, have obtained values differing from the experimental ones by 0.042–0.360 V. Redox potentials from OPBE and OLYP are more than 0.500 V different from the experimental value. It should be noted that the discrepancy between the redox potentials obtained from B3LYP and the experimental values (0.481 V) is less than the one obtained from OPBE and OLYP. However, as noted above, the structure of the doublet state obtained from B3LYP differs from the neutral molecule, so the first redox potential has to be irreversible, but it is reversible according to the experimental data. As to calculations with B3PW91 and B1B95, we obtained a very large error. For complex II, we have obtained the similar results. All the chosen pure functionals predict redox potentials quiet satisfactory, especially BP86 (deviation is

$$E^{n/n-1} = -\frac{G^0\Big([Fe_2(\mu_2 - SCNR)_2(NO)_4]^{n-1}\Big) - G^0\Big([Fe_2(\mu_2 - SCNR)_2(NO)_4]^n\Big)}{F} + 4.36, \tag{1}$$

where F is the Faraday constant, and G^0 is the standard free energy of the $[Fe_2(\mu_2 - SCNR)_2(NO)_4]^n$ complex, consisting of the sum of the total energy in the gas phase, a thermal correction to the Gibbs free energy, and the free solvation energy. 4.36 eV [37]—free energy associated with the reference electrode process $H^+(aq) + e^-(g) \rightarrow \frac{1}{2} H_2(g)$.

Calculated and experimental values of the redox potentials for complexes I and II in the dichloromethane solution are presented in Table 7. It should be noted that for our calculations of the redox potentials, we used the lowest energy structures of mono-anions, so it is the doublet spin state for all the chosen pure functionals for the both complexes and B3LYP for complex I, and it is the quartet spin state for B1B95 and B3PW91 for the both complexes and B3LYP for complex II. As shown in Table 7, when comparing with the experimental data of the

only 0.033 V), whereas all the chosen hybrid functionals give very large differences between calculated and experimental redox potentials (deviation is more than 1.000 V). So, the calculated redox potentials for I and II with the pure functionals (especially BP86) are close to the experimental values, whereas those obtained by using the hybrid functionals significantly deviate from the experiment. This is consistent with our previous calculations of redox potentials for Roussin's red esters [12]. For these complexes, BP86 functional provides excellent agreement with experiment, whereas B3LYP requires some correlation. But for these more complicated complexes, the difference between theory and experiment is large, and any correlations are not possible, so we can use the pure functionals only. Most likely it is directly due to problems that arise in calculations of low-spin states with the hybrid functionals.

Table 7 Calculated redox potentials (V/SHE) for $[Fe_2(\mu\text{-RS})_2(NO)_4]^-$ R = Me, Et, i-Pr, t-Bu and comparison with experiment. Differences between calculated and experimental redox potentials are shown in brackets (in volts)

	Exp	BP86	OLYP	OPBE	TPSS	PW91	B3LYP	B3PW91	B1B95
I	−0.306	−0.403 (0.097)	−0.861 (0.555)	−0.813 (0.507)	−0.588 (0.282)	−0.384 (0.078)	−0.787 (0.481)	−1.177 (0.871)	−1.386 (1.080)
II	−0.196	−0.229 (0.033)	−0.730 (0.534)	−0.655 (0.459)	−0.497 (0.301)	−0.334 (0.138)	−1.216 (1.020)	−1.302 (1.106)	−1.215 (1.019)

6 Conclusions

In the present work, the geometry and the electronic structure of the neutral molecules and the mono-anions of the iron nitrosyl complexes with "μ-SSCN" bridging ligands $[(\mu\text{-SC}_7H_4SN)Fe(NO)_2]_2$ (I) and $[(\mu\text{-SC}_2H_3N_4)Fe(NO)_2]_2$ (II) together with their redox potentials in the dichloromethane solutions have been computed by DFT: the pure functionals BP86, OLYP, OPBE, TPSS, PW91 and the hybrid functionals B3LYP, B1B95, B3PW91 along with 6-311++G**//6-31G*. For the geometry optimization of these complexes, the pure functionals are more appropriate than the hybrid functionals. TPSS for complex I and OPBE for complex II give the least average errors for the main optimized geometry parameters. For all the functionals, the geometry optimization of the mono-anions in two spin states—the doublet and the quartet—predicts the same structure as the neutral molecules except the structure of the mono-anion I in the doublet spin state obtained from B3LYP. We have found that all the chosen pure functionals predict the low-spin state to be more stable for the both anions, whereas the hybrid functionals favor the high-spin state to be more stable. A possible reason for such a difference is incorrect spin density distributions of the doublet spin state obtained with the hybrid functionals (high-spin contamination). For calculations of redox potentials, we can only use the pure functionals, they give the most accurate value (especially BP86), whereas the hybrid functionals are unsuitable for our systems due to incorrect doublet spin state, they give a very large error.

References

1. Sanina NA, Aldoshin SM (2004) Russ Chem Bull Int Ed 53:2428–2448
2. Traylorand TG, Sharma VS (1992) Biochemistry 31:2847–2849
3. Lai TS, Hausladen A, Slaughter TF, Eu JP, Stamlerand JS, Greenberg CS (2001) Biochemistry 40:4904–4910
4. Huang J, Hadimani SB, Rupon JW, Ballas SK, Kim-Shapiroand DB, King SB (2002) Biochemistry 41:2466–2474
5. Silaghi-Dumitrescu R, Kurtz DM Jr, Ljungdahland LG, Lanzilotta WN (2005) Biochemistry 44:6492–6501
6. Jaworska M, Stasicka Z (2005) New J Chem 29:604–612
7. Jaworska M, Stasicka Z (2006) J Mol Struct 785:68–75
8. Wang R, Camacho-Fernandez MA, Xu W, Zhangand J, Li L (2009) DaltonTrans (5):777–786
9. Hopmann KH, Conradie J, Ghosh A (2009) J Phys Chem B 113:10540–10547
10. Hopmann KH, Noodleman L, Ghosh A (2010) Chem Eur J 16:10397–10408
11. Sanina NA, Emel'yanova NS, Chekhlov AN, Shestakov AF, Sulimenkov IV, Aldoshin SM (2010) Russ Chem Bull Int Ed 59:1126–1136

12. Emel'yanova NS, Shestakov AF, Sanina NA (2012) DFT calculations of the redox-potentials for the nitrosyl complexes [Fe$_2$(μ-RS)$_2$(NO)$_4$] with R = alkyl. Int J Quantum Chem. doi: 10.1002/qua.24063

13. Surawatanawong P, Sproules S, Neese F, Wieghardt K (2011) Inorg Chem 50:12064–12074

14. Hughes TF, Friesner RA (2012) J Chem Theory Comput 8:442–459

15. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, 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 MLX, 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 (2004) Gaussian 03, Revision D.01. Gaussian Inc., Wallingford, CT

16. Tsai ML, Hsieh CH, Liaw WF (2007) Inorg Chem 46:5110–5117

17. Sanina NA, Korchagin DV, Shilov GV, Kulikov AV, Shestakov AF, Sulimenkov IV, Aldoshin SM (2010) Russ J Coord Chem 36:876–886

18. Cramer CJ, Truhlar DG (2009) Phys Chem Chem Phys 11:10757–10816

19. Boguslawski K, Jacob CR, Reiher M (2011) J Chem Theory Comput 7:2740–2752

20. Ghosh AJ (2006) Biol Inorg Chem 11:712–724

21. Conradie J, Ghosh A (2007) J Phys Chem B 111:12621–12624

22. Swart M (2008) J Chem Theory Comput 4:2057–2066

23. Swart M, Groenhof AR, Ehlers AW, Lammertsma K (2004) J Phys Chem A 108:5479–5483

24. Winget P, Cramer CJ, Truhlar DG (2004) Theor Chem Acc 112:217–227

25. Roy LE, Jakubikova E, Guthrie MG, Batista ER (2009) J Phys Chem A 113:6745–6750

26. Namazian M, Zare HR, Coote ML (2008) Biophys Chem 132:64–68

27. Kelly CP, Cramer CJ, Truhlar DG (2007) J Phys Chem B 111:408–422

28. Song Y, Zhou JF, Song Y, Wei Y, Wang H (2005) Bioorg Med Chem Lett 15:4971–4975

29. Song Y (2007) Spectrochim Acta, Part A 67:611–618

30. Speelman AL, Gillmore JG (2008) J Phys Chem A 112:5684–5690

31. Datta SN, Sudhamsu J, Pandey A (2004) J Phys Chem B 108:8007–8016

32. Tugsuz TJ (2010) Phys Chem B 114:17092–17101

33. Silva PJ, Ramos MJ (2011) Comput Theor Chem 966:120–126

34. Shimodaira Y, Miura T, Kudo A, Kobayashi HJ (2007) Chem Theory Comput 3:789–795

35. Baik MH, Friesner RA (2002) J Phys Chem A 106:7407–7412

36. Uudsemaa M, Tamm T (2003) J Phys Chem A 107:9997–10003

37. Tuttle TR, Malaxos S, Coe JV (2002) J Phys Chem 106:925–932

38. Lewis A, Bumpus JA, Truhlar DG, Cramer CJ (2004) J Chem Educ 81:596–604

39. Patterson EV, Cramer CJ, Truhlar DG (2001) J Am Chem Soc 123:2025–2031

REGULAR ARTICLE

Organometallic copper I, II or III species in an intramolecular dechlorination reaction

Albert Poater · Luigi Cavallo

Abstract The present paper gives insight into an intramolecular dechlorination reaction involving Copper (I) and an $ArCH_2Cl$ moiety. The discussion of the presence of a Cu^{III} organometallic intermediate becomes a challenge, and because of the lack of clear experimental detection of this proposed intermediate, and due to the computational evidence that it is less stable than other isomeric species, it can be ruled out for the complex studied here. Our calculations are completely consistent with the key hypothesis of Karlin et al. that $TMPA-Cu^I$ is the substrate of intramolecular dechlorination reactions as well as the source to generate organometallic species. However the organometallic character of some intermediates has been refused because computationally these species are less stable than other isomers. Thus this study constitutes an additional piece towards the full understanding of a class of reaction of biological relevance. Further, the lack of high energy barriers and deep energy wells along the reaction pathway explains the experimental difficulties to trap other intermediates.

Keywords Organometallic complexes · DFT calculations · TMPA · Dechlorination · CuIII · Dehalogenation · Organometallic

A. Poater (✉)
Departament de Química and Institut de Química
Computacional, Universitat de Girona, Campus de Montilivi,
17071 Gerona, Catalonia, Spain
e-mail: albert.poater@udg.edu

L. Cavallo
KAUST Catalyst Research Center, 4700 King Abdullah
University of Science and Technology, Thuwal 23955-6900,
Kingdom of Saudi Arabia

1 Introduction

Bearing in mind the high industrial attractiveness for Cu because of its low toxicity and low cost, most of recent studies are focused on the binding and activation of O_2 by a single Cu ion in biological systems [1–9], however it is still a challenge the study of the same compounds in reductive dehalogenation reactions occurred in anaerobic media [10, 11] such as

$$2LCuI + 2R - Cl \rightarrow 2LCuII - Cl + R - R \qquad (1)$$
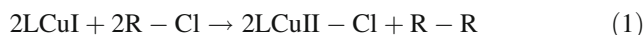
A large number of papers report on copper-ion-mediated R–X dehalogenations [12–14], a very active subfield that includes atom transfer radical addition (ATRA) and polymerization (ATRP) employing copper (I) complexes with tri- or tetraaza ligands [15, 16]. Research on R–X/Cu-enzyme reactivity is focused on the detoxification and remediation of halogenated organic pollutants [11, 17, 18]. However, comprehension of the stability and reactivity of the intermediates of intramolecular dechlorination still is incomplete [19–21].

The reaction of copper (I) complexes with exogenous organohalide substrates contrasts with a ligand that affords the dechlorination reaction intramolecularly from a chloromethylene substrate, such as TMPA (tris(2-pyridyl-methyl)amine) [19]. Overall Cu-alkyl intermediates are demonstrated to be present as intermediates according to

the results, however, details about the differentiation between Cu[III]-R and Cu[II]-R bonds are still missing. Experimental and theoretical research have described copper(III) intermediates in C–C coupling reactions by lithium organocopper chemistry [22–25]. However the characterization of formally Cu(III) species is still something odd and just a few examples are known [21–23, 26–30], with scarce computational insights [31]. Species like (TMPA)Cu[I] have been found active towards the dehalogenation of R–X substrates to form halogenated Cu[II] species as well as C–C coupling products (see Eq. 1) [10].

Karlin et al. proposed an organometallic copper(III) species, [19] in the context of an intramolecular dechlorination reaction, species **B** in Scheme 1. In the overall mechanism proposed, to start, the Cu[I] ligand adduct $[(6ClCH_2–TMPA)Cu^I]^+$ (**A**) rapidly undergoes a reversible intramolecular oxidative insertion of the C–Cl substrate, giving this intermediate **B**. A rate-determining electron transfer from a second molecule **A** ensues, giving **C** + **D**; the latter species maintains its chloromethylene group. Following this, protonation by solvent–water of the organometallic leads to a methyl-pyridyl containing product.

In this study a TMPA ligand includes an internal ligand R–Cl substrate [11], and a chloromethylene group in a pyridyl 6-position of one TMPA arm. There are evidences of organometallic species in the initial reaction stages [19]. However, new theoretical insights are required to describe which oxidation states the metal displays in these organometallic species and to describe the reaction mechanism. In this communication we provide a full theoretical characterization of all the species involved in the dechlorination reaction displayed in Scheme 1 [19].
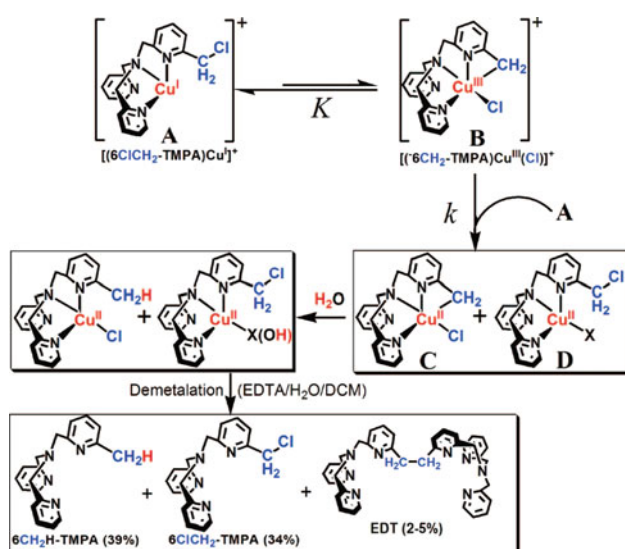
Scheme 1 Intramolecular dechlorination mechanism proposed by Karlin et al.

2 Theoretical methods

All geometry optimizations, with no symmetry constraints, have been performed with the Gaussian03 package [32], using the B3LYP functional [33–35] and the standard 6-31G(d) basis set [36, 37]. The geometries obtained at the B3LYP/6-31G(d) level were used to perform single-point energy calculations with a larger basis set, the 6-311G(d,p) basis set [38], and the same functional (B3LYP/6-311G(d,p)//B3LYP/6-31G(d)). Intrinsic reaction pathways were calculated to confirm that the transition states located connected the expected minima. Analytical Hessians were computed to determine the nature of all the stationary points we located, and to calculate zero-point energies (ZPEs) and thermodynamic properties at 298 K.

For open-shell states, the geometry optimizations were performed within the broken-symmetry unrestricted methodology, while for the closed-shell singlet states the restricted formalism was used. Theoretical treatment of biradical singlet species requires multiconfigurational or multireference methods due to strong static electron correlation. Unfortunately, these methods can only be applied to relatively small systems because computationally they are extremely demanding. As an alternative, we have used the unrestricted UBLYP method in broken symmetry (BS, using GUESS = MIX) [39]. This method improves the modeling of biradical singlet states at the expense of introducing some spin contamination from higher spin states [40–50]. Although this is not the most appropriate method to treat biradical singlet species, it has been shown that it can be used provided that the overlap between the open-shell orbitals is small (i.e. the unpaired electrons are located in separated atomic centers), which is the case of the systems that show predominant biradical character in this work [46]. In a previous paper, the application of the sum rule [45] to the energy of the biradical singlet state to remove the spin contamination error showed that this solution does not improve the calculated singlet–triplet energy gap but rather leads to larger errors when compared to the experimental value [39]. For this reason, closed shell singlet energies reported in this work does not contain spin contamination corrections.

Solvent effects including contributions of non electrostatic terms have been estimated in single point calculations on the gas phase optimized structures, based on the polarizable continuous solvation model (PCM) using CH_3CN as a solvent [51, 52]. The cavity is created via a series of overlapping spheres.

The relative free energies reported in this work include energies computed at the B3LYP/6-311G(d,p)//B3LYP/6-31G(d) level of theory together with solvent effects, while ZPEs and corrections to the Gibbs free energy were calculated at the B3LYP/6-31G(d) level at 298 K.

3 Results and discussion

The reaction path starts from interaction of $[Cu^I(CH_3CN)_4]^+$ with the 6ClCH$_2$-TMPA ligand to form $[Cu^I(6ClCH_2$-TMPA$)]^+$ complex, **A**, with the singlet electronic state. In Fig. 1, species **A** is suggested to be poorly in equilibrium with species **B**, with a Cu^{III} metal center, as proposed in Scheme 1. Species **A** displays only the closed-shell singlet ground state, because the open-shell singlet converges to the closed-shell configuration, and indeed the triplet state is nearly 50 kcal mol^{-1} above in energy. On the other hand, the closed-shell singlet ground state species **B** is located computationally 20.4 kcal mol^{-1} higher in energy with respect to **A^S**, with a barrier for the isomerisation from **A^S** placed at 26.3 kcal mol^{-1}. When exploring all isomerisation processes, it is possible to conclude that through a diradical singlet transition state complex **B** displays a triplet ground state multiplicity, being 0.5 kcal mol^{-1} less stable than **A^S**. Thus, the energy barrier of 17.8 kcal mol^{-1} to reach **B^S** from **A^S** suggests that this barrier itself is 8.5 kcal mol^{-1} with respect to the proposed one supposed to display the intermediate **B^S**, containing the Cu^{III} metal center. It must be pointed out that the ground state multiplicity of complex **B** cannot be distinguished between diradical singlet and triplet because the diradical singlet multiplicity state is less than 0.1 kcal mol^{-1} higher in energy. This complex **B^T** displays no organometallic bond, i.e. there is not the Cu-C bond with the 6ClCH$_2$-TMPA ligand that would allow the formation of the Cu^{III} intermediate. Moving to the Natural Population Analysis (NPA), the charges on Cu are 0.518, 1.024, and 1.288 e$^-$, for **A^S**, **B^T** and **B^S**, respectively, thus the later values are clear to understand their Cu^I, Cu^{II}, and Cu^{III} metal centers. On the other hand, the charges on the nearest C atom to the Cu metal center are -0.174, 0.225, and -0.434 e$^-$, which clearly describe the different nature of the cloromethylenic moiety, the free methylenic group, and finally the organometallic methylenic unit for **A^S**, **B^T** and **B^S**, respectively.

Considering the debate on which is the most suitable density functional to describe the low lying electronic states of Cu-Cl complexes, we tested our B3LYP conclusions with the pure BLYP functional. As expected, absolute energies are somewhat different, but the most important point is that relative energies are substantially the same, which indicates that the chemical scenario we described is rather independent of the specific computational method [53]. The difference between **A^S** and **B^S** was found to be a bit lower, 17.8 kcal mol^{-1} for BLYP with respect to 20.4 kcal mol^{-1} for B3LYP, thus confirming the validity of the computational method, because in qualitative terms the results were the same. Overall, we investigated all possible multiplicity configurations from species **A** to
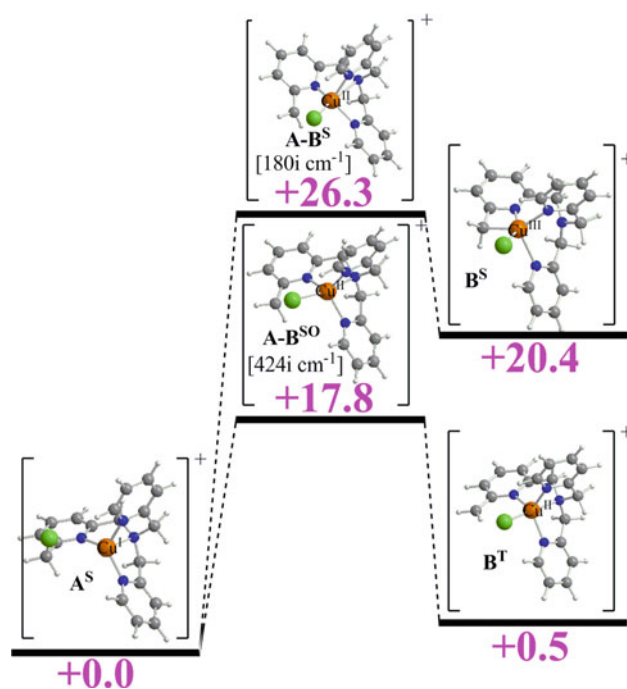

Fig. 1 Stationary points located for the equilibrium **A → B** step. Free energies relative to **A^S** (in kcal mol^{-1}) in solution. Calculated imaginary frequencies for transition state structures are given in brackets. Superindexes **S** (*closed-shell singlet*), **SO** (*open-shell singlet*), and **T** (*triplet*) refer to the multiplicity of the ground state

B with different functionals obtaining the same energetics order. To further validate the method we compared the optimized structures of **A**, **D**, and **E** with the corresponding X-ray data indicating that the DFT approach we choose can offer reliable insights on the geometry of other intermediates along the reaction path, for which no X-ray data are available.

The next step, depicted in Fig. 2, describes the step from **A^S** + **B^T** to **C^D** + **D^D**, which involves a disproportion, somewhat similar to other Cu complexes [21]. Bearing in mind the experimental evidence of the formation of an organometallic species, this species is theoretically found in **C^D**, from **B^T**, however with a Cu^{II} ion instead of the proposed Cu^{III} species. Species **C–F** display a doublet electronic ground state. To just equilibrate the disproportion equation, first an optimization of complex **D^D** without the insertion of a chloride anion was achieved, **D^D$_{bis}$**, i.e. the moiety **A** with a Cu^{II} metal center instead of Cu^I. This increase of the positive charge of the metal is stabilized by the rotation of the CH$_2$Cl moiety in order to create a favorable interaction between the metal and the chloride anion. Indeed the Cu-Cl distance is only 2.496 Å. However, after evaluating the energetics, the conclusion is that the disproportion is not favorable by 22.3 kcal mol^{-1}, which energy amount decreases to only 4.2 kcal mol^{-1}

when an entering chloride atom is bonded to Cu to form species D^D.

As found for B^T, in Fig. 2, species C^D displays no organometallic character, against the proposal of Scheme 1. The conversion from B^T to C^D includes first of all an increase of one electron, centered in the metal ion,
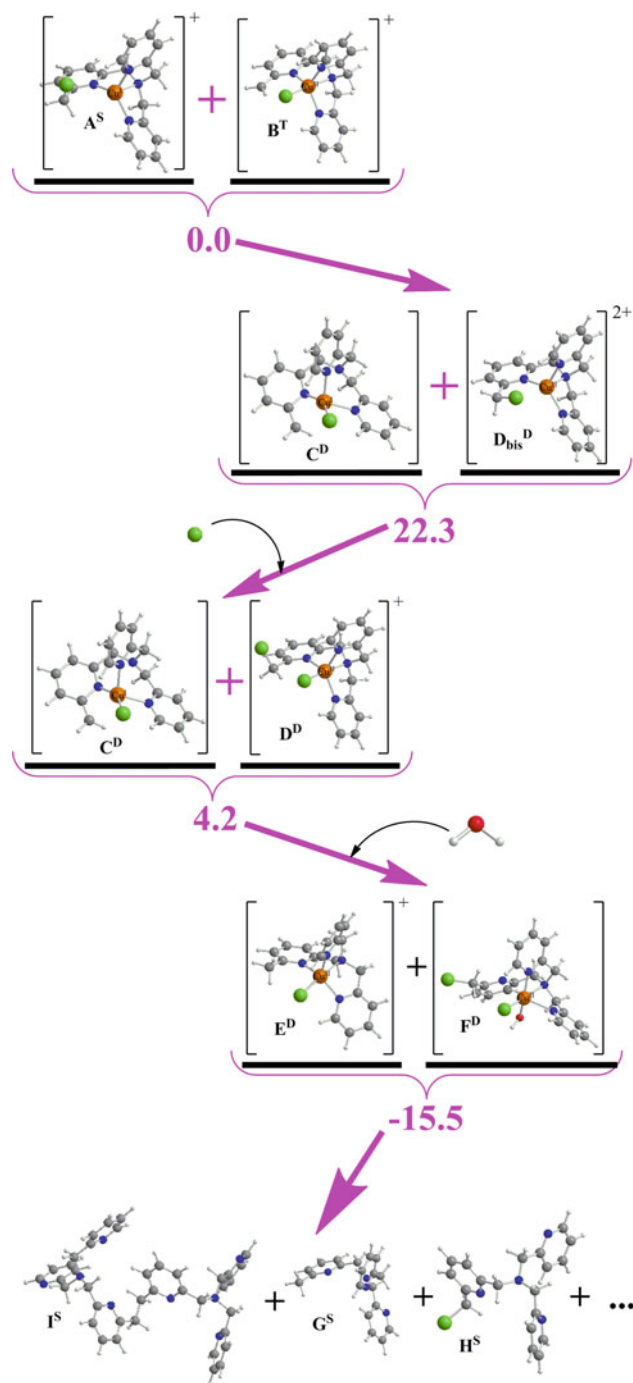

Fig. 2 Stationary points located along the reaction path of the dechlorination mechanism. Free energies relative to A^S (in kcal mol^{-1}) in solution. Superindexes **S** (*closed-shell singlet*), **D** (*doublet*), and **T** (*triplet*) refer to the multiplicity of the ground state

due to the disproportion. Overall species C^D presents a neutral charge. On the other hand, species C^D is different with respect to the homologous one proposed in Scheme 1 because computationally it is found to be not organometallic, i.e. there is not any Cu-C bond. However, the characterization of the proposed species **C** in Scheme 1, here called C_{iso}^D, using species C^D in Fig. 2 as its precursor is displayed in Fig. 3. C_{iso}^D indeed displays a similar shape with respect to species C^D, differentiated only by the presence of the Cu-C bond and the absence of a Cu–N bond. After overcoming a barrier of only 2.9 kcal mol^{-1} species C_{iso}^D is reached, which is 2.2 kcal mol^{-1} less stable than **C**. In this case, the formation of the Cu-C bond supposes an oxidative process, anyway, the organometallic nature of the intermediate C_{iso}^D is found energetically not favorable too here, however its presence cannot be ruled out because of the low isomerisation barrier between C^D and C_{iso}^D. Bearing in mind the different nature of both isomers of species **C**, the NPA reveals that the charges on Cu are 0.716 and 0.852 e$^-$, for C^D and C_{iso}^D, respectively, and the charge on the C of the organometallic C^D isomer is positive, 0.071 e$^-$, whereas for C_{iso}^D is negative, -0.185 e$^-$. The nature of the organometallic structure C^D is reinforced by a Mulliken analysis, because the spin density is 0.289 and 0.525 e$^-$ on Cu and C, respectively, to be compared with 0.478 and 0.356 for C_{iso}^D.

The presence of water allows the formation of complexes E^D and F^D from C^D and D^D. A proton of the water is used to protonate the methylidenic group of C^D and, whereas complex D^D takes the hydroxyl group bonded to the metal instead of acting as a simple counterion because energetically this Cu–O interaction stabilizes strongly the structure, being $E^D + F^D$ 19.7 kcal mol^{-1} more stable with respect to $C^D + D^D$.

Bearing in mind the products of the reaction after a process of demetalation are 6ClCH$_2$-TMPA (exp. 39 %),
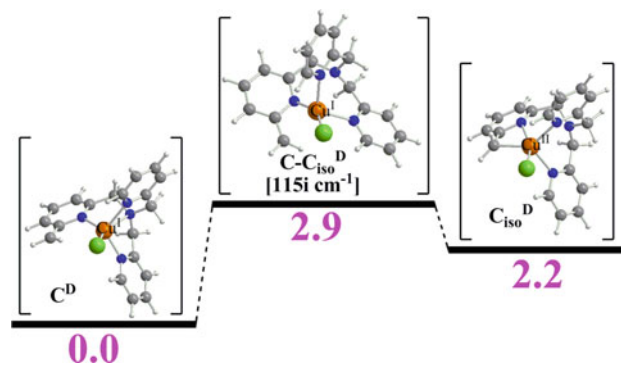

Fig. 3 Stationary points located for the equilibrium between the different isomers of species **C**. Free energies relative to complex **C** (in kcal mol^{-1}) in solution. A calculated imaginary frequency for the transition state is given in *brackets*. Superindex **D** (*doublet*) refers to the multiplicity of the ground state

G, 6ClCH$_2$-TMPA (exp. 34 %), **H**, and a dimer of ligands, **I** (exp. 2–5 %), combination of either **G** or **H** the analysis is not trivial. From **G**, the generation of the dimer costs 19.2 kcal mol^{-1}, *i. e.* with the subsequent formation of a hydrogen molecule, instead of the 29.7 kcal mol^{-1} of the dimer of chloro atoms from two **H** molecules. The lower stability for the dimer confirms the low experimental percentage of this species. However it is not possible to characterize species **G-I** energetically with respect to species **A** and **B** due to the difficulty to understand where the metal atom goes since the last step. Currently, further efforts are carried out to understand the inclusion of explicit solvent molecules.

Apart from the mentioned dimers by ligands it has been confirmed by X-Ray data the formation of dimers with charge +2, and triplet as ground state. There are two types of dimers, either with CH$_2$Cl groups (see Fig. 4) or CH$_3$ groups. The monomers of both types of dimers are the species **D** and **E**, respectively. However, in Fig. 4 the interaction between both monomers is demonstrated to be weak displaying Cu···Cl distances of 3.860 Å linking both units, very elongated with respect to the 3.302 Å corresponding to the Cu-Cl bond distances. Bearing in mind a high similarity between the dimer with CH$_3$ groups and the one with CH$_2$Cl groups, these dimers, reinforce the idea that there are chloride anions free in solution, which allows the formation of species **D** from **D$_{bis}$**. Furthermore, the dimers with CH$_3$ and CH$_2$Cl groups confirm the reliability of species **E** and **D**, respectively.

4 Conclusions

In summary, our calculations are completely consistent with the key hypothesis of Karlin et al. that TMPA-Cu^I is the substrate of intramolecular dechlorination reactions[19] as well as the source to generate organometallic species. However the organometallic character of some intermediate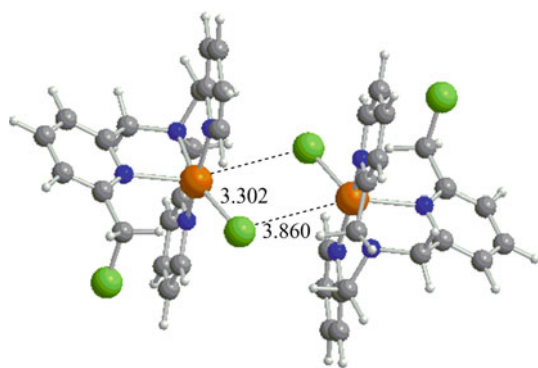s has been refused because computationally these species are less stable than other isomers. Thus this paper constitutes an additional piece towards the full understanding of a class of reaction of biological relevance. Further, the lack of high energy barriers and deep energy wells along the reaction pathway explains the experimental difficulties to trap other intermediates.

References

1. Itoh S (2006) Curr Opin Chem Biol 10:115–122
2. Gherman B, Heppner D, Tolman W, Cramer CJ (2006) J Biol Inorg Chem 11:197–205
3. Chen P, Solomon EI (2004) Proc Natl Acad Sci USA 101:13105–13110
4. Fukuzumi S, Karlin KD (2013) Coord Chem Rev 257:187–195
5. Quant Hatcher L, Karli KD (2004) J Biol Inorg Chem 9:669–683
6. De la Lande A, Gérard H, Parisel O (2008) Int J Quant Chem 108:1898–1904
7. De la Lande A, Parisel O, Gérard H, Moliner V, Reinaud O (2008) Chem Eur J 14:6465–6473
8. Poater A, Cavallo L (2009) Inorg Chem 48:4062–4066
9. De la Lande A, Gérard H, Moliner V, Izzet G, Reinaud O, Parisel O (2006) J Biol Inorg Chem 11:593–608
10. Osako T, Karlin KD, Itoh S (2005) Inorg Chem 44:410–415
11. Lucchese B, Humphreys KJ, Lee D-H, Incarvito CD, Sommer RD, Rheingold AL, Karlin KD (2004) Inorg Chem 43:5987–5998
12. Lipshutz BH, Parker D, Kozlowski JA, Miller RD (1983) J Org Chem 48:3334–3336
13. Nondek L, Hun LG, Wichterlova B, Krupicka S (1987) J Mol Catal 42:51–55
14. Kern J-M, Sauvage J-P (1987) J Chem Soc Chem Commun 546–548
15. Eckenhoff WT, Garrity ST, Pintauer T (2008) Eur J Inorg Chem 563–571
16. Matyjaszewski K, Xia JH (2001) Chem Rev 101:2921–2990
17. Wackett LP, Schanke CA (1992) Met Ions Biol Syst 28:329–356
18. Castro CE (1998) Rev Environ Contam Toxicol 155:1–67
19. Maiti D, Narducci Sarjeant AA, Itoh S, Karlin KD (2008) J Am Chem Soc 130:5644–5645
20. Casitas A, Poater A, Solà M, Stahl SS, Costas M, Ribas X (2010) Dalton Trans 39:10458–10463
21. Xifra R, Ribas X, Llobet A, Poater A, Duran M, Solà M, Stack TDP, Benet-Buchholz J, Donnadieu B, Mahia J, Parella T (2005) Chem Eur J 11:5146–5156
22. Bertz SH, Cope S, Murphy M, Ogle CA, Taylor BJ (2007) J Am Chem Soc 129:7208–7209
23. Bartholomew ER, Bertz SH, Cope S, Dorton DC, Murphy M, Ogle CA (2008) Chem Commun 1176–1177
24. Hu H, Snyder JP (2007) J Am Chem Soc 129:7210–7211
25. Gartner T, Henze W, Gschwind RM (2007) J Am Chem Soc 129:11362–11363
26. Ribas X, Jackson DA, Donnadieu B, Mahia J, Parella T, Xifra R, Hedman B, Hodgson KO, Llobet A, Stack TDP (2002) Angew Chem Int Ed 41:2991–2994
27. Furuta H, Maeda H, Osuka A (2000) J Am Chem Soc 122:803–807

Fig. 4 Dimer of species **D** (distances in Å)

28. Pratesi A, Zanello P, Giorgi G, Messori L, Laschi F, Casini A, Corsini M, Gabbiani C, Orfei M, Rosani C, Ginanneschi M (2007) Inorg Chem 46:10038–10040

29. Brasche G, Buchwald SL (2008) Angew Chem Int Ed 47: 1932–1934

30. King AE, Huffman LM, Casitas A, Costas M, Ribas X, Stahl SS (2010) J Am Chem Soc 132:12068–12073

31. Poater A, Cavallo L (2009) Inorg Chem 48:2340–2342

32. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, 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 Ö, 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 (2004) Gaussian 03, revision C02; Wallingford, CT

33. Becke AD (1993) J Chem Phys 98:5648–5652

34. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785–789

35. Stevens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) Phys Chem 98:11623–11627

36. Hehre WJ, Ditchfield R, Pople JA (1972) J Chem Phys 56:2257–2261

37. Hehre WJ, Radom L, Schleyer PVR, Pople JA (1986) Ab initio molecular orbital theory. Wiley, New York

38. For Cu, this basis corresponds to the (14s9p5d)/[9s5p3d] Wachters basis set (Wachters AJH (1970) J Chem Phys 52:1033–1036) with the contraction scheme 611111111/51111/311 supplemented with one f polarization function. In the Gaussian 03 implementation of the Cu basis set, the s and p functions come from Wachter's optimization for the Cu atom in its ^{2}S state, while the d functions come from Wachter's optimization for the Cu atom in its ^{2}D state

39. Caballol R, Castell O, Illas F, Moreira IDPR, Malrieu JP (1997) J Phys Chem A 101:7860–7866

40. Winkler M (2005) J Phys Chem A 109:1240–1246

41. Lindh R, Bernhardsson A, Schütz M (1999) J Phys Chem A 103:9913–9920

42. Cramer CJ (1999) J Chem Soc Perkin Trans 2:2273–2283

43. Kikuchi A, Ito H, Abe J (2005) J Phys Chem B 109:19448–19453

44. Borden WT (1998) Diradicals. In: Schleyer PVR, Allinger NL, Clark T, Gasteiger J, Kollman PA, Schaeffer HF III (eds) The enciclopedia of computational chemistry. Wiley, Chichester, UK, pp 708–722

45. Poater J, Bickelhaupt FM, Solà M (2007) J Phys Chem A 111:5063–5070

46. Gräfenstein J, Kraka E, Filatov M, Cremer D (2002) Int J Mol Sci 3:360–394

47. Poater A, Ribas X, Llobet A, Cavallo L, Solà M (2008) J Am Chem Soc 130:17710–17717

48. Poater A, Cavallo L (2013) Theor Chem Acc 132:1336–1348

49. Gherman BF, Tolman WB, Cramer CJ (2006) J Comp Chem 27:1950–1961

50. Borden WT, Davidson ER (1996) Acc Chem Res 29:67–75

51. Barone V, Cossi M (1998) J Phys Chem A 102:1995–2001

52. Tomasi J, Persico M (1994) Chem Rev 94:2027–2094

53. Siegbahn PEM (2003) J Biol Inorg Chem 8:577–585

Alkyl mercury compounds: an assessment of DFT methods

**M. Merced Montero-Campillo · Al Mokhtar Lamsabhi ·
Otilia Mó · Manuel Yáñez**

Abstract A series of hybrid and pure density functionals have been assessed in order to provide a reliable description of infrared spectra of alkyl mercury compounds. The theoretical assessment also includes other physicochemical properties, as ionization energies and binding energies. The accuracy of the results was tested through the available experimental data and CCSD(*T*) calculations. B3LYP functional provides the smallest average error of all functionals considering the whole set of properties, and it offers very good results in IR spectra in particular, although other functionals as BP86, M06-L or mPW1PW91 can give even better results when looking at one particular property. Hybrid functionals B3PW91 and X3LYP and double-hybrid B2PLYP follow B3LYP in the list of most well-behaved functionals.

M. M. Montero-Campillo (✉) · A. M. Lamsabhi · O. Mó · M. Yáñez
Departamento de Química, Módulo 13, Universidad Autónoma de Madrid, Campus de Excelencia UAM-CSIC, Cantoblanco, 28049 Madrid, Spain
e-mail: mm.montero@uam.es

1 Introduction

Mercury is a well-known pollutant included in several regulatory laws because of its persistence in the environment, its bioaccumulation and its health effects in human beings [1–6]. Because in most of its compounds, the *d* shell is filled, mercury does not behave from a chemical point of view as a typical transition element. Some of those compounds are the so-called organomercury compounds, characterized by the presence of mercury–carbon bonds. Within this family, we have focused our interest in alkyl mercury compounds, which are highly toxic and unstable [7], their synthesis and characterization presenting several experimental difficulties. A group of alkyl mercury hydrides of general formula RHgH and dialkyl compounds RHgR′ have been obtained in the last years, although many physicochemical properties are still unknown [8–15].

For computational chemists, mercury chemistry is particularly interesting when compared with other neighbor heavy metals (Au) or other elements of group 12 (Zn, Cd) because of its electronic configuration and shares with the first ones the importance of relativistic effects [16, 17]. Several ab initio methods have been used in small mercury compounds, although the fast development of new density functionals in recent times and the consolidation of some of them due to encouraging results in the accuracy of many properties have revealed the convenience of their use, as they represent the best compromise between accuracy and computational cost [18–33]. This last feature is crucial in some cases, in particular the vibrational aspect, since only these quantum chemical methods can be used to study the temporal evolution of these systems through molecular dynamics schemes. Using non-relativistic density functional theory (DFT) methods, relativistic effects can be treated through the use of effective potentials that replace

the mercury core electrons, even though it is also possible to use relativistic Hamiltonians with an appropriate all-electron basis set. As no property that depends on the electron density near the nucleus has to be determined in this work, effective core potentials are perfectly suited to calculate the equilibrium structures of the organomercury compounds of our interest and estimate their infrared spectra, ionization energies and binding energies. The set of molecules RHgH and RHgR′ selected to perform the theoretical assessment was chosen taking into account the available experimental data and their structural similarities with the new mercury compounds that are going to be studied in future works from an experimental and theoretical point of view. The aim of this paper is to test different DFT functionals in order to have a general protocol to obtain reliable results when predicting the IR spectra of these systems as well as other properties.

2 Computational details

All calculations were carried out with Gaussian 09 code [34]. Electronic structure calculations were done with density functionals M06, M06-L, B3LYP, BLYP, LC–BLYP, B2PLYP, X3LYP, B97-D, PBE0, MPW1PW91, B3PW91 and BP86. From these twelve functionals, six of them are hybrid functionals (M06, B3LYP, X3LYP, PBE0, MPW1PW91 and B3PW91), one is a long-range corrected functional (LC–BLYP), and another one is a double-hybrid (B2PLYP). M06 was developed by Truhlar group, specifically for organometallic and inorganometallic chemistry calculations as well as for properly considering non-covalent interactions [35]. M06-L is a pure functional from the same group parameterized for transition metal chemistry [36]. B3LYP is a functional that combines the three-coefficient-dependent hybrid functional for the exchange energy proposed by Becke (B3) with the correlation functional proposed by Lee et al. [37, 38]. This LYP correlation functional can also be combined with the Becke's exchange functional B (BLYP), Xu and Goddard's X3 (X3LYP) or Grimme's B2P (B2PLYP) [39–41]. When adding long-range corrections to BLYP, it results into the LC–BLYP functional, which is highly recommended by Hirao and coworkers to study ionization potentials [42, 43]. Grimme's B97-D is a GGA functional for general chemistry to be used in large systems where dispersion forces could be important [44]. Adamo and Barone's PBE0 is built by mixing 25 % of exact Hartree–Fock exchange with 75 % of GGA exchange functional developed by Perdew et al. [45, 46], the correlation potential still represented by the corresponding functional in PBE. mPW1PW91 combines Perdew–Wang exchange modified by Adamo and Barone with the PW91 correlation functional [47]. PW91

correlation functional can also be combined with Becke's exchange B3 resulting in hybrid B3PW91 [48, 49].[1] Finally, BP86 is a pure functional obtained using Becke's exchange functional B and Perdew's correlation functional P86[2] [50]. This set of functionals allows us to see the response of different types of DFT methods for the same set of mercury compounds.

Different basis sets were tested with the B3LYP functional. In order to properly include the relativistic effects, the mercury active electrons have been represented by different valence basis sets plus small-core effective relativistic potentials as defined in the Gaussian 09 code and the EMSL basis set library (LANL2DZ, MWB60, cc–PVTZ–PP, DEF2–TVZPP, –TVZPPD) [51–62].[3,4,5] The remaining atoms have been defined by all-electron basis sets analogous to those of mercury or triple-zeta Pople basis set as specified in the Sect. 3. No scaling factors were used for harmonic frequency calculations.

3 Results and discussion

Different functionals could be used in principle to work with alkyl mercury compounds. The most used DFT functional in the literature for these systems is B3LYP, although there are several examples in which B3PW91 and BP86 were preferred. Very recent ab initio and DFT studies of mercury–halogen and Hg–OH species have also been reported in which B3PW91 was found to be the preferred option [32, 33]. Mercury hydrides and small saturated dialkyl mercury compounds that have been previously calculated with different ab initio, and DFT methods are HgH_2, MeHgH, MeHgMe and EtHgH (Me = $-CH_3$, Et = $-CH_2CH_3$) [18, 19, 31]. There are also experimental data available from IR spectra for HgH_2, MeHgH, MeHgMe, EtHgH and EtHgEt, and ionization energies for MeHgMe, EtHgMe, EtHgEt, iPrHgMe and PrHgMe (iPr = $-CH-(CH_3)_2$, Pr = $-CH_2-CH_2-CH_3$) [13–15].[6] All these compounds were fully optimized with the different functionals used in the assessment, and their structures and main bond distances are shown in Figure S1 in the

---

[1] See Ref. [37].

[2] See Ref. [39].

[3] MWB is the acronym for "multielectron-fit wood-boring", a quasirelativistic pseudopotential. MWB60 is a small core ECP (60 electrons).

[4] http://www.theochem.uni-stuttgart-de/pseudopotentials/clickpse.html.

[5] See Ref. [54].

[6] Experimental data were obtained from the NIST database on Internet (National Institute of Standards and Technology database: http://webbook.nist.gov/chemistry/). All tables on the results and discussion section contain the bibliographic references for each compound.

Supplementary Material file. As expected, the (atom–Hg–atom) angles are very close to 180°, and bond distances Hg–C and –H have values around 2.1 and 1.6 angstrom, respectively [18].

3.1 Selection of the basis set

All basis set tests were assessed with B3LYP functional and compared with the available infrared spectra experimental data. As a first approximation, the structures were previously optimized with a small double-zeta basis set (LANL2DZ), and then reoptimized with three widely used triple-zeta valence basis set plus small-core pseudopotential found for Hg in the recent literature: Dunning's cc–pVTZ–PP (the same basis set for the remaining atoms), ECPMWB60 (Pople's 6-311G(d,p) for the remaining atoms) and DEF2–TZVPP [6-311G(d,p) for the remaining atoms]. Although the three sets gave almost the same results, DEF2–TZVPP for Hg atom offered the best relation accuracy/computational cost (see Table S1 and additional considerations in the Supplementary Material file). The final structures were calculated adding diffusion functions (DEF2–TVZPPD for Hg and 6-311+G(d,p) for the remaining atoms). The diffusion function for Hg was not reoptimized and has been used as defined by their developers [62]. This larger basis set slightly improves the energy values and was then used to assess the different density functionals, as far as the description of IR spectra, ionization energies and MeHg$^+$ binding energies is concerned.

3.2 Selection of the functional and comparison with experimental and CCSD(T) results

Once the basis set was chosen, different density functionals were used in order to know their capability to predict the IR spectra of those organomercury compounds selected as test molecules. Then, ionization energies and binding energies were calculated with those same functionals.

3.3 Infrared spectra

Table 1 shows the wave numbers of the theoretical spectra as well as the experimental values for HgH$_2$, MeHgH, EtHgH, MeHgMe, EtHgEt, HHgOH and Hg(OH)$_2$ molecules with twelve different functionals (see Computational Details). The results of the table are divided in three groups: Hg–H, –C and –O stretching modes. The Hg–H stretching mode is the most characteristic fingerprint of the spectra in these compounds because of its intensity. As we intend to predict the properties of new organomercury hydrides and their alkyl derivatives, the most simple mercury hydride has to be included in the set (HgH$_2$). Also, HHgOH and Hg(OH)$_2$ were included in order to evaluate

the quality of those same functionals when considering the presence of Hg–O bonds in the set.

For Hg–H bond, all the studied functionals have a relative error with respect to the experimental value under 3.3 %, with the only exception of LC–BLYP. The ability of almost the whole set of functionals to predict the position of this stretching mode is really good and, moreover, three of them, that is, BP86, B3LYP and X3LYP, with negligible errors (0.53, 0.56 and 0.76 %, respectively). Functional M06-L has also a very good behavior (1.50 % of error). Some DFT results are really encouraging, as for example, the HgH$_2$ molecule with BP86 and B3LYP, which are comparable with more sophisticated calculations [68]. Figure 1 illustrates the prediction of BP86 functional for MeHgH, in particular its nearly zero deviation for the Hg–H stretching mode.

However, for the less intense stretching modes associated with the Hg–C and –O bonds, the results are reasonable but not so good (see Table S2 in the Supplementary Material file). For Hg–C in particular, some functionals show spectacular results with errors below 0.5 %, as mPW91PW91 (0.27 %), PBE0 (0.29 %) or B2PLYP (0.44 %). Concerning the Hg–O vibrational mode present in the test, the results in Hg(OH)$_2$ are similar but slightly worst of those obtained for Hg–C bonds, with very small errors for mPW1PW91 (0.54 %), PBE0 (0.54 %), B3PW91 (1.55 %) or B2PLYP (1.55 %). The LYP hybrids that predicted so well the Hg–H mode (B3LYP, X3LYP) or BP86 have much more modest results when looking at Hg–C and –O stretching modes in particular.

Taking a look to the total average error in Table 1, the best functionals for a general description of the IR spectra are mPW1PW91 (1.64 %), PBE0 (1.65 %) and B3PW91 (1.78 %), followed by X3LYP (2.01 %) and B3LYP (2.08 %) because of their ability for describing the main peak of the Hg–H vibration.

3.4 Ionization energies

The ionization energies for MeHgMe, EtHgMe, EtHgEt, iPrHgMe and PrHgMe defined as the energy required to remove one electron from the molecule were calculated with the same 12 density functionals set used in the IR section. One approximation to the ionization energy (IE) value is the energy of the highest occupied molecular orbital (HOMO), the well-known Koopmans' theorem. Within the DFT context, Perdew and co-workers [70] concluded that "the *exact* highest occupied Kohn–Sham eigenvalue is minus the IE of a Z electron system." Of course, the quality of the result obtained with this assumption directly depends on the accuracy of the orbital energies. This option is completely useless in our case, as shown in Table S3 in the Supplementary Material file.

Table 1 Comparison of theoretical harmonic wave numbers of the stretching vibrations of Hg–H, –C and –O bonds with the available experimental data in the selected mercury compounds

	ω_{Exp} (cm^{-1})	B3LYP	BLYP	LC–BLYP	B2PLYP	X3LYP	B97-D
HgH$_2$	1,918.1	1,923.6	1,878.3	2,045.3	1,974.4	1,930.2	1,893.5
	1,912.8						
HHgOH	2,118.2	2,130.9	2,060.2	2,269.1	2,194.2	2,135.0	2,041.6
	2,116.5						
MeHgH	1,953.8	1,969.1	1,906.9	2,114.1	2,024.3	1,974.9	1,923.9
	1,954.9						
EtHgH	1,946.8	1,941.8	1,872.6	2,095.9	1,998.7	1,946.4	1,894.8
MeHgH	530.3	511.6	487.5	570.6	531.0	512.9	490.4
	534.0						
MeHgMe	540.0	519.2	495.0	578.5	537.9	521.9	501.8
EtHgEt	519.0	498.6	473.6	551.8	515.4	500.8	477.6
Hg(OH)$_2$	644.2	616.1	575.8	687.7	630.8	619.1	569.0
	637.3						
Error Hg–H (%)		0.56	2.56	7.42	3.32	0.76	2.18
Error (%)		2.08	5.54	7.27	2.16	2.01	5.26
	ω_{Exp} (cm^{-1})	M06	M06-L	BP86	PBE0	mPW1PW91	B3PW91
HgH$_2$	1,918.1	1,819.7	1,877.3	1,910.9	1,963.2	1,965.2	1,954.1
	1,912.8						
HHgOH	2,118.2	2,119.5	2,131.5	2,098.3	2,167.2	2,167.6	2,156.4
	2,116.5						
MeHgH	1,953.8	1,862.9	1,914.6	1,954.1	2,021.3	2,017.7	2,006.9
	1,954.9						
EtHgH	1,946.8	1,886.1	1,924.7	1,920.1	1,894.8	1,998.9	1,978.6
MeHgH	530.3	507.9	498.9	506.0	533.3	531.4	525.3
	534.0						
MeHgMe	540.0	516.8	505.1	514.2	541.9	539.9	533.8
EtHgEt	519.0	499.3	491.6	489.8	518.4	517.0	510.9
Hg(OH)$_2$	644.2	613.0	612.2	598.4	640.6	639.4	630.8
	637.3						
Error Hg–H (%)		3.24	1.50	0.53	2.75	2.73	2.10
Error (%)		3.73	3.36	2.86	1.65	1.64	1.78

See Refs. [63–68]

As explained in the NIST website, the most precise data in infrared spectroscopic measurements are those obtained in the gas-phase, but many times the chemical properties of the system do not allow their characterization with this technique. The use of matrix isolation sampling in nitrogen or rare gases is a common option. Ideally, the smallest matrix shift occurs for neon matrix measurements: "The generalization that matrix interactions are minimal for neon and that they increase as the mass of the rare gas is increased (…) is supported by experimental observations in larger molecules, as well" [63]. However, due to the presence of very different types of matrices in the experimental data literature for our set of compounds and due to the impossibility of choosing only gas-phase or only neon matrix data, all the available values were used in order to minimize the effect of the matrix

The average relative error for Hg–H bonds and all three types of bonds are shown

A typical hybrid functional as B3LYP gives an average relative error of 26.46 %, while the consideration of long-range corrections can improve that result up to 4.89 % obtained with LC–BLYP. This last value is not good enough, and it is still very far from the CCSD(T) values.

A more accurate solution with DFT is to apply the IE definition itself and calculate the energy difference between the neutral and the radical molecule for each case (see Table 2). The vertical value was used, the radical molecule energy being obtained by a single-point calculation on the optimized neutral molecule geometry. As shown in Table 2, most of functionals yield average relative errors between 2 and 3.5 %. The importance of the long-range correction is highlighted by the good result of
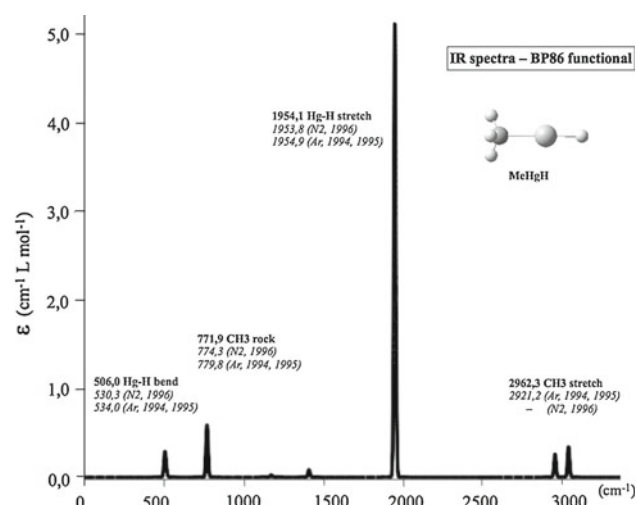
Fig. 1 Theoretical spectra for MeHgH at BP86/DEF2–TVZPPD level. Experimental values are shown between *brackets* [69] (references [9–17] there in)

LC–BLYP functional, with an average error value of 3.04 %, in comparison with its non-corrected equivalent functional BLYP (5.90 %). This corrected functional implemented by Hirao and co-workers is recommended by their authors to estimate ionization energies, a subject that used to be an Achilles heel of DFT calculations.

The lowest average relative errors correspond to hybrid functionals mPW1PW91 and B3PW91 (2.32 and 2.43 %), both containing the Perdew and Wang correlation

functional. Functionals PBE0 and M06 also show good performance. There is an interesting "second group" of functionals including the LYP correlation functional: B2PLYP (2.86 %), LC–BLYP (3.04 %), B3LYP (3.07 %), X3LYP (3.41 %). A quite remarkable feature is the systematic underestimation of the IE values, with the only exception of LC–BLYP. As shown in Table S4 in Supplementary Material file, it worths noting that CCSD(T)// B3LYP level of calculation does not improve the best DFT results for ionization potentials of organomercury compounds when considering the energy difference between neutral and radical forms, while the results of the much more expensive CCSD(T)//CCSD calculations are just slightly under the mPW1PW91 error.

3.5 MeHg$^+$ binding energies

Methyl mercury cation MeHg$^+$ is one of the most interesting alkyl mercury ions, as it has been proved its ability to bind DNA bases, while its reactivity and toxicity have generated many publications in the last years [1–7, 75]. The whole set of functionals was used to predict the binding energies of methyl mercury cation with three neutral bases, namely NH_3, H_2CO and HCN, which can be a good model for large biochemical systems, in which amino groups, carboxylic groups or unsaturated nitrogen atoms are common active sites. The DFT binding energies were compared with those obtained at the CCSD(T) level

Table 2 Theoretical ionization energies (vertical values) with different density functionals compared with the available experimental data

| ΔE (neu-rad)	Exp (eV)	B3LYP	BLYP	LC–BLYP	B2PLYP	X3LYP	B97-D
MeHgMe	9.33	9.04	8.79	9.57	9.03	9.01	8.80
EtHgMe	8.84	8.57	8.28	9.10	8.59	8.54	8.31
EtHgEt	8.45	8.18	7.87	8.75	8.22	8.15	7.92
iPrHgMe	8.48	8.24	7.94	8.73	8.27	8.20	7.98
PrHgMe	8.78	8.50	8.28	9.06	8.52	8.47	8.23
Error (%)		3.07	5.90	3.04	2.86	3.41	6.04
ΔE (neu-rad)	Exp (eV)	M06	M06-L	BP86	PBE0	mPW1PW91	B3PW91
MeHgMe	9.33	9.05	8.89	9.10	9.11	9.11	9.11
EtHgMe	8.84	8.59	8.38	8.57	8.62	8.63	8.63
EtHgEt	8.45	8.20	7.98	8.15	8.23	8.25	8.23
iPrHgMe	8.48	8.25	8.03	8.22	8.28	8.29	8.28
PrHgMe	8.78	8.53	8.33	8.49	8.57	8.58	8.56
Error (%)		2.87	5.18	3.08	2.44	2.32	2.43

See Refs. [71–74]

For this set of compounds, the most recent vertical ionization energy values of the literature from photoelectron spectroscopy technique were chosen (in most of the cases there is only one available value or coincident values)

We calculated for the whole set the adiabatic ionization energies with B3LYP and the results are considerably different from the theoretical vertical values (see Table 2): MeHgMe 8.63 eV, EtHgMe 8.13 eV, EtHgEt 7.74 eV, iPrHgMe 7.75 eV, PrHgMe 8.07 eV. There are significant changes in the geometry of the ionized molecules and the average error is very high (9.04%) with respect to the experimental values shown in Table 2, which are vertical values as explained in Ref. [71–74]

Table 3 Binding energies, BE (kJ/mol) of methyl mercury cation complexes with different density functionals in comparison with CCSD(*T*) results [76]

B. E. (kJ/mol)	CCSD(*T*)	B3LYP	BLYP	LC–BLYP	B2PLYP	X3LYP	B97-D
MeHgNH₃⁺	−215.72	−220.79	−215.31	−275.47	−226.41	−224.41	−220.96
MeHgHCN⁺	−157.66	−156.18	−147.69	−213.49	−161.41	−159.76	−147.43
MeHgH₂CO⁺	−137.02	−143.34	−138.52	−193.92	−146.73	−146.75	−138.34
Error (%)		2.63	2.53	34.88	4.81	4.12	3.29
B. E. (kJ/mol)	CCSD(*T*)	M06	M06-L	BP86	PBE0	mPW1PW91	B3PW91
MeHgNH₃⁺	−215.72	−223.81	−217.41	−231.01	−237.81	−233.06	−227.98
MeHgHCN⁺	−157.66	−164.96	−158.74	−160.75	−171.35	−167.51	−161.94
MeHgH₂CO⁺	−137.02	−144.24	−142.33	−147.08	−153.26	−149.68	−144.83
Error (%)		4.55	1.78	5.46	10.26	7.84	4.70

Basis set superposition error was not calculated as they use to be very small with DFT functionals. See for example [76]

Fig. 2 Average *errors* for the different density functionals
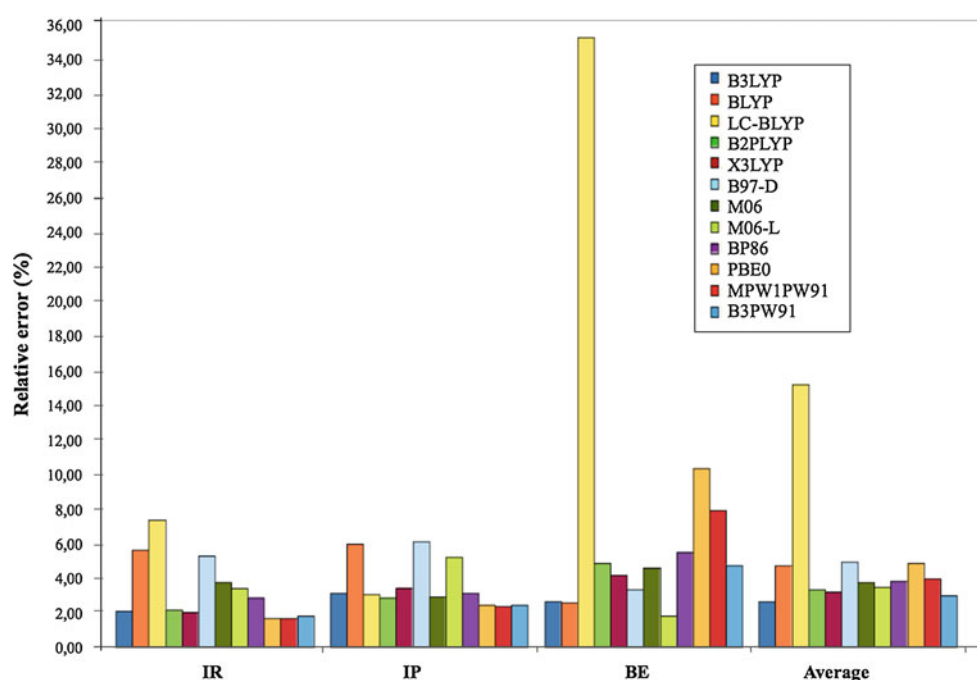
using the same basis set, as shown in Table 3. Binding energies were evaluated by subtracting from the energy of the complex the energy of the monomers. Those CCSD(*T*)/DEF2–TVZPPD reference values are single-point calculations on top of fully optimized CCSD/DEF2–TVZPP geometries.

Truhlar group M06-L functional is the one that better mimics the CCSD(*T*) results, with an average relative error of only 1.78 %. The following best functionals are BLYP and B3LYP, around a 2.5 % of error. This property presents a more irregular behavior than IR or IE data when changing the density functional and comparing with the reference values (although the reference values in the previous cases were experimental measurements). These deviations from the CCSD(*T*) values are only due to the density functional nature in each case and not to significant

geometrical changes, with the exception of the LC–BLYP result. This last functional completely overestimates the strength of bonding between methyl mercury cation and the neutral molecules, presenting the same relative geometrical disposition between atoms but much shorter bond distances than the coupled-cluster optimized structures. The comparison with its BLYP non-corrected related functional justifies this large error.

As mentioned at the beginning of the Sect. 3, B3LYP, BP86 and B3PW91 have been used in the recent literature for mercury compounds. For our alkyl mercury compounds in particular, which is specifically oriented to organomercury hydrides and derivatives, B3LYP and B3PW91 are good choices, while BP86 is particularly recommended to predict Hg–H stretching modes in IR spectra.

Figure 2 (see also Table S5 in Supp. Information file) shows a comparison of the average relative error of each functional in each property (IR, IP and BE) and their total average errors. This comparison reveals some common features and differences at a glance. The most evident result is that there is no one single functional that gives the smallest error in the three properties. On one hand, mPW1PW91 achieves the best predictions for IR spectra and IE values, even though BP86 is particularly good at predicting the Hg–H vibrational stretching frequency; on the other hand, M06-L has the lowest errors for binding energies. Also, as expected because of its design and the applications it was designed for, long-range corrected functionals such as LC–BLYP should be used with caution and only after a careful assessment with the experimental data. Even B3LYP is not the best functional for any one of the three measured properties, it gives the best average error value (2.59 %) of all functionals; that is, it is the "most regular" one, followed by B3PW91 (2.97 %) and X3LYP (3.18 %). It is also remarkable the great difference in some functionals when looking at different properties, as for example PBE0, that gives bad binding energies but very good ionization energies or IR frequencies.

4 Conclusions

DFT is an appropriate methodology to estimate the infrared spectra of alkyl mercury compounds and offers very good results in comparison with experimental data and CCSD(T) calculations. BP86, B3LYP or X3LYP correctly predict the experimental values of Hg–H characteristic stretching mode in the IR spectra with very little error when a DEF2–TVZPPD basis set is used for Hg atom, even though other basis sets provide similar results. Also, density functionals mPW1PW91 and M06-L with the same basis set give good ionization potentials and binding energies for these compounds, respectively. B3LYP shows the most regular behavior in the three studied properties, followed by B3PW91 and X3LYP. It is very interesting the ability of LC–BLYP to provide reasonable ionization energies within the Koopmans' theorem, while the other exchange–correlation functionals do not, and also, it is able to reduce to a half the error of BLYP when using the vertical approximation.

References

1. Atwood DA (ed) (2006) Recent developments in mercury science. Springer, Heidelberg
2. Fitzgerald WF, Lamborg CH, Hammerschmidt CR (2007) Chem Rev 107:641–662
3. Nriagu JO, Pacyna JM (1988) Nature 333:134–139
4. Nriagu JO (1989) Nature 338:47–49
5. Mousavi A, Chavez RD, Ali AMS, Cabaniss SE (2011) Environ Forensics 12:14–18
6. Mahamud S, Ariya PA, Dastoor AP (2011) Atmos Environ 45:5664–5676
7. Sigel A, Sigel H, Sigel RKO (eds) (2010) Metal ions in life sciences, vol 7: organometallics in environment and toxicology. RSC, London, pp 365–401, 403–434
8. Devaud M (1981) J Organomet Chem 220:C27–C29
9. Shayesteh A, Yu S, Bernath PF (2005) J Phys Chem A 109:10280–10286
10. Craig PJ, Mennie D, Needham M, Oshah N, Donard OFX, Martin FJ (1993) J Organomet Chem 447:5–8
11. Craig PJ, Garraud H, Laurie SH, Mennie D, Stojak GH (1994) J Organomet Chem 468:7–11
12. Kwetakat K, Kitching W (1994) J Chem Soc Chem Commun (3):345–347. doi:10.1039/C39940000345
13. Guillermin JC, Bellec N, Kiz-Szetsi S, Nyulaszi L, Veszpremi T (1996) Inorg Chem 35:6586–6591
14. Goubet M, Motiyenko RA, Margulès L, Guillemin JC (2012) J Phys Chem A 116:5405–5409
15. Bellec M, Guillemin JC (1995) Tetrahedron Lett 36:6883–6886
16. Chaudret R, Contreras-García J, Delcey M, Yang W, Parisel O, Piquemal JP Understanding microsolvation of Au(I) and Hg(II) cations: antagonism effects of relativity and correlation (unpublished work in preparation)
17. Arcisauskaite V, Melo JI, Hemmingsen L, Sauer SPA (2011) J Chem Phys 135:044306-1–044306-11
18. Barone V, Bencini A, Totti F, Uytterhoeven MG (1997) Int J Quant Chem 61:361–367
19. Kaupp M, Malkina OL (1998) J Chem Phys 108:3648–3659
20. Lin TP, Wade CR, Pérez LM, Gabbai FP (2010) Angew Chem Int Ed 49:6357–6360
21. Afaneh AT, Schreckenbach G, Wang F (2012) Theor Chem Acc 131:1174–1191
22. Castro L, Dommergue A, Larose C, Ferrari C, Maron L (2011) J Phys Chem A 115:5602–5608
23. Li X, Liao RZ, Zhou W (2010) Chem Phys Chem Chem Phys 12:3961–3971
24. Castro L, Dommergue A, Ferrari C, Maron L (2009) Atmos Environ 43:5708–5711
25. Maron L, Dommergue A, Ferrari C, Delacour-Larose M, Faïn X (2008) Chem Eur J 14:8322–8329
26. Alfonso M, Contreras-García J, Espinosa A, Tárraga A, Molina P (2012) Dalton Trans 41:4437–4444
27. Kretschmer R, Schlangen M, Kaupp M, Schwarz H (2012) Organometallics 31:3816–3824
28. Beck W, Klapötke TM (2008) J Mol Struct (Theochem) 848:94–97
29. Soldán P, Lee EPF, Wright TG (2002) J Phys Chem A 106:8619–8626
30. Sládek V, Lukes V, Ilcin M, Biskupic S (2012) J Comput Chem 33:767–778
31. Cundari TR, Yoshikawa A (1998) J Comput Chem 19:902–911
32. Castro L, Dommergue A, Renard A, Ferrari C, Ramírez-Solís A, Maron L (2011) Phys Chem Chem Phys 13:16772–16779
33. Hernández-Cobos J, Ramírez-Solís A, Maron L, Ortega-Blake I (2012) J Chem Phys 136:014502–014510

34. 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 Jr JA, 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 Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09, revision A1. Gaussian, Wallingford
35. Zhao Y, Truhlar DG (2008) Theor Chem Acc 120:215–241
36. Zhao Y, Truhlar DG (2006) J Chem Phys 125:194101–194118
37. Becke AD (1993) J Chem Phys 98:5648–5652
38. Lee C, Yang W, Parr R (1988) Phys Rev B 37:785–789
39. Becke AD (1988) Phys Rev A 38:3098–3100
40. Xu X, Goddard WA III (2004) Proc Natl Acad Sci USA 101:2673–2677
41. Grimme S (2006) J Chem Phys 124:034108–034124
42. Iikura I, Tsuneda T, Yanai T, Hirao K (2001) J Chem Phys 115:3540–3544
43. Hirao K (2012) Recent advances in LC–DFT. In: Electronic structure theory for strongly correlated systems. Oral Communication, Palermo
44. Grimme S (2006) J Comp Chem 27:1787–1799
45. Adamo C, Barone V (1999) J Chem Phys 110:6158–6169
46. Perdew JP, Burke K, Ernzerhof M (1996) Phys Rev Lett 77:3865–3868
47. Adamo C, Barone V (1998) J Chem Phys 108:664–675
48. Perdew JP (1991) In: Ziesche P, Eschrig H (eds) Electronic structure of solids. Akademie Verlag, Berlin
49. Burke K, Perdew JP, Yang W (1998) In: Dobson JF, Vignale G, Das MP (eds) Electronic density functional theory: recent progress and new directions. Springer, Heidelberg
50. Perdew JP (1986) Phys Rev B 33:8822–8824
51. Feller D (1996) J Comp Chem 17:1571–1586
52. Schuchardt KL, Didier BT, Eisethagen T, Sun L, Gurumoorthi V, Chase J, Li J, Windus TL (2007) J Chem Inf Model 47:1045–1052
53. Dunning Jr TH, Hay PJ (1976) In: Schaefer III HF (ed) Modern theoretical chemistry, vol 3. Plenum, New York
54. Hay PJ, Wadt WR (1985) J Chem Phys 82:270–283
55. Wadt WR, Hay PJ (1985) J Chem Phys 82:284–298
56. Hay PJ, Wadt WR (1985) J Chem Phys 82:299–310
57. Andrae D, Haeussermann U, Dolg M, Stoll H, Preuss H (1990) Theor Chim Acta 77:123–141
58. Martin JML, Sundermann A (2001) J Chem Phys 114:3408–3420
59. Peterson KA, Puzzarini C (2005) Theor Chem Acc 114:283–296
60. Figgen D, Rauhut G, Dolg M, Stoll H (2005) Chem Phys 311:227–244
61. Weigend F, Alrichs R (2005) Phys Chem Chem Phys 7:3297–3305
62. Rappoport D, Furche F (2010) J Chem Phys 133:134105–134116
63. Jacox ME (2002) Chem Soc Rev 31:108–115
64. Wang X, Andrews L (2005) Phys Chem Chem Phys 7:750–759
65. Legay-Sommaire N, Legay F (1996) Chem Phys 211:367–375
66. Greene TM, Andrews L, Downs AJ (1995) J Am Chem Soc 117:8180–8187
67. Shimanouchi T (1972) J Phys Chem Ref Data 6(3):993–1102
68. Wang X, Andrews L (2005) Inorg Chem 44:108
69. Botschwina P, Sebald P, Figgen D, Stoll H (2007) Mol Phys 105:1193–1205
70. Perdew JP, Levy M (1997) Phys Rev B 56:16021–16028
71. Rosenstock HM, Draxl K, Steiner BW, Herron JT (2012) In: Linstrom PJ, Mallard WG (eds) Ion energetics data in NIST chemistry WebBook. NIST standard reference database number 69. National Institute of Standards and Technology, Gaithersburg, MD, 20899. http://webbook.nist.gov. Retrieved 18 Oct 2012
72. Fehlner TP, Ulman J, Nugent WA, Kochi JK (1976) Inorg Chem 15:2544–2548
73. Creber DK, Bancroft GM (1980) Inorg Chem 19:643–648
74. Nugent WA, Wu M.M.-H., Fehlner TP, Kochi JK (1976) J Chem Soc Chem Commun (12):456–457. doi:10.1039/C39760000456
75. Sigel H (1993) Chem Soc Rev 22:255–267
76. Eizaguirre A, Yáñez M, Tortajada J, Salpin J-Y (2008) Chem Phys Lett 464:240–244

REGULAR ARTICLE

On the transferability of fractional contributions to the hydration free energy of amino acids

Josep M. Campanera · Xavier Barril ·
F. Javier Luque

Abstract This study reports the application of the quantum mechanical self-consistent reaction field MST method to compute the solvation profile in water of the twenty natural amino acids. The aim is to derive intrinsic fractional contributions to the hydration free energy and to examine their transferability to peptides. To this end, IEF-MST calculations have been performed at the B3LYP/6-31G(d) level for the series of acetyl amino acid amides, which were chosen as model compounds. In order to account for the flexibility of both the backbone and the side chain in deriving the hydration fractional contributions, calculations have been performed for representative conformers taken from the Dunbrack's backbone-dependent conformational library. The results allow us to dissect the hydration free energy into backbone and side chain contributions and examine the conformational dependence of the fragmental contributions to hydration. For the backbone, different hydration contributions are found for α-helical and β-sheet conformations, which mainly reflect differences in the electrostatic contribution to hydration of the carbonyl group. In contrast, the conformational flexibility of the side chain is found to have little impact on the fractional contribution to hydration. These findings should be valuable to refine semiempirical methods for predicting solvation properties of peptides and proteins in large-scale genomic studies.

Published as part of the special collection of articles derived from the 8th Congress on Electronic Structure: Principles and Applications (ESPA 2012).

Electronic supplementary material The online version of this article (doi:10.1007/s00214-013-1343-y) contains supplementary material, which is available to authorized users.

J. M. Campanera (✉) · F. J. Luque (✉)
Department of Physical Chemistry, Faculty of Pharmacy
and Institute of Biomedicine (IBUB), University of Barcelona,
Campus de l'Alimentació Torribera, Avgda. Prat de la Riba 171,
08921 Santa Coloma de Gramenet, Spain
e-mail: campanera@ub.edu

F. J. Luque
e-mail: fjluque@ub.edu

X. Barril
Department of Physical Chemistry, Faculty of Pharmacy
and Institute of Biomedicine (IBUB), University of Barcelona,
Avgda. Diagonal 643, 08028 Barcelona, Spain

X. Barril
Catalan Institution for Research and Advanced Studies (ICREA),
Passeig Lluís Companys 23, 08010 Barcelona, Spain

1 Introduction

Solvation plays a crucial role in modulating the structure and flexibility of proteins, as well as in mediating their interaction with small ligands and other macromolecular entities. The study of these challenging biological processes and their functional implications is associated with an understanding of the intrinsic solvation properties of amino acid residues. In this context, it is not surprising that many experimental studies have examined the solvation preferences of amino acids in different environments, paying particular attention to solvation in aqueous solution and to the transfer free energies from water to a nonpolar medium [1–9]. Alternatively, computational studies, generally based on the use of discrete treatments of solvation coupled to free energy calculations, have also been used to determine the solvation properties of amino acids [10–13].

The solvation free energy (ΔG_{solv}) measures the reversible work needed to transfer the solute from the gas phase to solution at constant pressure, temperature and concentration [14, 15]. It is convenient to decompose the solvation of a solute into three steps: (1) creation of a solute-shaped cavity in the solvent, (2) switching on the steric properties of the solute inside the cavity and (3) building up of the charge distribution of the solute. This latter term gives rise to the electrostatic component of ΔG_{solv}, whereas the two former contributions are generally grouped into the "steric" or non-electrostatic term. A plethora of theoretical methods relying on different formalisms has been developed to estimate the solvation free energy [16–19]. Among them, quantum mechanical self-consistent reaction field (QM-SCRF) continuum methods have found widespread acceptance for the study of small compounds in solution [20, 21]. The success of these methods can be attributed to the rigorous physical basis of their formalisms, the small increase in computational cost compared to *in vacuo* calculations and the possibility to examine the effect of solvation on the solute properties from the wave function in solution. Finally, it is also worth noting that the most elaborate QM-SCRF methods are capable of estimating the solvation free energy of small (bio)organic compounds with chemical accuracy, as noted in the SAMPL challenges for prediction of hydration free energies [22–29].

The application of these methods to larger systems such as peptides and even proteins is limited by their large size and the complexity of their conformational space. Although the solvation free energy is a property of the entire solute, it is however, useful to partition this quantity into group contributions, since not all the fragments of the solute interact with the same strength with the solvent molecules. Thus, such a partitioning would be valuable to identify chemical groups in the solute that display a prominent contribution to the solvation, to explore the dependence of the solvation free energy on conformational changes that affect the spatial arrangement of fragments or to examine their sensitivity to changes in the local environment [30–32]. This partitioning scheme is implicitly assumed for peptides and proteins, as the total solvation free energy can be expressed as the sum of the solvation contributions from the amino acid residues, which are typically weighted by the ratio of the solvent accessible surface area (SAS) of the residue in the protein and in the fully solvent-exposed free residue [33–42]. In the framework of QM-SCRF methods, partitioning schemes have been developed for the Generalized Born-based SMx model [43, 44], the COSMO-RS method [45, 46] and within the MST model [47–50].

The aim of this study is to determine the fractional contributions of backbone and side chain to the hydration free energy of the twenty amino acid residues using the partitioning strategy implemented in the IEF-MST model [51]. Keeping in mind the wide range of physicochemical properties covered by the natural amino acids and the conformational flexibility of both backbone and side chains, the transferability of these fragmental contributions to the hydration is not fully guaranteed. Accordingly, our first goal is to quantify the contribution of both backbone and side chain to the hydration free energy of amino acids. In addition, the influence of conformation flexibility on the transferability of the fragmental contributions to hydration will also be examined. Finally, the potential application of the fragmental hydration contributions for predicting the total hydration free energy of secondary structural elements will be determined.

2 Methods

2.1 Molecular models and geometries

N-acetyl-ʟ-amino acid amides (CH_3–CO–NH–CHR–CONH–CH_3) were chosen as molecular models to study the hydration contribution of the backbone and the side chain for the twenty natural amino acids. Inclusion of acetyl and methylamine capping groups allow us to build models of amino acids in a context that mimics the local environment of residues in a polypeptide. For the set of acidic (Asp, Glu) and basic (Lys, Arg) amino acids, calculations were performed for both the neutral and ionized species.

In order to examine the conformational dependence on the hydration free energy, calculations were performed for the set of most relevant conformations found in proteins following the backbone-dependent conformational library reported by Dunbrack and Karplus [52, 53]. This library contains the probabilities of side chain conformations (defined by a set of χ angles) as a function of the backbone dihedrals (φ and ψ, based on a grid of 20 × 20°) in 132 protein chains taken from the Brookhaven Protein Database (PDB). For our purposes here, calculations have been performed for all rotamers with a probability contribution higher to 5 % over the total conformational space of a specific amino acid. In the case of Gly and Ala, which are the simplest models of the peptide backbone, we have chosen eight conformations that encompass the torsional angles defined by the regions of the Ramachandran plot associated with α-helical and β-sheet conformations. Within these premises, up to 408 conformations were selected for the whole set of natural amino acids. By using this set of conformations, more than 90 % of the conformational space is covered for 13 amino acids (Gly, Ala, Leu, Ile, Val, Pro, Phe, Tyr, Ser, Thr, Cys, Asn and Asp). For Met, Trp, His, Glu, Lys and Gln, it covers between 50

and 90 % of the rotamers found in proteins. Finally, only for Arg the coverage of the conformational space is lower than 50 %, which reflects the higher flexibility of the side chain for this residue. On average, more than 15 rotamers are used to describe the conformational space of each residue covering either α-helical or β-sheet structures (see Table 1).

The initial structures of N-acetyl-L-amino acid amides were generated automatically by using the obrotate program from OpenBabel suite [54], which permits to rotate the dihedral angle of a specified bond to fit a target value. Then, all the structures were geometrically optimized while keeping the backbone dihedrals fixed to the torsional values of the Dunbrack's library. The geometrical optimization was performed at the B3LYP/6-31G(d) level.

In addition, four pentapeptides that adopt α-helical and β-sheet structures were extracted from selected crystallographic or NMR structures of polypeptides taken from the Protein Data Bank (PDB) [55] and used as models to examine the transferability of the intrinsic hydration free

Table 1 Number of conformers considered for each residue (with a population larger than or equal to 5 %) and distribution between α-helical and β-sheet conformations

Residue	Total no. conformers	α-helix conformers	β-sheet conformers	% Conformational space covered
Gly	8	4	4	99.9
Ala	8	4	4	99.9
Leu	10	5	5	94.3
Ile	9	5	4	96.0
Met	28	12	16	80.3
Val	7	3	4	98.9
Pro	8	4	4	97.9
Phe	12	6	6	91.5
Tyr	14	6	8	96.2
Trp	24	10	14	84.1
His[a]	18	8	10	89.6
Ser	9	4	5	99.0
Thr	7	3	4	97.5
Cys	8	4	4	99.9
Asn	21	8	13	93.9
Gln	23	12	11	72.3
Asp	17	8	9	97.2
Glu	28	16	12	81.7
Arg	25	13	12	42.0
Lys	27	15	12	52.5
Total	408	201	207	–

The percentage of conformational space covered by the selected conformers is also given

[a] HIE, histidine residue protonated in the epsilon position

energy contributions of amino acid residues. The residue segments 27–31 of PDB entry 1SPF (Ile-Val-Gly-Ala-Leu; from NMR [56]) and 38–42 of PDB entry 2P5K (Gln-Ala-Thr-Val-Ser; from X-ray at 1.0 Å resolution [57]) were chosen as α-helical model structures, while segments 265–269 and 37–41 of PDB entries 1R6J (Val-Thr-Ile-Thr-Ile; from X-ray at 0.73 Å resolution [58]) and 3PUC (Thr-Ala-Ile-Trp-Thr; from X-ray at 0.96 Å resolution [59]), respectively, were taken for β-sheet structures. In all cases, the pentapeptides were capped with acetyl and methylamine groups at both N- and C-terminus, and the hydration free energy was determined from IEF-MST B3LYP/6-31G(d) calculations.

2.2 Solvation calculations

The hydration free energy of molecular systems was determined by using the B3LYP/6-31G(d) IEF-MST version [60, 61], which relies on the IEF version of the PCM model [62]. The reader is addressed to Ref. [51] for details about the formalism of the MST model and the parametrization of the electrostatic, cavitation and van der Waals components of ΔG_{solv}. Here, we limit ourselves to describe the essential features of the partitioning scheme implemented in the MST method [38, 39], where the hydration free energy (ΔG_{hyd}) is partitioned into atomic contributions as noted in Eq. 1.

$$\Delta G_{hyd} = \sum_{i=1}^{N} \Delta G_{hyd,i} = \sum_{i=1}^{N} \left(\Delta G_{ele,i} + \Delta G_{cav,i} + \Delta G_{vW,i} \right)$$

(1)

where N is the number of atoms.

Following a perturbative treatment of the solvent response [47], the fractional electrostatic contribution of a given atom is determined from the interaction energy between the whole charge distribution of the molecule and the apparent charges located at the surface elements pertaining to the portion of the cavity generated from that atom (Eq. 2).

$$\Delta G_{ele,i} = \sum_{\substack{j=1 \\ j \in i}}^{M} \left\langle \Psi^o | \frac{1}{2} \frac{q_j^{sol}}{|r_j - r|} | \Psi^o \right\rangle$$

(2)

where Ψ^o denotes the wave function of the solute in the gas phase, and q_j^{sol} stands for the apparent charge created on the surface element j (located at r_j) in response to the fully polarized solute in solution.

The cavitation (ΔG_{cav}) and van der Waals (ΔG_{vW}) terms are easily decomposed into atomic components taking advantage of the relationship with the atomic contribution to the solvent-exposed surface of the solute, as noted in Eqs. 3 and 4.

$$\Delta G_{\mathrm{cav},i} = \frac{S_i}{S_{\mathrm{tot},i}} \Delta G_{P,i} \tag{3}$$

where $\Delta G_{P,i}$ stands for the cavitation free energy of the isolated atom i in Pierotti's formalism [63], S_i is the solvent-exposed surface of such an atom and $S_{\mathrm{tot},i}$ denotes the total surface of the atom.

$$\Delta G_{\mathrm{vW},i} = S_i \xi_i \tag{4}$$

where ξ_i is the atomic surface tension determined by fitting experimental values.

On the basis of Eqs. 2–4, one can easily partition the hydration free energy between contributions due to the atoms pertaining to the backbone and to the side chain, excluding the atoms in the capping groups (acetyl and methylamine moieties), as well as to separate the contribution of apolar and polar fragments of the side chain.

2.3 Computational details

Following the standard parameterization of the IEF-MST model, molecular geometries of the amino acid residues were optimized at the B3LYP/6-31G(d) level, but for the torsional angles that define a given conformation in the backbone-dependent conformational library (see above). The optimized geometries were then kept frozen in MST calculations in water. A similar strategy was adopted for the pentapeptide systems, though the backbone dihedrals were fixed to the values found in the crystallographic structure. All calculations were performed using Gaussian03 package [64]. In addition, besides MST calculations, the hydration free energy of pentapeptides was also determined at both MM-GBSA and MM-PBSA levels for the sake of comparison. These calculations were carried out by mmpbsa.pl script under the AMBER11 suite [65]. Finally, the multivariate partial least squares (PLS [66]) regression technique was used to extract the relevant trends between the hydration free energy and the fragmental contributions following the module implemented in the R statistics package [67].

Due to the large number of calculations required to examine the conformational dependence of hydration properties and the analysis of the huge amount of data derived from atomic contributions to ΔG_{hyd}, Kepler [68, 69] was used to set up a workflow environment to extract the relevant conformers from Dunbrack's library, to build automatically the molecular geometries with fixed torsional angles, to perform QM and QM-SCRF calculations and to store all data in a Mysql database. This workflow is available upon request to the authors. Finally, the library of hydration fractional contributions derived from the partitioning scheme is also available upon request.

3 Results and discussion

3.1 Hydration free energies: comparison with experimental data and other theoretical values

Table 2 reports the side chain contributions to the hydration free energy determined from IEF-MST calculations (weighted by population of each conformer) for the amino acid residues, as well as the experimental data determined from dynamic vapour pressure measurements for a set of molecules that can be considered analogues of amino acid side chains [70]. The involvement of different molecular structures in both IEF-MST calculations and experimental measurements (i.e., the side chain bound to the peptide backbone versus a small molecule that mimics the side chain) makes it necessary to impose some caution regarding the strict comparison of the corresponding hydration values. As an example, let us note that the side chains of small hydrophobic residues (Ala, Val, Leu and Ile) are predicted to be less hydrophobic than the molecular analogues. This difference, however, can be attributed to the partial burial of the side chain atoms by the backbone and, hence, less exposure to water compared to the analogue compounds, reflecting the solvent-excluding effect of the backbone on the fractional hydration of the side chain [38]. In spite of these considerations, it is worth noting that the two sets of data exhibit a nice correlation ($r = 0.96$; Fig. 1). The mean absolute error is close to 1 kcal/mol for a range of hydration free energies of 13.2 kcal/mol (experimental data varying from $+2.3$ kcal/mol for isobutane to -10.9 kcal/mol for N-propylguanidine as mimics of Leu and Arg, respectively). Overall, these findings suggest that the IEF-MST fragmental contributions reflect the influence of the distinctive physicochemical properties of side chain on the hydration of amino acid residues.

Table 2 also reports the hydration free energies determined for analogues of the amino acid side chains by using other theoretical approaches, including either a solvent interaction potential based on quantum mechanically derived charges [71] or free energy calculations coupled to Monte Carlo [38] and molecular dynamics simulations [72]. The agreement between IEF-MST results and the other theoretical values is significant (Fig. 1), as noted in correlation coefficients ranging between 0.92 and 0.98 and average absolute errors varying from 1.2 to 1.9 kcal/mol.

3.2 Backbone/side chain partitioning of the hydration free energy

Table 3 shows the hydration free energy of the 20 N-acetyl-L-amino-acid amides decomposed into the considered backbone and side chain moieties, including the contributions to apolar/polar fragments in the side chain as

well. For the sake of consistency with the molecular model, Table 3 also reports the contributions of the capping (acetyl, methylamine) groups added to mimic the environment of the residue in a peptide chain. For neutral residues, the hydration free energy of the N-acetyl-L-amino-acid amides varies from −10.4 (Pro) and −24.7 (Arg) kcal/mol, whereas the ΔG_{hyd} values of the residues with charged side chains range from −68.0 (Asp) to −83.1 (Lys) kcal/mol.

The partitioning of the hydration free energy shows that the contribution of the capping groups is almost constant for the neutral residues. Their average values amount to −3.6 (acetyl; ΔG_{acetyl}) and −2.5 (methylamine; $\Delta G_{methylamine}$) kcal/mol for methylamine, accounting for 40–50 % of ΔG_{hyd}. This indicates that the contribution of the capping groups is rather independent from the nature of the neutral residue and thus can be considered to be additive. However, this is not valid for ionic residues, as the capping group contribution varies from −7.9 to 1.4 kcal/mol. In fact, the fractional hydration can be realized from the net charge of the residue and the length of the side chain. Thus, the reaction field created by a positive charge tends to favour the hydration of the methylamine group, while a negative charge favours hydration of the acetyl group, this effect being larger for residues with shorter side chains.

By excluding the contribution of the capping groups, the hydration of the neutral residues ranges between −7.3 and −18.5 kcal/mol. The backbone ($\Delta G_{backbone}$) exerts a favourable contribution to the hydration free energy, which on average amounts to −5.1 kcal/mol (see Table 3; Fig. 2). For ionic residues, however, the contribution becomes
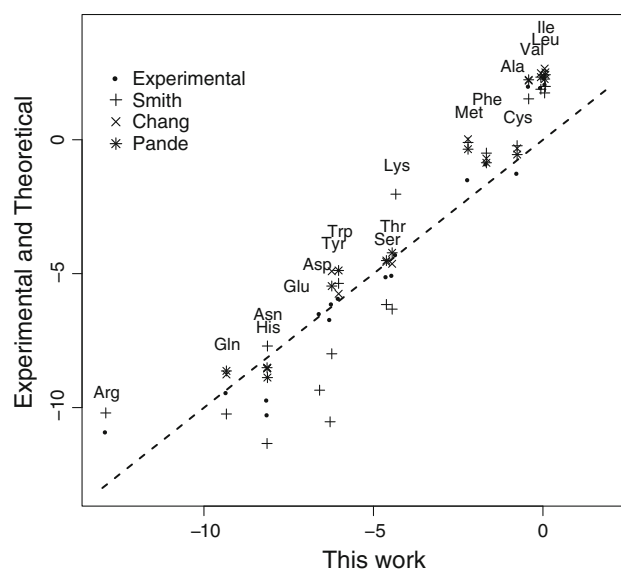

Fig. 1 Correlation between the IEF-MST hydration free energy (x axis) for the side chain of neutral amino acids and experimental [70] or theoretical (Smith [71]; Chang [38]; Pande [72]) values determined for side chain analogues. The *dashed line* represents a perfect regression line with slope unity. Values in kcal/mol

more negative for Asp (−11.9 kcal/mol) and Glu (−10.0 kcal/mol), but more positive for Arg (−3.7 kcal/mol) and Lys (−3.4 kcal/mol). As noted above, these trends reflect the perturbing effect due to the net charge of the side chain. Moreover, as will be discussed later, these trends suggest that the hydration of the backbone is primarily determined by the contribution of the CO group compared to the NH unit. On the other hand, the contribution of the

Table 2 Comparison between the IEF-MST hydration free energy of the side chain and the values determined for side chain analogues from computational and experimental studies

Residue	This work	Side chain analogue	Smith[a]	Chang[b]	Pande[c]	Exp.[d]
Ala	−0.42	Methane	1.53	2.21	2.24	2.01
Leu	0.05	Isobutane	1.75	2.66	2.27	2.32
Ile	0.07	Butane	1.99	2.52	2.43	2.08
Met	−2.21	Methyl ethyl sulphide	−0.10	0.02	−0.35	−1.48
Val	−0.06	Propane	1.89	2.49	2.34	1.96
Phe	−1.67	Toluene	−0.50	−0.70	−0.86	−0.89
Tyr	−6.23	p-cresol	−7.99	−4.91	−5.46	−6.12
Trp	−6.03	Methylindole	−5.36	−5.75	−4.88	−5.91
His	−8.13	Methylimidazole	−7.70	−8.52	−8.88	−10.26
Ser	−4.62	Methanol	−6.15	−4.56	−4.51	−5.10
Thr	−4.45	Ethanol	−6.32	−4.63	−4.22	−5.05
Cys	−0.76	Methanethiol	−0.22	−0.32	−0.55	−1.24
Asn	−8.14	Acetamide	−11.34	−8.54	−8.51	−9.71
Gln	−9.34	Propionamide	−10.24	−8.76	−8.63	−9.43
Asp	−6.28	Acetic acid	−10.53	–	–	−6.70
Glu	−6.59	Propionic acid	−9.35	–	–	−6.48
Lys	−4.34	N-butylamine	−2.03	–	–	−4.28
Arg	−12.90	N-propylguanidine	−10.20	–	–	−10.90

Values in kcal/mol
[a] Ref. [71]; [b] Ref. [38]; [c] Ref. [72]; [d] Ref. [70]

Table 3 Decomposition of the conformation-weighted hydration free energy determined from IEF-MST calculations for the set of *N*-acetyl-L-amino-acid amides

Residues	ΔG_{hyd}	$\Delta G_{backbone}$	$\Delta G_{side\ apolar}$	$\Delta G_{side\ polar}$	ΔG_{acetyl}	$\Delta G_{methylamine}$
Neutral						
Gly	−14.7 (1.7)	−6.7 (0.6)	−0.6 (0.3)	0.0 (0.0)	−4.8 (1.0)	−2.7 (0.9)
Ala	−11.8 (1.5)	−6.0 (1.0)	−0.2 (0.2)	0.0 (0.0)	−3.5 (0.6)	−2.0 (1.1)
Leu	−11.5 (1.2)	−5.3 (1.3)	0.1 (0.2)	0.0 (0.0)	−3.7 (0.3)	−2.6 (0.5)
Ile	−11.1 (1.0)	−4.8 (1.3)	0.1 (0.2)	0.0 (0.0)	−3.8 (0.2)	−2.7 (0.7)
Met	−13.9 (1.4)	−5.3 (1.4)	−0.5 (0.3)	−1.7 (0.2)	−3.7 (0.4)	−2.7 (0.6)
alpha	−15.0 (0.4)	−6.4 (0.1)	−0.7 (0.2)	−1.7 (0.2)	−4.0 (0.1)	−2.2 (0.1)
beta	−12.6 (0.4)	−3.9 (0.6)	−0.2 (0.1)	−1.9 (0.2)	−3.4 (0.3)	−3.2 (0.4)
Val	−11.2 (1.0)	−4.6 (1.3)	−0.1 (0.3)	0.0 (0.0)	−3.8 (0.2)	−2.7 (0.6)
Pro	−10.4 (0.5)	−3.9 (0.5)	−0.7 (0.2)	0.0 (0.0)	−3.6 (0.3)	−2.3 (0.4)
Phe	−13.1 (1.4)	−5.0 (1.3)	−1.7 (0.4)	0.0 (0.0)	−3.8 (0.3)	−2.7 (0.7)
Tyr	−17.5 (1.2)	−4.9 (1.3)	−1.5 (1.7)	−4.7 (0.1)	−3.7 (0.3)	−2.7 (0.6)
alpha	−18.6 (0.6)	−6.4 (0.5)	−1.7 (0.5)	−4.6 (0.2)	−3.9 (0.2)	−2.0 (0.1)
beta	−16.7 (0.5)	−4.0 (0.6)	−1.3 (0.6)	−4.7 (0.1)	−3.6 (0.3)	−3.1 (0.4)
Trp	−17.4 (1.4)	−5.2 (1.4)	−2.4 (0.4)	−3.7 (0.2)	−3.7 (0.4)	−2.4 (0.6)
His	−18.6 (0.9)	−4.7 (0.6)	−0.6 (0.1)	−7.6 (0.2)	−3.5 (0.2)	−2.2 (0.1)
Ser	−16.2 (1.2)	−5.4 (1.0)	−0.4 (0.1)	−4.2 (0.7)	−3.6 (0.4)	−2.6 (0.4)
Thr	−15.5 (2.1)	−4.7 (1.2)	−0.4 (0.1)	−4.1 (0.3)	−3.8 (0.4)	−2.5 (0.8)
Cys	−11.7 (1.4)	−4.6 (1.2)	−0.5 (0.1)	−0.3 (0.2)	−3.4 (0.3)	−2.9 (0.8)
Asn	−19.7 (1.8)	−5.6 (1.6)	−0.4 (0.2)	−7.8 (0.9)	−3.3 (0.3)	−2.6 (0.8)
Gln	−21.1 (1.7)	−5.6 (1.4)	−0.9 (0.3)	−8.5 (0.6)	−3.6 (0.3)	−2.6 (0.7)
alpha	−22.0 (1.2)	−6.4 (0.5)	−1.5 (0.2)	−8.5 (0.7)	−3.8 (0.2)	−2.3 (0.1)
beta	−19.7 (0.8)	−4.3 (0.8)	−0.6 (0.2)	−8.5 (0.5)	−3.3 (0.2)	−3.0 (0.7)
Asp	−17.6 (1.6)	−5.4 (1.2)	−0.6 (0.2)	−5.7 (0.3)	−3.5 (0.3)	−2.4 (0.5)
Glu	−18.5 (1.6)	−5.8 (1.2)	−0.7 (0.3)	−5.9 (0.4)	−3.7 (0.4)	−2.5 (0.5)
Arg	−24.7 (1.5)	−5.6 (1.2)	−0.7 (0.4)	−12.2 (0.5)	−3.7 (0.4)	−2.5 (0.5)
Lys	−16.2 (1.2)	−5.7 (1.2)	−0.1 (0.2)	−4.3 (0.2)	−3.7 (0.4)	−2.4 (0.5)
Charged						
Asp	−68.0 (4.3)	−11.9 (1.6)	−0.3 (0.6)	−48.7 (3.9)	−7.9 (1.3)	0.9 (0.8)
Glu	−73.5 (4.6)	−10.0 (1.7)	−1.6 (1.1)	−56.0 (5.3)	−7.4 (1.1)	1.4 (0.5)
alpha	−73.4 (1.5)	−10.6 (0.9)	−1.5 (0.2)	−55.5 (0.4)	−7.0 (0.2)	1.2 (0.2)
beta	−73.8 (0.6)	−8.8 (0.7)	−1.8 (0.1)	−56.9 (0.3)	−8.0 (0.3)	1.8 (0.5)
Arg	−76.1 (4.5)	−3.7 (1.6)	−9.8 (1.3)	−54.5 (3.1)	−1.8 (0.9)	−6.3 (1.4)
Lys	−83.1 (5.3)	−3.4 (1.4)	−18.8 (1.7)	−53.2 (4.0)	−1.6 (0.9)	−6.1 (1.2)

The conformation-weighted average and the standard deviation (in parenthesis) are shown for each hydration component. For selected residues, the values determined for α-helical and β-sheet conformations are also tabulated. Values in kcal/mol

side chain exhibits the largest variance between residues, thus reflecting the distinct chemical nature of the groups present in the side chain (see Table 3; Fig. 2). In fact, there is an excellent correlation between the fractional contributions of the side chain ($\Delta G_{sidechain}$) and the total hydration free energy of the capped amino acids, as noted in Eq. 5 (see also Fig. 3), where the independent term nicely fits the addition of the average contributions due to the backbone ($\overline{\Delta G_{backbone}}$; −5.1 kcal/mol) and capping acetyl ($\overline{\Delta G_{acetyl}}$; −3.6 kcal/mol) and methylamine ($\overline{\Delta G_{methylamine}}$;

−2.5 kcal/mol) groups. Overall, these findings strongly support the separability of contributions due to both backbone and side chains for neutral residues.

$$\Delta G_{hyd} = 1.00\Delta G_{side\_chain} - 11.5(r = 0.98) \qquad (5)$$

The separability between backbone and side chain contributions is also supported by the results determined for ionized residues (see Fig. 3). Compared to neutral residues, however, one must take into account the different

hydration contribution of the backbone, which is reflected in the shift of the values determined for Asp and Glu from the ideal regression line, as well as in the larger unsigned mean deviation from the dashed line for charged residues (3.3 kcal/mol) compared to the neutral ones (1.2 kcal/mol). This trend is less apparent for Lys and Arg due to the larger length of the side chain.

The effect of the conformational flexibility on the hydration free energy is encoded in the standard deviation of the conformation-weighted fractional contributions. For neutral residues, the standard deviation of the side chain component is generally lower than 0.7 kcal/mol, but for the backbone, it is generally comprised in the range 1.2–1.6 kcal/mol, which indicates a larger conformational dependence of the backbone hydration. In fact, the PLS multivariate analysis of the distinct fractional contributions reveals a distinctive trend between the conformation of the backbone and its hydration, so that the α-helical conformation is found to be better hydrated (by around 2 kcal/mol) than the β-sheet structure (see Fig. 4). This effect is illustrated for Met, Tyr and Gln in Table 3, which reports the values determined for the structures pertaining to α-helical and β-sheet conformations. It is worth noting the drastic reduction in the standard deviation within both subgroups compared with the whole set of conformations for those residues. As noted in previous studies [73], this trend can be realized by the fact that the parallel dipoles of the adjacent peptide groups in α-helices reinforce the interaction with water molecules compared to the
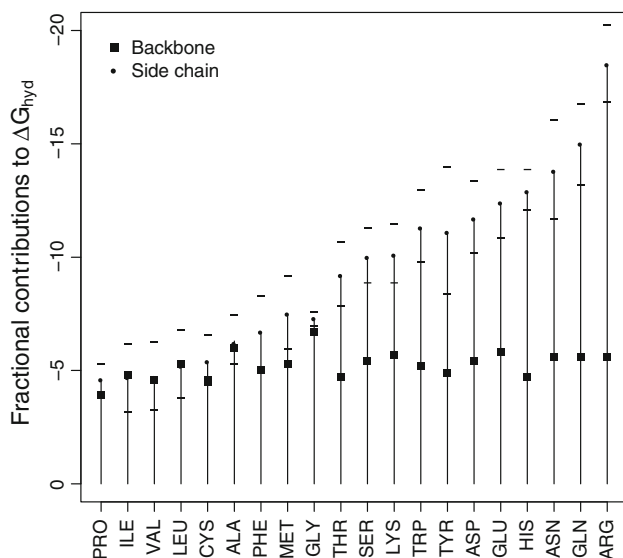

Fig. 3 Representation of the total hydration free energy of residues (corrected by subtracting the average contributions due to backbone and capping groups) versus the fractional contribution of the side chain for the whole set of conformations chosen for neutral and ionized residues. The *dashed line* represents the hypothetical behaviour where backbone, side chain and capping groups show perfect additivity. Values in kcal/mol

antiparallel dipole arrangement found in β-sheets. The other fractional components, particularly the side chain, have a minor effect in modulating the hydration free energy in α-helix and β-sheet structures.

To gain further insight into the backbone hydration, we have extended the PLS analysis to the contributions due to the backbone atoms. The results point out that the dominant component to the hydration, hence, to the conformational dependence of the backbone, comes from the carbonyl oxygen (see Fig. 5). Thus, the fractional hydration of this atom amounts to −3.5 kcal/mol in the α-helical conformation and to −2.5 kcal/mol in the β-sheet arrangement. Then, this single atom accounts for up to 60–70 % of the backbone contribution to the hydration free energy. For the sake of comparison, the average contribution of selected side chain heteroatoms amounts to −1.5 for S in Met, −2.3 for O in Ser, −3.9 for O in Gln, −2.0 for N in Gln and −2.0 for N in His.

In summary, the analysis of the fractional hydration contributions supports the additivity of backbone and side chain contributions for neutral and charged residues, even though it seems necessary to assign different hydration contributions to the backbone of negatively charged residues (Asp, Glu). On the other hand, while the conformational preferences of the side chain are found to have little impact on the hydration contribution, the backbone exhibits a larger sensitivity to the peptide conformation, suggesting the convenience to assign distinct hydration contributions for predicting solvation patterns of α-helical and β-sheet conformations.

Fig. 2 Representation of the fractional contributions to the hydration free energy due to backbone and side chain (polar/apolar) fragments for the set of neutral amino acids. The results are ordered according to the total hydration free energy of the residue. The plot also includes the standard deviation around the total hydration free energy (*tick marks*). Values in kcal/mol
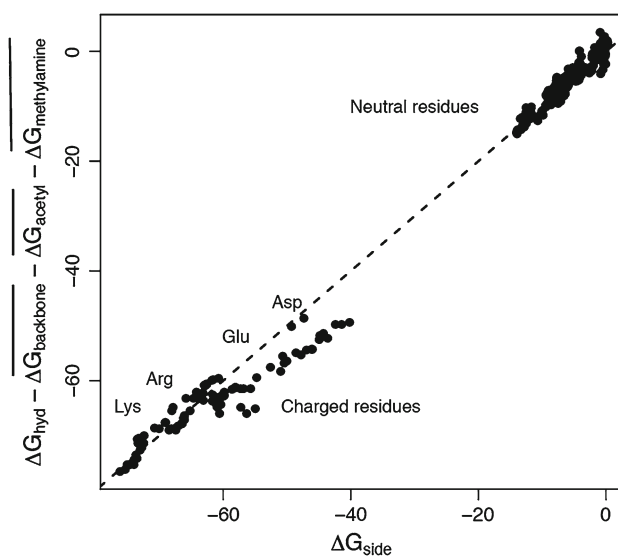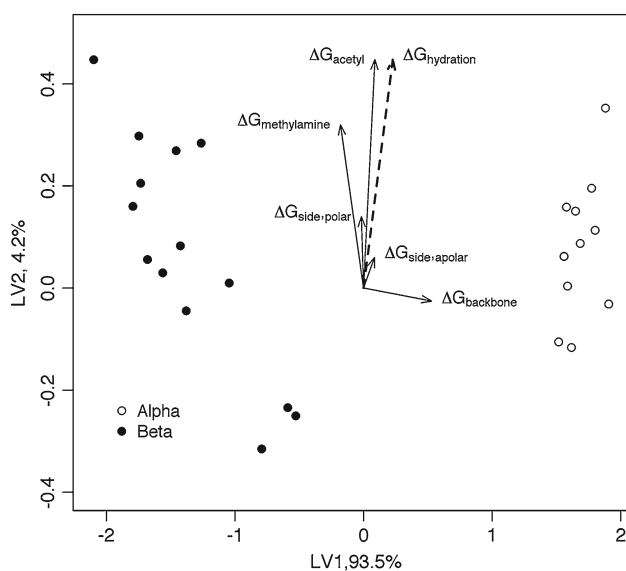
Fig. 4 PLS multivariate analysis between the total hydration free energy (ΔG_{hyd}) and its backbone ($\Delta G_{backbone}$), side chain (including polar, $\Delta G_{side,polar}$ and apolar, $\Delta G_{side,apolar}$, moieties) and capping group (acetyl: ΔG_{acetyl}; methylamine: $\Delta G_{methylamine}$) components for the 28 conformations chosen for *N*-acetyl methionine amide. The plot shows the distribution of the conformers according to the two main latent variables (LV1 and LV2) derived from the PLS analysis, which account for 93.5 and 4.2 % of the variance in hydration free energies. *Dots* denote the conformations and *arrows* correspond to fractional contributions
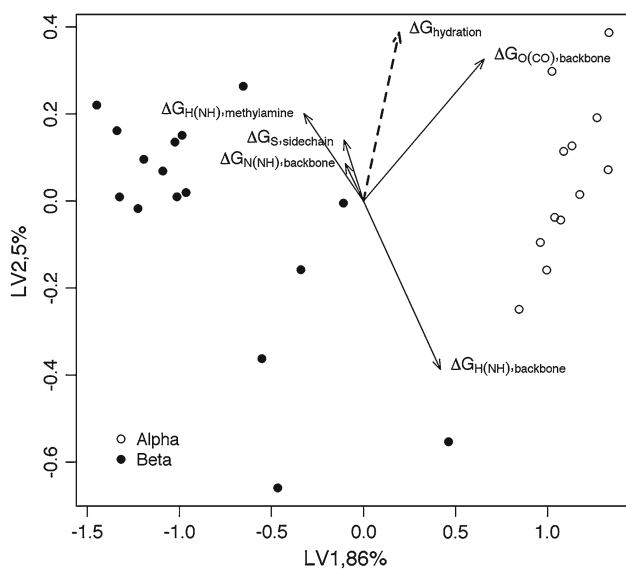
Fig. 5 PLS multivariate analysis between the total hydration free energy (ΔG_{hyd}) and the atomic contributions for the 28 conformations chosen for *N*-acetyl methionine amide. The plot shows the distribution of the conformers according to the two main latent variables (LV1 and LV2) derived from the PLS analysis, which account for 86.0 and 2.5 % of the variance in hydration free energies. *Dots* denote the rotamers and *arrows* correspond to the fractional contributions

3.3 Free energy-type decomposition of the hydration free energy

For computational purposes, it is widely accepted that the solvation free energy can be decomposed into electrostatic and non-electrostatic components. Accordingly, it is convenient to examine the conformation-weighted contributions of electrostatic, cavitation and van der Waals components to the hydration of both backbone and side chain (see Table 4).

For the backbone, the cavitation free energy has a constant contribution of 6.6 kcal/mol, which is compensated by the similar contribution of the van der Waals and electrostatic terms, which have average values of -5.9 and 6.3 kcal/mol, respectively. The electrostatic component accounts for the conformation-dependent hydration of the backbone, as it is found to favour α-helices by 1.8 kcal/mol compared to β-sheets. The different size of the side chain gives rise to a sizable difference in both van der Waals and cavitation contributions. For instance, this latter term varies from 5.2 kcal/mol for Ala to 18.8 kcal/mol for Trp. It is also worth noting that even for polar residues the van der Waals contribution is similar or even larger than the electrostatic one. Furthermore, even though the presence of a net charge makes the electrostatic term to be the dominant component to the side chain hydration, the van der Waals term still has a significant contribution to the hydration (ranging from 16 % for Asp to 31 % for Arg).

For neutral residues, the conformational dependence of the hydration free energy is associated with the electrostatic component of the backbone ($\Delta G_{back,ele}$). This is illustrated for the distinct conformations of Met in Fig. 6, which shows that the discrimination between α-helical and β-sheet conformers is dictated by the $\Delta G_{back,ele}$ component. With regard to ionic residues, the variability introduced by the electrostatic term of the backbone is overcome by the electrostatic contribution of the side chain, which is the main source for the conformational variability of the hydration free energy. In fact, the PLS analysis of the free energy components at the atomic level (data not shown) reveals that the electrostatic contribution of the carbonyl oxygen of the backbone is the main responsible for the distinction between α-helical and β-sheet conformations. Minor contributions to the conformational dependence of hydration are also due to the hydrogen atom of the backbone NH unit and to the heteroatoms in the side chain. The rest of atomic fractional contributions only play a marginal role. Overall, these findings point out that the electrostatic term is implicated in the conformational dependence of hydration, in agreement with previous studies [74, 75] that highlighted the importance of electrostatics in explaining the hydration of peptides.

Table 4 Decomposition of the conformation-weighted hydration contribution of backbone and side chain into electrostatic, cavitation and van der Waals components for the set of N-acetyl-L-amino-acid amides

Residues	ΔG_{hyd}	Backbone			Side chain		
		ΔG_{ele}	ΔG_{vW}^a	ΔG_{cav}^a	ΔG_{ele}	ΔG_{vW}^a	ΔG_{cav}^a
Neutral							
Gly	−7.3 (0.3)	−7.1 (0.6)	−6.7	7.0	−0.6 (0.3)	−1.4	1.5
Ala	−6.2 (1.1)	−6.3 (1.0)	−6.4	6.7	−0.4 (0.1)	−5.0	5.2
Leu	−5.2 (1.5)	−5.6 (1.3)	−6.3	6.6	−0.5 (0.2)	−15.0	15.6
Ile	−4.7 (1.5)	−5.1 (1.3)	−6.2	6.5	−0.5 (0.2)	−14.9	15.5
Met	−7.5 (1.6)	−5.6 (1.4)	−6.3	6.7	−2.7 (0.3)	−14.6	15.0
alpha	−8.8 (0.3)	−6.8 (0.1)	−6.4	6.7	−2.08 (0.2)	−14.6	15.0
beta	−5.9 (0.5)	−4.3 (0.6)	−6.3	6.6	−2.5 (0.1)	−14.6	14.8
Val	−4.7 (1.5)	−5.0 (1.3)	−6.3	6.6	−0.5 (0.3)	−11.7	12.2
Pro	−4.5 (0.7)	−4.6 (0.5)	−4.8	5.5	−1.1 (0.2)	−10.3	10.7
Phe	−6.7 (1.6)	−5.4 (1.3)	−6.3	6.7	−2.0 (0.4)	−16.6	17.0
Tyr	−11.1 (1.5)	−5.2 (1.3)	−6.3	6.7	−6.5 (0.4)	−17.9	18.1
alpha	−12.8 (1.2)	−6.7 (1.0)	−6.4	6.7	−6.6 (0.3)	−17.9	18.1
beta	−10.1 (0.7)	−4.3 (0.6)	−6.3	6.6	−6.4 (0.3)	−17.9	18.1
Trp	−11.3 (1.6)	−5.6 (1.4)	−6.4	6.7	−6.6 (0.3)	−20.4	20.9
His	−12.8 (0.9)	−5.0 (0.6)	−6.3	6.6	−8.6 (0.3)	−12.6	13.1
Ser	−10.0 (1.2)	−5.7 (1.0)	−6.3	6.7	−4.5 (0.8)	−6.5	6.4
Thr	−9.2 (1.4)	−5.1 (1.3)	−6.2	6.6	−4.4 (0.3)	−9.9	9.8
Cys	−5.4 (1.2)	−4.9 (1.2)	−6.3	6.6	−2.5 (0.2)	−6.0	7.8
Asn	−13.8 (2.2)	−6.0 (1.6)	−6.3	6.6	−8.4 (0.9)	−9.6	9.8
Gln	−14.9 (1.8)	−5.9 (1.4)	−6.3	6.7	−9.7 (0.7)	−12.9	13.4
alpha	−16.0 (1.1)	−6.8 (0.5)	−6.4	6.7	−9.9 (0.6)	−12.9	13.4
beta	−13.3 (1.7)	−4.6 (0.8)	−6.2	6.6	−9.4 (0.6)	−12.9	13.4
Arg	−18.5 (1.7)	−5.9 (1.2)	−6.4	6.7	−12.5 (0.8)	−19.8	19.4
Glu	−12.4 (1.5)	−6.1 (1.2)	−6.4	6.7	−7.1 (0.4)	−12.3	12.9
Lys	−10.1 (1.3)	−6.1 (1.2)	−6.4	6.7	−4.5 (0.2)	−17.1	17.4
Asp	−11.6 (1.6)	−5.6 (1.3)	−5.0	5.2	−6.7 (0.5)	−9.0	9.4
Charged							
Asp	−60.9 (5.0)	−12.2 (1.6)	−6.5	6.8	−50.0 (3.9)	−7.9	8.9
Glu	−67.5 (5.2)	−10.3 (1.7)	−6.3	6.7	−58.7 (5.4)	−11.2	12.3
alpha	−67.5 (1.3)	−10.9 (0.9)	−6.3	6.9	−58.2 (0.4)	−11.2	12.3
beta	−67.5 (0.8)	−9.2 (0.7)	−6.3	6.7	−59.8 (0.3)	−11.2	12.4
Arg	−68.0 (5.0)	−4.0 (1.6)	−6.3	6.7	−64.1 (3.3)	−19.8	19.8
Lys	−75.4 (5.8)	−3.7 (1.4)	−6.3	6.7	−72.4 (4.2)	−17.3	17.7

The conformation-weighted average and the standard deviation (in parenthesis) are shown for each hydration component

For selected residues, the values determined for α-helical and β-sheet conformations are also tabulated. Values in kcal/mol

[a] Standard deviation values <0.1 kcal/mol

3.4 Hydration free energy of peptide models

A direct application of the intrinsic *per residue* hydration of amino acids is the prediction of the solvation properties of peptides, since it can be assumed that the solvation of a given residue can be determined by weighting the intrinsic hydration by the fraction of the residue surface that remains exposed to the solvent [41]. This affords a simple but computationally fast approach that might be useful for estimating the solvation properties in structural genomics and other large-scale studies. In this context, the suitability of the IEF-MST fractional hydration contributions has been calibrated by predicting the hydration free energy of four pentapeptides extracted from PDB entries 1R6J, 3PUC, 1SPF and 2P5K (see Sect. 2). They were chosen to encompass α-helical and β-sheet arrangements and
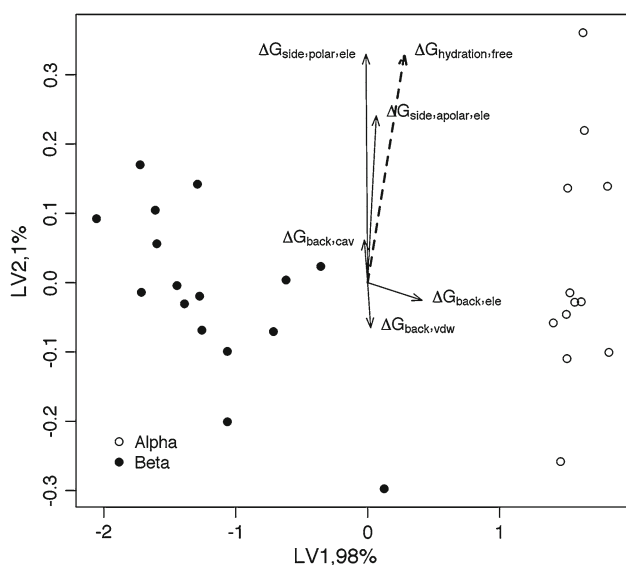
Fig. 6 PLS multivariate analysis between the total hydration free energy (ΔG_{hyd}) and the free energy components (electrostatic, cavitation and van der Waals) for the 28 conformations chosen for *N*-acetyl methionine amide. The plot shows the distribution of the conformers according to the two main latent variables (LV1 and LV2) derived from the PLS analysis, which account for 98.0 and 1.0 % of the variance in hydration free energies. *Dots* denote the rotamers and *arrows* correspond to the fractional contributions

sequences with distinct polar/apolar character. The hydration free energy determined for the four pentapeptides from IEF-MST B3LYP/6-31G(d) calculations amounts to -28.5, -32.6, -28.4 and -43.1 kcal/mol, respectively.

Two additive schemes have been used to estimate the hydration free energy from fractional hydration contributions. In the first case, the hydration free energy was determined by combining the conformation-weighted fractional hydration contributions of both backbone and

$$\lambda_X^i = \frac{S_X^i}{S_X^{i,\text{free}}} \qquad (7)$$

The second approach consists of an atom-based scheme, which exploits the conformation-weighted atomic fractional contributions derived from the set of conformation chosen for residue *i* (Eq. 8).

$$\Delta G_{\text{hyd}} = \sum_{i=1}^{N_{\text{res}}} \sum_{j=1}^{N_{\text{at}}} \zeta_j^i \Delta G_j^{i,\text{free}} \qquad (8)$$

where N_{at} is the number of atoms in the peptide, and ζ_j^i is the fraction of solvent-exposed surface of atom *j* in residue *i*.

The results point out that both fragment-based and atom-based partitioning schemes yield very similar hydration free energies (see Table 5). Compared to the IEF-MST B3LYP/6-31G(d) values, the two fractional schemes slightly underestimate the hydration free energy of the two β-sheet peptides (by 1–3 kcal/mol). For the two α-helix peptides, however, the fractional schemes underestimate the hydration by 9 and 6 kcal/mol for 1SPF and 2P5K, respectively.

In order to realize the different trends found for α-helical and β-sheet peptides, we have compared the free energy contributions to the hydration free energy determined from the IEF-MST partitioning scheme and from the average contributions derived for the *N*-acetyl amino acid amide compounds (see Table 6). For the sake of simplicity, cavitation and van der Waals components are given together, whereas the electrostatic component of both backbone and side chain are given separately. The results show that there is a close agreement for the non-electrostatic (cavitation + van der Waals) components upon scaling of the conformation-weighted atomic contributions of residues by

$$\Delta G_{\text{hyd}} = \sum_{i=1}^{N_{\text{res}}} \left(\lambda_{\text{backbone}}^i \Delta G_{\text{backbone}}^{i,\text{free}} + \lambda_{\substack{\text{sidechain} \\ \text{apolar}}}^i \Delta G_{\substack{\text{sidechain} \\ \text{apolar}}}^{i,\text{free}} + \lambda_{\substack{\text{sidechain} \\ \text{polar}}}^i \Delta G_{\substack{\text{sidechain} \\ \text{polar}}}^{i,\text{free}} \right) \qquad (6)$$

side chain of amino acid residues by the fraction of the solvent-exposed surface of the fragments in the peptide and the free residue (as determined for the *N*-acetyl-L-amino acid amide; Eq. 6).where N_{res} is the number of residues in the peptide, $\Delta G_X^{i,\text{free}}$ stands for the fractional contribution of the fragment (*X*: backbone, apolar side chain, polar side chain), and λ_X^i denotes the fraction of solvent-exposed surface of the fragment (Eq. 7).

the solvent-exposed surface. In fact, the RMSD determined for the separate residues in the peptide models is only 0.2 kcal/mol. It is also worth noting the resemblance between the electrostatic contributions of the side chain to the hydration of the peptides. Thus, comparison of the IEF-MST and fractional contributions for the separate residues in the peptides yields a RMSD of 0.7 kcal/mol. Nevertheless, a distinct trend is found for the backbone electrostatic term:

Table 5 Per residue decomposition of the hydration free energy of selected pentapeptide models

Peptide	IEF-MST	Fragment-based[a]	Atom-based[b]
β-sheet			
1R6J			
Val	−4.0	−3.2	−3.3
Thr	−8.8	−8.0	−8.1
Ile	−3.6	−2.9	−3.3
Thr	−8.4	−7.7	−7.6
Ile	−3.7	−2.9	−3.4
Total	−28.5	−24.6	−25.7
3PUC			
Thr	−8.7	−8.0	−8.0
Ala	−4.1	−3.8	−3.3
Ile	−4.3	−2.6	−2.9
Trp	−9.5	−8.7	−9.6
Thr	−6.0	−7.9	−7.7
Total	−32.6	−31.0	−31.5
α-helix			
1SPF			
Ile	−4.7	−3.3	−2.9
Val	−3.3	−2.4	−2.3
Gly	−6.8	−5.5	−5.3
Ala	−6.9	−5.8	−5.3
Leu	−6.7	−4.7	−4.0
Total	−28.4	−21.6	−19.8
2P5K			
Gln	−16.0	−12.7	−12.4
Ala	−5.3	−4.2	−4.6
Thr	−6.0	−7.2	−6.7
Val	−5.0	−4.1	−4.0
Ser	−10.9	−9.1	−9.4
Total	−43.1	−37.2	−37.0

Values in kcal/mol

[a] Eq. 6; [b] Eq. 8

Table 6 Decomposition of the hydration free energy of selected pentapeptide models

Peptide	Cavitation + van der Waals		Electrostatic (backbone)		Electrostatic (side chain)	
	IEF-MST	Atom-based[a]	IEF-MST	Atom-based	IEF-MST	Atom-based
β-sheet						
1R6J	3.0	2.4	−20.9	−20.6	−10.7	−10.4
3PUC	2.9	2.7	−21.2	−22.8	−14.2	−16.3
α-helix						
1SPF	3.5	3.0	−27.9	−34.1	−4.0	−2.5
2P5K	2.6	2.0	−25.8	−32.8	−20.0	−19.6

Values in kcal/mol

[a] Scaled by the fraction of solvent-exposed surface

and 2 is hydrogen-bonded to the NH unit of residue 5 and capping methylamine, respectively (Fig. 7). For these residues, the deviation between IEF-MST and fractional values of the backbone electrostatic term is, on average, 2.7 kcal/mol, which should be attributed to the shielding of the carbonyl oxygen due to intrahelical hydrogen bonds. In contrast, for residues 3–5, where the carbonyl oxygen is not directly involved in hydrogen bonds, the deviation between IEF-MST and fractional values amounts to only 0.4 kcal/mol. Overall, whereas the scaling by the solvent-exposed surface is well suited for the non-electrostatic term, the inclusion of correction factors that account for the shielding of polar groups through hydrogen bonding might be more advisable for the electrostatic component. In fact, the addition of the surface-scaled non-electrostatic terms, the hydrogen-bonded corrected backbone electrostatic term and the unscaled side chain electrostatic component yield hydration free energies very close to the IEF-MST values, as noted in Table 7.

Finally, let us note that even though MM/PBSA and MM/GBSA calculations lead to hydration free energies that overestimate the IEF-MST values obtained for the pentapeptides (see Table 7), there is a good correspondence between the hydration contributions of residues determined from the conformation-weighted fractional contributions and from MM/PB(GB)SA computations, as noted in the regression equations $y = 1.26x$ ($r = 0.92$) and $y = 1.22x$ ($r = 0.89$), where y stands for the *per residue* MM/PB(GB)SA hydration free energies and x for the fractional ones (see Table S1 and Figure S1 in Supporting Material). Keeping in mind the distinct nature of QM-SCRF and classical calculations, the correspondence found between residue contributions is remarkable and opens the way to further calibrations of the fractional hydration scheme for the widespread application to more complex polypeptides, including the analysis of less populated

while this component nicely fits the IEF-MST values for the β-sheet peptides, it overestimates the contribution for the α-helical ones.

It is worth noting that this effect cannot be relieved by correcting the backbone electrostatic component for the fraction of the solvent-exposed surface in the peptide and in the free N-acetyl amino acid amide, as this approach leads to an underestimation of the scaled values (−20.5 and −20.4 kcal/mol for peptides 1SPF and 2P5K, respectively) with regard to the IEF-MST ones (−27.9 and −25.8 kcal/mol; see Table 6). In fact, this effect reflects the different geometrical arrangement of the backbone in α-helical and β-sheet peptides, and specifically the occlusion of the backbone by the formation of hydrogen bonds in α-helices. At this point, note that the carbonyl oxygen of residues 1

Table 7 Hydration free energy of selected pentapeptide models determined from IEF-MST B3LYP/6-31G(d) calculations, conformation-weighted fractional contributions, MM/GBSA and MM/PBSA methods

Peptide	IEF-MST	Atom-based[a]	MM/GBSA	MM/PBSA
β-sheet				
1R6J	−28.5	−28.6	−37.1	−37.0
3PUC	−32.6	−36.5	−48.4	−47.3
α-helix				
1SPF	−28.4	−28.2	−30.8	−27.4
2P5K	−43.1	−45.1	−56.0	−53.5

Values in kcal/mol

[a] Derived by adding the surface-scaled non-electrostatic terms, the hydrogen-bonded corrected backbone electrostatic term and the unscaled side chain electrostatic component
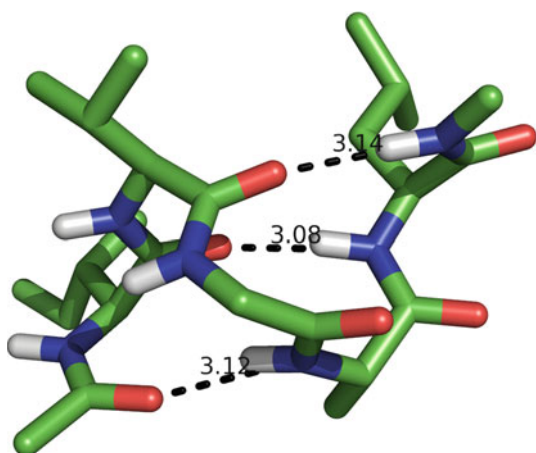
Fig. 7 Representation of the model peptide taken from PDB entry 1SPF. The *dashed lines* represent the hydrogen bonds formed between (1) the carbonyl oxygen of the capping acetyl unit with the NH of residue 4, (2) the carbonyl oxygen of residue 1 with the NH unit of residue 5 and (3) the carbonyl oxygen of residue 2 and the NH unit of capping methylamine

main-chain conformations [76, 77] and nonadditivity effects due to neighbouring residues [78, 79].

4 Conclusions

The IEF-MST formalism has been applied to investigate hydration using a dataset that comprises the most representative backbone-dependent conformational preferences of amino acids and to dissect the hydration free energy into fractional contributions. The results support the additivity of fragmental contributions determined for backbone and side chains for neutral residues. The hydration free energy of the backbone has a dominant contribution to the total hydration energy with an average contribution of −5.1 kcal/mol, though it also has a significant dependence on conformation. Thus, hydration of the backbone in

α-helix is found to be 1 kcal/mol more favourable than in β-sheets, making it convenient to assign distinct fractional contributions to these structures. On the other hand, the different physicochemical properties of the side chain is reflected in hydration free energies of the side chain ranging from 0.1 kcal/mol in Leu or Ile to −12.9 in the neutral Arg. For ionic residues, however, the permanent charge in the side chain leads to nonadditive effects in the fragmental contributions, which makes it necessary to assign specific charge-dependent contributions to the backbone.

The availability of conformational-weighted fragmental contributions at the residue level permits to devise fast strategies for estimating the solvation properties of peptides and proteins in large-scale genomic studies. The approach examined here relies on three main features: (1) the partitioning of the hydration free energy between backbone and side chain, (2) the distinction between fractional contributions for the backbone representative of distinct secondary structure conformations and (3) the correction of the intrinsic hydration components by the fraction of solvent accessibility for the non-electrostatic term and by specific (hydrogen bond) factors for the electrostatic term. The results obtained for a set of pentapeptides are encouraging and support the computational efficiency of this simple strategy compared to MM/PB(GB)SA calculations. Future studies will be performed to evaluate the implementation of additional refinements in the formalism, mainly targeting the electrostatic response in the backbone, and to calibrate the range of applicability of the fractional-based solvation formalism.

References

1. Pliska V, Schmidt M, Fauchere JL (1981) Partition coefficients of amino acids and hydrophobic parameters π of their side-chains as measured by thin-layer chromatography. J Chromatogr 216:79–92
2. Fauchere JL, Pliska V (1983) Hydrophobic parameters-π of amino-acid side-chains from the partitioning of N-acetyl-amino-acid amides. Eur J Med Chem 18:369–375
3. Kim A, Szoka FC (1992) Amino-acid side-chain contributions to free energy of transfer of tripeptides from water to Octanol. Pharm Res 9:504–514
4. Radzicka A, Wolfenden R (1988) Comparing the polarities of the amino acids—side-chain distribution coefficients between the vapor phase cyclohexane, 1-octanol and neutral aqueous solution. Biochemistry 27:1664–1670

5. Radzicka A, Pedersen L, Wolfenden R (1988) Influences of solvent water on protein folding—free energies of solvation of cis and trans peptides are nearly identical. Biochemistry 27: 4538–4541

6. Wimley WC, Creamer TP, White SH (1996) Solvation energies of amino acid side chains and backbone in a family of host-guest pentapeptides. Biochemistry 35:5109–5124

7. Auton M, Bolen DW (2004) Additive transfer free energies of the peptide backbone unit that are independent of the model compound and the choice of concentration scale. Biochemistry 43:1329–1342

8. Shaytan AK, Shaitan KV, Khokhlov AR (2009) Solvent accessible surface area of amino acid residues in globular proteins: correlation of apparent transfer free energies with experimental hydrophobicity scales. Biomacromolecules 10:1224–1237

9. Moon CP, Fleming KG (2011) Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. Proc Natl Acad Sci 108:10174–10177

10. Gu W, Rahi SJ, Helms V (2004) Solvation free energies and transfer free energies for amino acids from hydrophobic solution to water solution from a very simple residue model. J Phys Chem B 108:5806–5814

11. Villa A, Mark AE (2002) Calculation of the free energy of solvation for neutral analogs of amino acid side chains. J Comp Chem 23:548–553

12. Shirts MR, Pitera JW, Swope WC, Pande VS (2003) Extremely precise free energy calculations of amino acid side chain analogs: comparison of common molecular mechanics force fields for proteins. J Chem Phys 119:5740–5761

13. Deng YQ, Roux B (2004) Hydration of amino acid side chains: nonpolar and electrostatic contributions calculated from staged molecular dynamics free energy simulations with explicit water molecules. J Phys Chem B 108:16567–16576

14. Ben-Naim A, Marcus Y (1984) Solvation thermodynamics of non-ionic solutes. J Chem Phys 81:2016–2027

15. Tomasi J, Persico M (1994) Molecular interactions in solution: an overview of methods based on continuous distributions of the solvent. Chem Rev 94:2027–2094

16. Gao J (1996) Hybrid quantum and molecular mechanical simulations: an alternative avenue to solvent effects in organic chemistry. Acc Chem Res 29:298–305

17. Bashford D, Case DA (2000) Generalized born models of macromolecular solvation effects. Ann Rev Phys Chem 51:129–152

18. Orozco M, Luque FJ (2000) Theoretical methods for the description of the solvent effect in biomolecular systems. Chem Rev 100:4187–4226

19. Prabhu N, Sharp K (2006) Protein-solvent interactions. Chem Rev 106:1616–1623

20. Cramer CJ, Truhlar DG (1999) Implicit solvation models: equilibria structure spectra and dynamics. Chem Rev 99:2161–2200

21. Tomasi J, Mennucci B, Cammi R (2005) Quantum mechanical continuum solvation models. Chem Rev 105:2999–3093

22. Guthrie JP (2009) A blind challenge for computational solvation free energies: introduction and overview. J Phys Chem B 113:4501–4507

23. Geballe MT, Skillman AG, Nicholls A, Guthrie JP, Taylor PJ (2010) The SAMPL2 blind prediction challenge: introduction and overview. J Comput Aided Mol Des 24:259–279

24. Soteras I, Forti F, Orozco M, Luque FJ (2009) Performance of the IEF-MST solvation continuum model in a blind test prediction of hydration free energies. J Phys Chem B 113: 9330–9334

25. Marenich AV, Cramer CJ, Truhlar DG (2009) Performance of SM6, SM8, and SMD on the SAMPL1 test set for the prediction of small-molecule solvation free energies. J Phys Chem B 113:4538–4543

26. Klamt A, Eckert F, Diedenhofen M (2009) Prediction of the free energy of hydration of a challenging set of pesticide-like compounds. J Phys Chem B 113:4508–4510

27. Soteras I, Orozco M, Luque FJ (2010) Performance of the IEF-MST solvation continuum model in the SAMPL2 blind test prediction of hydration and tautomerization free energies. J Comput-Aided Mol Des 24:281–291

28. Ribeiro RF, Marenich AV, Cramer CJ, Truhlar DG (2010) Prediction of SAMPL2 aqueous solvation free energies and tautomeric ratios using the SM8, SM8AD, and SMD solvation models. J Comput-Aided Mol Des 24:317–333

29. Klamt A, Diedenhofen M (2010) Blind prediction test of free energies of hydration with COSMO-RS. J Comput Aided Mol Des 24:357–360

30. Jones S, Thornton JM (1997) Analysis of protein–protein interaction sites using surface patches. J Mol Biol 272:121–132

31. Vonheijne G (1994) Membrane-proteins—from sequence to structure. Annu Rev Biophys Biomol Struct 23:167–192

32. Bartlett GJ, Porter CT, Borkakoti N, Thornton JM (2002) Analysis of catalytic residues in enzyme active sites. J Mol Biol 324:105–121

33. Eisenberg D, McLachlan AD (1986) Solvation energy in protein folding and binding. Nature 319:199–203

34. Eisenberg D, Wilcox W, McLachlan AD (1986) Hydrophobicity and amphiphilicity in protein structure. J Cell Biochem 31:11–17

35. Sippl MJ (1993) Boltzmann principle knowledge-based mean fields and protein folding—an approach to the computational determination of protein structures. J Comput Aided Mol Des 7:473–501

36. Jackson RM, Sternberg MJE (1995) A continuum model for protein–protein interactions: application to the docking problem. J Mol Biol 250:258–275

37. Lazaridis T, Karplus M (1999) Discrimination of the native from misfolded protein models with an energy function including implicit solvation. J Mol Biol 288:477–487

38. Chang J, Lenhoff AM, Sandler SI (2007) Solvation free energy of amino acids and side-chain analogues. J Phys Chem B 111:2098–2106

39. Morreale A, de la Cruz X, Meyer T, Gelpi JL, Luque FJ, Orozco M (2005) Partition of protein solvation into group contributions from molecular dynamics simulations. Proteins-Struct Funct Bioinf 58:101–109

40. Morreale A, Gelpi JL, Luque FJ, Orozco M (2003) Continuum and discrete calculation of fractional contributions to solvation free energy. J Comput Chem 24:1610–1623

41. Talavera D, Morreale A, Meyer T, Hospital A, Ferrer-Costa C, Gelpí JL, de la Cruz X, Soliva R, Luque FJ, Orozco M (2006) A fast method for the determination of fractional contributions to solvation in proteins. Prot Sci 15:2525–2533

42. Arab S, Sadeghi M, Eslahchi C, Pezeshk H, Sheari A (2010) A pairwise residue contact area-based mean force potential for discrimination of native protein structure. BMC Bioinf 11:16

43. Giesen DJ, Chambers CC, Cramer CJ, Truhlar DG (1997) What controls partitioning of the nucleic acid bases between chloroform and water? J Phys Chem B 101:5084–5088

44. Hawkins GD, Cramer CJ, Truhlar DG (1998) Universal quantum mechanical model for solvation free energies based on gas-phase geometries. J Phys Chem B 102:3257–3271

45. Hornig M, Klamt A (2005) COSMOfrag: a novel tool for high-throughput ADME property prediction and similarity screening based on quantum chemistry. J Chem Inf Model 45:1169–1177

46. Thormann M, Klamt A, Hornig M, Almstetter M (2006) Cosmosim: bioisosteric similarity based on cosmo-RS σ profiles. J Chem Inf Model 46:1040–1053

47. Luque FJ, Bofill JM, Orozco M (1995) New strategies to incorporate the solvent polarization in self-consistent reaction field and

free-energy perturbation simulations. J Chem Phys 103: 10183–10191

48. Luque FJ, Barril X, Orozco M (1999) Fractional description of free energies of solvation. J Comput Aided Mol Des 13:139–152

49. Muñoz J, Barril X, Hernandez B, Orozco M, Luque FJ (2002) Hydrophobic similarity between molecules: a MST-based hydrophobic similarity index. J Comput Chem 23:554–563

50. Muñoz-Muriedas J, Perspicace S, Bech N, Guccione S, Orozco M, Luque FJ (2005) Hydrophobic molecular similarity from MSDT fractional contributions to the octanol/water partition coefficient. J Comput-Aided Mol Des 19:401–419

51. Luque FJ, Curutchet C, Munoz-Muriedas J, Bidon-Chanal A, Soteras I, Morreale A, Gelpi JL, Orozco M (2003) Continuum solvation models: dissecting the free energy of solvation. J Comput Chem 5:3827–3836

52. Dunbrack JRL, Karplus M (1994) Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains. Nat Struct Biol 1:334–340

53. Dunbrack JRL, Karplus M (1993) Backbone-dependent rotamer library for proteins application to side-chain prediction. J Mol Biol 230:543–574

54. Guha R, Howard MT, Hutchison GR, Murray-Rust P, Rzepa H, Steinbeck C, Wegner J, Willighagen EL (2006) The blue obelisk—interoperability in chemical informatics. J Chem Inf Model 46:991–998

55. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28:235–242

56. Johansson J, Szyperski T, Curstedt T, Wuthrich K (1994) The NMR structure of the pulmonary surfactant-associated polypeptide SP-C in an apolar solvent contains a valyl-rich alpha-helix. Biochemistry 33:6015–6023

57. Garnett JA, Baumberg S, Stockley PG, Phillips SE (2007) A high-resolution structure of the DNA-binding domain or AhrC, the arginine repressor/activator protein fro Bacillus subtilis. Acta Crystallogr Sect F 63:914–917

58. Kang BS, Devedjiev Y, Derewenda U, Derewenda ZS (2004) The PDZ2 domain of syntenin at ultra-high resolution: bridging the gap between macromolecular and small molecule crystallography. J Mol Biol 338:483–493

59. Sauer F, Wilmanns M. doi:10.2210/pdb3puc/pdb

60. Curutchet C, Orozco M, Luque FJ (2001) Solvation in octanol: parametrization of the continuum MST model. J Comput Chem 22:1180–1193

61. Soteras I, Curutchet C, Bidon-Chanal A, Orozco M, Luque FJ (2005) Extension of the MST model to the IEF formalism: HF and B3LYP parametrizations. J Mol Struct Theochem 727:29–40

62. Cancès E, Mennucci B, Tomasi J (1997) A new integral equation formalism for the polarizable continuum model: theoretical background and applications to isotropic and anisotropic dielectrics. J Chem Phys 107:3032–3041

63. Pierotti RA (1976) Scaled particle theory of aqueous and non-aqueous solutions. Chem Rev 76:717–726

64. Gaussian 03 Revision C02 Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JJA, Vreven T, 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, Bakken V, 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 (2004) Gaussian Inc, Wallingford CT

65. Case DA, Darden TA, Cheatham TE, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts BP, Wang B, Hayik S, Roitberg A, Seabra G, Kolossváry I, Wong KF, Paesani F, Vanicek J, Liu J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Hsieh HJ, Cui G, Roe DR, Mathews DH, Seetin MG, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A, Kollman PA (2010) AMBER 11

66. Massart DL, Vandeginste BGM, Buydens LMC, De Jong S, Lewi PJ, Smeyers-Verbeke J (1998) In handbook of chemometrics and qualimetrics. Elsevier Science, Amsterdam

67. R-Development Core Team (2011) R: a language and environment for statistical computing. http://www.r-project.org (accessed 2011)

68. Deelman E, Gannon D, Shields M, Taylor I (2009) Workflows and e-science: an overview of workflow system features and capabilities. Future Gener Comput Syst 25:528–540

69. Altintas I, Berkley C, Jaeger E, Jones M, Ludäscher B, Mock S (2004) In Kepler: an extensible system for design and execution of scientific workflows. In: Proceedings of the international conference on scientific and statistical database management SSDBM, vol 16, pp 423–424

70. Wolfenden R, Andersson L, Cullis PM, Southgate CCB (1981) Affinities of amino acid side chains for solvent water. Biochemistry 20:849–855

71. Smith BJ (1999) Solvation parameters for amino acids. J Comput Chem 20:428–442

72. Shirts MR, Pande VS (2005) Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. J Chem Phys 122:134508

73. Avbelj F (2000) Amino acid conformational preferences and solvation of polar backbone atoms in peptides and proteins. J Mol Biol 300:1335–1359

74. Lin B, Pettitt BM (2011) Electrostatic solvation free energy of amino acid side chain analogs: implications for the validity of electrostatic linear response in water. J Comput Chem 32: 878–885

75. Baldwin RL (2002) Relation between peptide backbone solvation and the energetics of peptide hydrogen bonds. Biophys Chem 101:203–210

76. Gould IR, Cornell WD, Hillier IH (1994) A quantum mechanical investigation of the conformational energetics of the alanine and glycine dipeptides in the gas phase and in aqueous solution. J Am Chem Soc 116:9250–9256

77. Beachy MD, Chasman D, Murphy RB, Halgren TA, Friesner RA (1997) Accurate ab initio quantum chemical determination of the relative energetics of peptide conformations and assessment of empirical force fields. J Am Chem Soc 119:5908–5920

78. Staritzbichler R, Gu W, Helms V (2005) Are solvation free energies of homogeneous helical peptides additive? J Phys Chem B 109:19000–19007

79. Avbelj F, Baldwin RL (2009) Origin of the change in solvation enthalpy of the peptide group when neighboring peptide groups are added. Proc Natl Acad Sci USA 106:3137–3141

REGULAR ARTICLE

A time-dependent DFT/molecular dynamics study of the proton-wire responsible for the red fluorescence in the LSSmKate2 protein

**Carlos Randino · Marc Nadal-Ferret ·
Ricard Gelabert · Miquel Moreno ·
José M. Lluch**

Abstract Fluorescent proteins (FP) have become a major topic in the recent biochemical research due to their applications as in vivo markers in biological systems. In particular, Red fluorescent proteins (RFP) present some advantages since they require less harmful radiations to be excited and show less light-scattering. In this paper, we are focusing on the LSSmKate2 protein, a RFP that, together with LSSmKate1 and mKeima, is well known for the outstanding difference between absorption and emission wavelengths, which is usually referred as Large Stokes Shift (LSS). It is commonly accepted that an excited state proton transfer accounts for the fluorescence observed in the three proteins. In this work, a molecular dynamics simulation of the LSSmKate2 protein has been carried out, and from different snapshots, a series of excited states have been calculated and analyzed. Our molecular dynamics simulation has proved the availability of the two-link proton-wire suggested by Piatkevich et al. and has furnished a new one-link relay, more prone to take place. The statistical treatment of the excited states can reproduce the electronic absorption spectrum in a reasonable way, and the analysis of the involved orbitals confirms that one absorption wavelength maximum corresponds to an acidification of the chromophore, regardless of the hydrogen-bonded acceptor residue. All this work constitutes an important step in what should be a thorough and complete study of the photochemistry of the LSSproteins.

1 Introduction

Since the extraction, purification and characterization of Green Fluorescent Protein (GFP) by Shimomura from the jellyfish *Aequorea victoria* in the early 1960s [1, 2], a wide range of natural and artificial fluorescent proteins (FP) have appeared [3], contributing to add new improvements for these in vivo markers in biological systems [4–7]. Many theoretical studies have appeared in recent years due to the interest of fluorescent proteins [8–21]. Among these, the red fluorescent proteins (RFPs) are particularly interesting due to their better biochemical features, suitable to minimize some drawbacks caused by the use of energetic and dangerous wavelengths in living organisms. Moreover, the RFPs reduce the autofluorescence on the surrounding cells and show deeper light penetration and lower light-scattering [22].

In 2010, two new red fluorescent proteins, LSSmKate1 and LSSmKate2, were developed by Piatkevich et al. [23, 24] from the far-red fluorescent protein mKate. LSSmKate1 and LSSmKate2 differ from each other in the amino acid sequence at position 160 (Glu160 for LSSmKate1 and Asp160 for LSSmKate2) and in the conformation of the chromophore (*cis* for LSSmKate1 and *trans* for LSSmKate2). Both proteins

C. Randino · M. Nadal-Ferret · R. Gelabert (✉) · M. Moreno · J. M. Lluch
Departament de Química, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain
e-mail: Ricard.Gelabert@uab.cat

J. M. Lluch
Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

share the special property of showing a big gap between absorption and emission wavelengths, commonly named in the field of biology large stokes shift (LSS). This feature allows the possibility of obtaining a multicolor image useful to select and distinguish different parts of biological systems. LSSmKate1 and LSSmKate2 present absorption maxima at 463/460 nm and emission maxima at 624/605 nm, respectively.

Because of the similarities that LSSmKate1 and LSSmKate2 share with GFP, the suggested photochemical process could follow a similar scheme. For both proteins, the mechanism involves two species: the protonated and deprotonated chromophore. The protonated chromophore absorbs light and starts a proton transfer reaction. Based on their X-ray structures, a one-link excited state proton transfer (ESPT) for LSSmKate1 and a two-link ESPT for LSSmKate2 are proposed. LSSmKate1 would support a direct proton migration from the hydroxyl group of the chromophore to the carboxylic group of Glu160. For LSSmKate2 (Fig. 1), the first proton moves from the hydroxyl group of the chromophore to the hydroxyl group of Ser158 and the second one from Ser158 to the carboxylic group of Asp160. Once the different proton transfers occur, the fluorescence in the deprotonated chromophores takes place. Our group's research interest is focused on the multiple proton transfer reactions [11–13, 25], and so in this paper, we will center on LSSmKate2 which has been proposed to operate via a two-step proton-wire.

Due to the size and the wide mobility of this kind of biochemical systems, the study becomes highly complex. Focusing only on the available crystallographic structure forces us to work with a simple and static model of the whole protein and therefore, restricting us to one small and scarce representation, representative only of the solid state of the crystal structure. However, the protein as encountered naturally is not in the solid phase: it is surrounded by
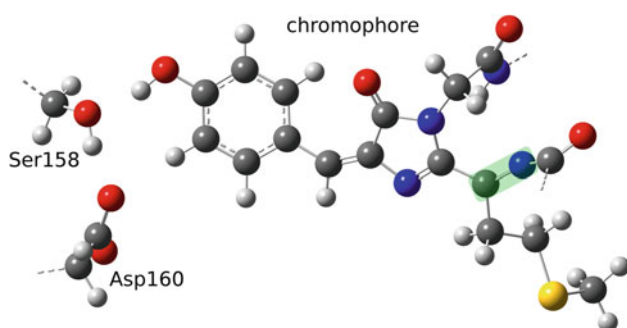
solvent (water) and ions and at some finite temperature. In these conditions, the protein is a very dynamic entity, flexible and that jostles around due to thermal motion. To understand the protein's properties in these conditions, it is necessary to sample a number of configurations of it, obtained through for instance, molecular dynamics. To achieve this aim, we will carry out a molecular dynamics simulation using classical mechanics. So the main purpose of this work is to provide this variety of geometries from a theoretical point of view, including a suitable environment. With this information, we will simulate the electronic absorption spectrum for LSSmKate2, and we will analyze along the molecular dynamics the thermal fluctuations of the three residues directly involved in the two proton transfers responsible for the fluorescence.

2 Methods

2.1 Parameterization and classical dynamics simulation

A 40-ns molecular (classical) dynamics (MD) simulation has been run with a dual aim: on the one hand, to check the dynamical stability of the crystallographic structure [24] supplied at the Protein Data Bank, and on the other hand, to furnish an ensemble of structures from which further excited state electronic calculations have been done, thus providing a tool to introduce thermal effects in the spectrum, as explained in the following subsection.

To properly run the MD simulation, the crystallographic structure has had to undergo some refinements. First of all, as hydrogen atoms are not detected in the X-ray structure, they have been included by using the PDB2PQR [26, 27] server at a neutral pH. As for the chromophore, which is not a standard amino acid, it has been protonated as Piatkevich et al. [24] suggest. Secondly, crystallographic structures are obtained at very low temperatures and do not contain bulk solvent molecules. Physiological conditions imply the inclusion of the solvent effects and require heating and equilibrating the system to physiological temperatures before running the MD simulation. All these processes (including the MD simulation) have been performed using the CHARMM22 [28–30] force field by means of the CHARMM-35b1 [31, 32] software. In all, 34952 atoms have been included in the molecular dynamics simulation.

Nevertheless, before tackling these very steps, another delicate issue has had to be dealt with, namely, the fact that the chromophore is not a standard residue, new parameters are therefore needed. Since the LSSmKate2 chromophore is similar to the GFP one, [33] most of the necessary parameters have been taken from the parameterization done for GFP's by Thiel and co-workers, [34] consistent

Fig. 1 Representation of the chromophore and the acceptor amino acids involved in the ESPT in LSSmKate2, Ser158 and Asp160. The C and N atoms, highlighted in *green*, have had to be reparametrized, as their likes are not present in the standard CHARMM force field. Color code: *dark gray*, carbon; *white*, hydrogen; *red*, oxygen; *blue*, nitrogen; *yellow*, sulfur

with the CHARMM22 force field. Compared to GFP's, the LSSmKate2 chromophore has a methionine instead of a serine, so for these very atoms, the CHARMM22 force field parameters have been taken. However, there are two atoms that differ from any other amino acid: the nitrogen and carbon atoms that form a double bond (Fig. 1, highlighted in green). Parameters for all stretchings, bendings, torsions and improper torsions, partial charges and Van Der Waals interactions involving these two atoms have had to be found. For the sake of consistency, we have tried to follow Thiel's method [34] as thoroughly as possible, although we have calculated each parameter separately from the rest (except for the nonbonding and partial charge), as explained elsewhere [35].

The system has been solvated with a cubic box of water molecules of 69 Å edge, and periodic boundary conditions have been set to account for the long-range electrostatic interactions of the protein. Two rough energy minimizations (each one with 200 steps of Adopted Basis Newton–Raphson method) have been performed to avoid bad contacts, firstly involving just the water molecules, and then the whole system. Next, the system must be heated and equilibrated. The heating and equilibration have been carried out using the CHARMM-implemented Verlet integrator algorithm with steps of 1 fs. The temperature has been increased 25 K every 5 ps. This way the system is softly heated from 0 to 300 K. The periodic boundary conditions imply the construction of a cubic lattice, in which the protein moves freely and the potential energy is calculated with a 20 Å cutoff. The motion of the backbone is constrained by harmonic forces. When the system has reached 300 K, the constrained dynamics simulation is restarted with the same integration step for 160 ps, without further increase of temperature and gradually releasing all the harmonic constraints applied on the system. Afterward, 360 ps more has been run without any harmonic constraints. Once the system is fully equilibrated, the MD simulation is run for 40 ns production time.

2.2 Theoretical simulation of the electronic absorption spectrum

Excited state electronic structure calculations have been performed on 160 snapshots from the ensemble of structures obtained from the molecular dynamics simulation (one every 0.25 ns). The calculated absorption wavelengths constitute a fairly good approximation to the experimental absorption spectrum as it takes into account the thermal effects.

The electronic calculations have been done as follows: each structure resulting from the MD simulation has been divided in two regions, namely, a quantum mechanics (QM) region and a molecular mechanics (MM) region.

The QM region includes the chromophore in its entirety with the carbonyl group of the previous amino acid (Phe60) and the amino group of the next amino acid (Ser66). The side chain of Ser158 and a deprotonated Glu160, the two responsible amino acids participating in the ESPT, are also included (in total, 53 QM atoms + 4 link atoms). All the QM atoms of the model are represented as spheres in Fig. 1. The rest of the protein has been treated as MM region to introduce polarization effects on the QM part. The polarization of the protein due to the chromophore is not taken into account, but we think that this effect will not be very significant in the spectrum, see for instance the work by Murugan et al. [36] The QM region has been calculated quantum-mechanically using the Gaussian 09 [37] software, and the MM atoms have been treated as point charges according to the CHARMM-22 force field. We have used the Chemshell [38] package to prepare the inputs. Link atoms have been used to treat the QM/MM boundary.

The quantum mechanical calculations have been done with the density functional theory (DFT) for the ground state, and the time-dependent density functional theory (TDDFT) for the excited states, all of them with the CAM-B3LYP [39] functional and the 6–31 + G(d,p) basis set. TDDFT is considered a proper method to tackle the analysis of an ESPT, as it has been stated in published works about other chromophores [40, 41]. The alleged involvement of charge-transfer excited states in LSSmKate2 [42, 43] fully justifies the election of the hybrid Coulomb-attenuated CAM-B3LYP, specifically designed to avoid the spurious effects resulting from the inability of pure DFT methods to properly describe long-range exchange effects [39, 44]. As a matter of fact, we have recently shown that CAM-B3LYP provides good results for LSSmKate2, specially compared with conventional hybrid functional such as B3LYP [43].

15 excited states have been calculated for each of the 160 frames. Then, all the different excited states with oscillator strength (f) greater than 0.1 have been sorted by their frequencies, and a wavelength relative histogram has been done with a bin size of 10 nm. This histogram represents an average oscillator strength for each 10 nm wavelength window, weighted by the number of snapshots which have an excited state accessible ($f \geq 0.1$) through excitation at the wavelength range from the ground state. The sum of the counted oscillator strengths around a given wavelength value gives an approximate idea of how likely it is to detect absorption around it regardless of the nature of the excited state involved. It is not a thorough reproduction of the spectrum, but it is a feasible and plausible way to simulate it. A similar procedure has already been used by Thiel and co-workers [19] and other groups [20, 45] to characterize the absorption spectrum of different fluorescent proteins. Other approaches have also been

applied to obtain the optical properties of varied chemical systems [36, 46–49].

3 Results

3.1 Classical dynamics simulation

The 40-ns MD simulation has yielded interesting results to figure out possible paths available for the excited state proton transfer. To begin with, it is seen that the protein is stable along the dynamics simulation, that is to say, there are no relevant conformational changes. As for the chromophore and the residues that are proposed to host the proton transfer, namely, Ser158 and Asp160, several hydrogen bonds are observed. By analyzing the distances between the oxygen atoms of Ser158 and Asp160 (Fig. 2a) and the distances between the hydrogen atom of Ser158

and the oxygen atom of Asp160 (Fig. 2a), several features can be noticed.

First of all, leaving apart some pronounced and random fluctuations that appear occasionally, the range of the oscillations is small (about 0.5 Å, slightly more for the hydrogen–oxygen distance). Besides, the distance between the oxygen atoms oscillates between 2.5 and 3 Å, and the distance between the oxygen and the hydrogen atoms oscillates between 1.5 and 2 Å. This constant difference of 1 Å coincides with the average hydrogen chemical bond distance, which implies that the three atoms are almost collinear and the hydrogen bond is well formed and is maintained along the MD simulation. Following the same criteria, we can also indentify in Fig. 2b, c when the chromophore is hydrogen bonded either to Ser158 or Asp160. On the contrary, the corresponding hydrogen bond does not exist when the range of oscillation of the oxygen–oxygen distance is about 1 Å, while the oxygen–oxygen distance is

Fig. 2 Distances between the residues involving the hydrogen bonds in the putative proton-wire along the MD simulation, taking into account the oxygen–oxygen distances (*left*) or the oxygen-hydrogen distances (*right*) for Ser158-Asp160 **a**, Cro-Ser158 **b** and Cro-Asp160 **c** hydrogen bonds
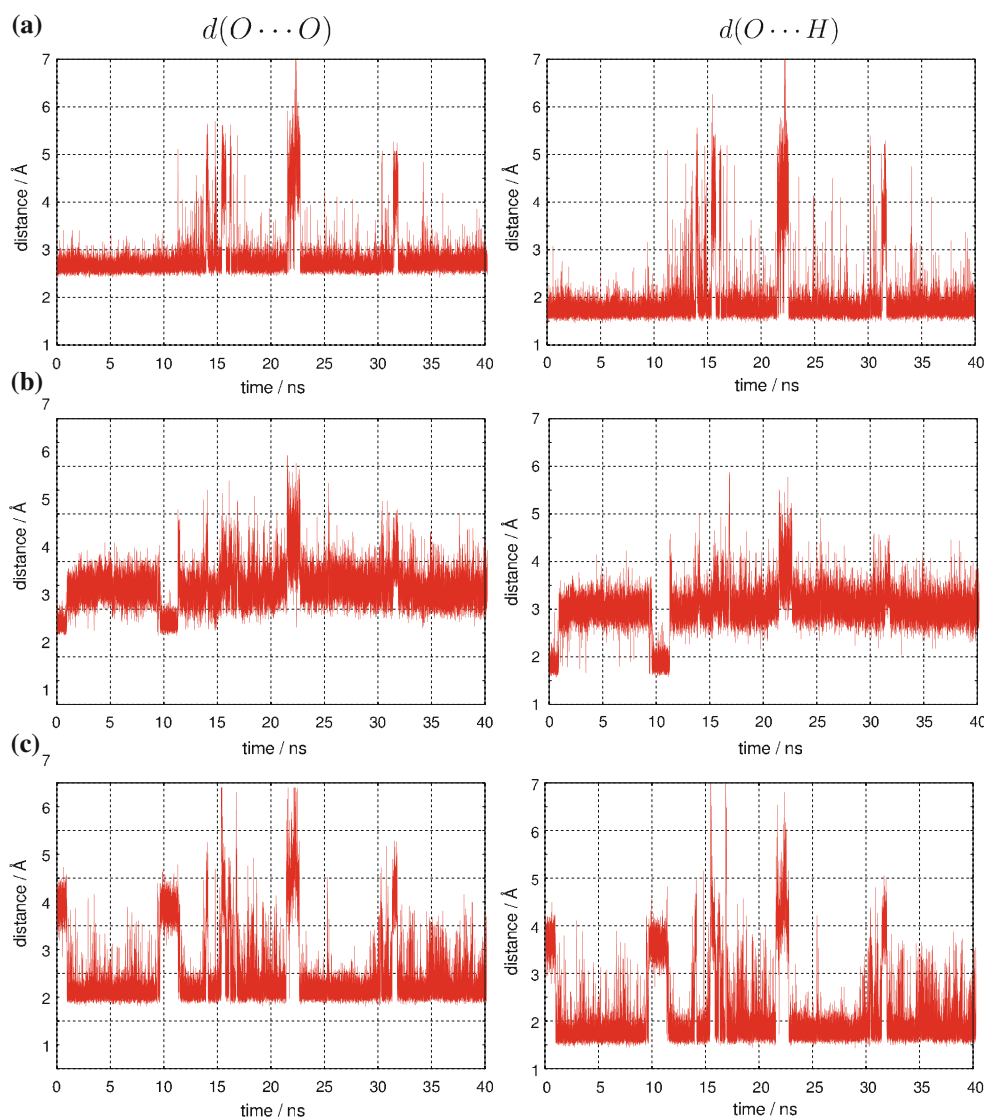
beyond 3 Å (Ser158) or 4 Å (Asp160). Then, except for short temporary hydrogen bond disruptions, the chromophore is always hydrogen bonded to either to Ser158 or Asp160, but never to both simultaneously.

Considering all that, we can identify two kinds of hydrogen bond networks between the other residues that constitute two qualitatively different situations. Firstly, one in which, as it appears in the crystallographic structure, the chromophore is hydrogen bonded to Ser158 and Ser158 is hydrogen bonded to Asp160, thus enabling a putative two-link proton relay. This situation, as can be observed in Fig. 2a, b, is not the most probable in terms of residence time as it only occurs between 0 and 1.5 ns, and between 9.5 and 11.5 ns approximately. Secondly, another situation can be observed, in which the chromophore is hydrogen bonded to Asp160, which is also hydrogen bonded to Ser158 (Fig. 2a, c). This is the most likely situation in terms of residence time, as it happens between 1.5 and 9.5 ns, and from 11.5 ns to the end of the simulation. Nonetheless, from about 14–17 ns, about 22–23 ns and about 31–32 ns, all the distances shown in Fig. 2 increase. This corresponds to temporary disruptions of the hydrogen bonds, and in these periods, the above-mentioned residues are hydrogen bonded to other residues (mainly water molecules), but they always return to the latter situation, in which the chromophore is hydrogen bonded to Asp160 despite some spurious peaks. This latter configuration would enable a single link proton relay, in which the proton could be transferred directly from the chromophore to Asp160, analogously to what is suggested to occur for LSSmKate1, [24] where the phenoxy moiety proton in the chromophore is thought to be directly transferred to Glu160 (already in the X-ray structure).

As it has been checked by the analyses of the excited state electronic structure calculations presented in the following subsection, there are no appreciable differences between those two different configurations in terms of the nature of excited states found for each of them. Furthermore, both configurations would eventually lead to the same chemical species: an anionic chromophore and an aspartic residue.

3.2 Simulated electronic absorption spectrum

The simulated electronic absorption spectrum of LSSmKate2 (Fig. 3) built as we have explained in the methods section takes into account several configurations of the protein. Among them, both proton-relay configurations (that is to say, the Cro-Ser158-Asp160 and Cro-Asp160 ones, described in the former subsections) have been considered for the simulated absorption spectrum. To figure out whether both proton-relay configurations present different absorption frequencies or not, the absorption frequencies

corresponding to the 15 excited states calculated for each of the 160 snapshots (including only those transitions with oscillator strength equal to or greater than 0.1) have been represented versus $d_{Cro-Ser}-d_{Cro-Asp}$, where $d_{Cro-Ser}$ is the oxygen–oxygen distance between the chromophore and Ser158 and $d_{Cro-Asp}$ is the oxygen–oxygen distance between the chromophore and Asp160. When $d_{Cro-Ser}-d_{Cro-Asp}$ is positive, the chromophore is closer to Asp160 than to Ser158 (typical of the Cro-Asp160 proton-relay configuration). Conversely, when the value is negative, the chromophore is closer to Ser158 than to Asp160 (which would be typical of the Cro-Ser-Asp proton relay). The results are shown in Fig. 4. Most points are concentrated on the positive abscissa axis, this being a consequence of the predominance of the Cro-Asp proton-relay configuration along the whole dynamics. Notwithstanding this, no obvious correlation can be observed between the abscissa value
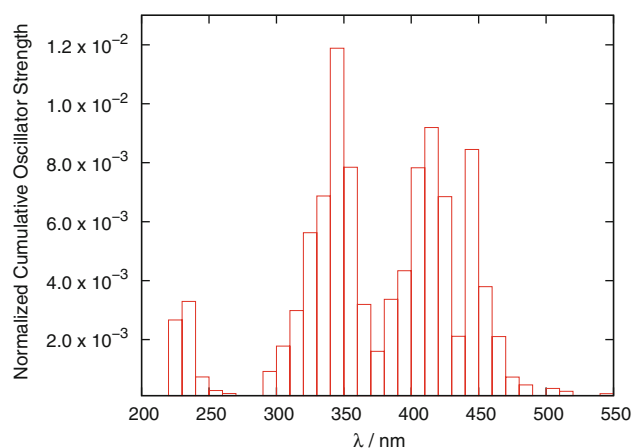

Fig. 3 Simulated electronic absorption spectrum of LSSmKate2. The *red columns* represent the sum of oscillator strength for each set of 10 nm (properly *scaled*)
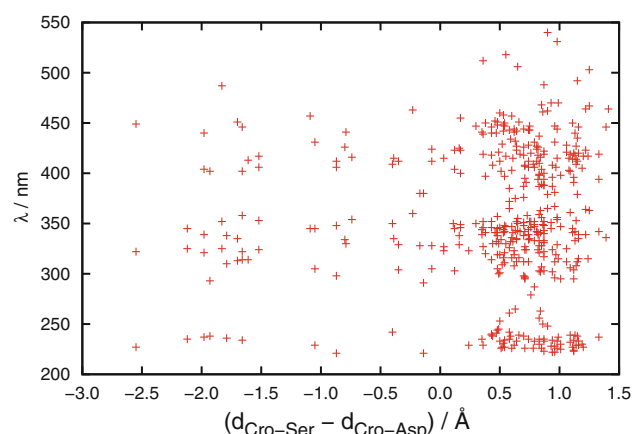

Fig. 4 Representation of the wavelengths obtained while calculating the spectrum versus the difference of O–O distances between Cro-Ser158 and Cro-Asp160

and the absorption wavelength, so both configurations appear equally spread along the spectrum.

The simulated absorption spectrum of LSSmKate2 (shown in Fig. 3) clearly presents two main bands in agreement with the available experimental data. The more energetic band has a maximum at about 345 nm and the second one shows two peaks at 415 and 445 nm. The latter absorption band, the one experimentally analyzed and known to produce red fluorescence at 605 nm, is well reproduced when compared with the experimental absorption maximum (460 nm). Nevertheless, the higher energy band between 300 and 375 nm is red-shifted by about 65 nm with respect to its corresponding experimental absorption band (graphically estimated at 280 nm [24]). Shifts of similar magnitude have been found in related studies [36]. Both bands present narrower bandwidths than the corresponding experimental values.

The spectrum depicted in Fig. 3 has been obtained from data consisting of 160 snapshots each of them contributing up to 15 excited states. From a statistical point of view, a possible concern could be the degree of convergence of this spectrum. To assess this point, we have proceeded as follows: we have computed different spectra based on data with increasingly large length of the simulation in steps of 10 ns and computed the spectra using exactly the same procedure and convoluted with a Gaussian function of $\sigma = 10$ nm to attenuate statistical noise. The results are shown in Fig. 5, and it can be seen that the overall shape of the spectra (number, intensity, width and position of the bands) is already stable from the beginning. As for the precise positions of the peaks, they undergo some small shifts and could change slightly if more samples were included.
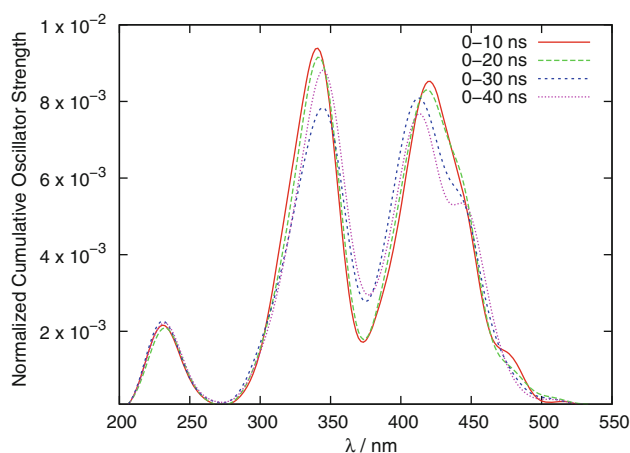
The use of molecular dynamics simulations has allowed us to obtain complete absorption bands that are attributable to thermal fluctuations (configurational changes) on the chromophore and its environment, adding contributions from different excited states reproducing the whole spectrum.

Analyzing the description of the electronic states accessed upon photoexcitation in terms of the orbital description, we can observe three π molecular orbitals that participate in the ππ* excitations leading to the UV–vis spectrum shown in Fig. 3: HOMO, LUMO and LUMO + 1. All of them are located over the chromophore, preserving the approximate molecular plane as a node. In Fig. 6, we have depicted these three molecular orbitals for Cro-Asp160 (taken at 4 ns of the MD simulation) and Cro-Ser158-Asp160 (taken at 10.5 ns of the MD simulation) hydrogen-bonded structures. It is clear that these molecular orbitals are identical for the two structures, so justifying that both are compatible with the experimental electronic absorption spectrum.

The right band (low energy) of the spectrum could be assigned to a HOMO → LUMO transition. As can be seen in Fig. 6, the HOMO → LUMO excitation represents an electron shift from the phenoxy moiety to the pyrimidine moiety of the chromophore. This represents effectively a "charge transfer" of intramolecular type and justifies the methodological choice made. Based on the nature of the molecular orbitals involved in this excited state, the excitation results in a depletion of charge of the phenoxy moiety
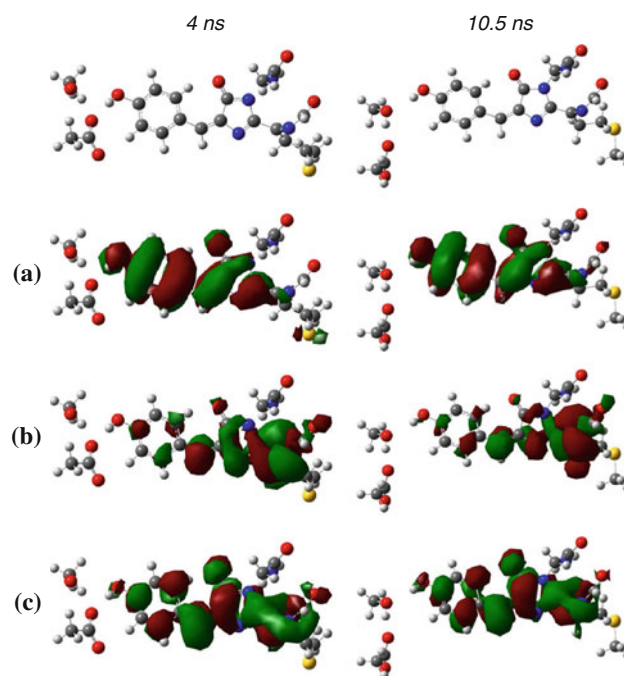

Fig. 5 Convergence of spectrum along the simulation. Each of the spectra has been computed on an increasingly longer part of the simulation. A convoluting Gaussian function of $\sigma = 10$ nm has been used to smooth the statistical noise

Fig. 6 Orbital representations of two selected frames at 4 and 10.5 ns of the MD simulation. The HOMO (a), LUMO (b) and LUMO + 1 (c) involved in the ππ* transitions are shown

which likely results in an increase of the acidity of the chromophore. Thus, this charge shift provides support for the ESPT hypothesis as a basis of the mechanism of generation of the fluorescent species. On the other hand, the higher energy band corresponds mostly to a HOMO →LUMO + 1 excitation. In this case, LUMO + 1 is a orbital evenly spread over the full chromophore, and the main difference with HOMO lies in the increased number of nodal planes perpendicular to molecular plane on the phenoxy moiety of the chromophore (from HOMO with 2 to LUMO + 1 with 3).

4 Conclusions

In the present work, we have run a MD simulation of the LSSmKate2 protein. From this simulation, we have been able to analyze the evolution of the plausible structures prone to host the proton transfer reaction that accounts for the fluorescence. Besides, from a series of snapshots, the electronic excited states have been calculated, furnishing a reasonable simulation of the experimental electronic absorption spectrum and allowing the analysis of the molecular orbitals involved in the absorption bands.

As for the molecular dynamics, the 40-ns simulation has proved the general structure of the protein to be dynamically stable, as already stated. However, apart from corroborating the availability of the two-link proton-wire configuration reported in the crystallographic structure, [24] a new single link proton-wire involving the chromophore and Asp160 has emerged as an even more likely configuration at least in terms of residence time. This new scenario would imply an analogy with the LSSmKate1, in which a one-link proton relay is suggested [24] to allow for the formation of the anionic chromophore, the species that would be responsible for the fluorescence. This series of different paths can provide interesting insights about the actual mechanism of proton transfer.

Focusing now on the methodology, the use of molecular dynamics simulations together with the calculation of excited states via TDDFT and the functional CAM-B3LYP has allowed us to simulate a full absorption spectrum that reasonably reproduces the available experimental data.

Regardless of the fact that two different configurations allow the formation of the anionic chromophore along, respectively a one-link or two-link proton transfer, the π molecular orbitals involved in the electronic excitation are not appreciably modified, which justifies that both configurations are compatible with the experimental electronic absorption spectrum. The two bands present a ππ* orbital transition. A transition between the HOMO to the LUMO accounts for the absorption band responsible for the red fluorescence. The migration of the electronic density toward the pyrimidine ring facilitates the shift of the proton

of the chromophore in the excited state, making a more favorable reaction. Conversely, the UV-band can be assigned to the HOMO → LUMO + 1 transition, implying a very minor displacement of the electronic density over the chromophore.

References

1. Prasher DC, Eckenrode VK, Ward WW, Prendergast FG, Cormier MJ (1992) Primary structure of the *Aequorea victoria* green-fluorescent protein. Gene 111:229–233
2. Shimomura O, Johnson FH, Saiga Y (1962) Extraction, purification and properties of Aequorin, a bioluminescent protein from luminous hydromedusan. Aequorea J Cell Compar Phys 59(3):223–239
3. Matz MV, Fradkov AF, Labas YA, Savitsky AP, Zaraisky AG, Markelov ML, Lukyanov SA (1999) Fluorescent proteins from nonbioluminescent Anthozoa species. Nat Biotechnol 17(10):969–973
4. Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA (2010) Fluorescent proteins and their applications in imaging living cells and tissues. Physiol Rev 90(3):1103–1163
5. Davidson MW, Campbell RE (2009) Engineered fluorescent proteins: innovations and applications. Nat Methods 6(10):713–717
6. Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC (1994) Green fluorescent protein as a marker for gene-expression. Science 263(5148):802–805
7. Day RN, Davidson MW (2009) The fluorescent protein palette: tools for cellular imaging. Chem Soc Rev 38(10):2887–2921
8. Jonasson G, Teuler J-M, Vallverdu G, Mérola F, Ridard J, Lévy B, Demachy I (2011) Excited state dynamics of the green fluorescent protein on the nanosecond time scale. J Chem Theor Comput 7(6):1990–1997
9. Laurent AD, Assfeld X (2010) Effect of the enhanced cyan fluorescent protein framework on the UV/visible absorption spectra of some chromophores. Interdiscip Sci Comput Life Sci 2(1):38–47
10. Laurent AD, Mironov VA, Chapagain PP, Nemukhin AV, Krylov AI (2012) Exploring structural and optical properties of fluorescent proteins by squeezing: modeling high-pressure effects on the mStrawberry and mCherry red fluorescent proteins. J Phys Chem B 116(41):12426–12440
11. Vendrell O, Gelabert R, Moreno M, Lluch JM (2006) Potential energy landscape of the photoinduced multiple proton-transfer process in the green fluorescent protein: classical molecular dynamics and multiconfigurational electronic structure calculations. J Am Chem Soc 128:3564–3574
12. Vendrell O, Gelabert R, Moreno M, Lluch JM (2008) A potential energy function for heterogeneous proton-wires. Ground end photoactive states of the proton-wire in the green fluorescent protein. J Chem Theor Comput 4:1138–1150

13. Vendrell O, Gelabert R, Moreno M, Lluch JM (2008) Operation of the proton wire in green fluorescent protein. A quantum dynamics simulation. J Phys Chem B 112(17):5500–5511

14. Bravaya KB, Bochenkova AV, Granovskii AA, Nemukhin AV (2008) Modeling of the structure and electronic spectra of green fluorescent protein chromophore. Russ J Phys Chem B 2(5):671–675

15. Li X, Chung LW, Mizuno H, Miyawaki A, Morokuma K (2010) Competitive mechanistic pathways for green-to-red photoconversion in the fluorescent protein kaede: a computational study. J Phys Chem B 114(49):16666–16675

16. Filippi C, Buda F, Guidoni L, Sinicropi A (2012) Bathochromic shift in green fluorescent protein: a puzzle for QM/MM approaches. J Chem Theor Comput 8(1):112–124

17. Bravaya KB, Grigorenko BL, Nemukhin AV, Krylov AI (2012) Quantum chemistry behind bioimaging: insights from Ab Initio studies of fluorescent proteins and their chromophores. Acc Chem Res 45(2):265–275

18. Zhang H, Sun Q, Wang S, Olsen SL, Smith SC (2011) Hydrogen bonding and transfer in the excited state. Wiley, New York, p 815

19. Sanchez-Garcia E, Doerr M, Thiel W (2010) QM/MM study of the absorption spectra of DsRed.M1 chromophores. J Comput Chem 31(8):1603–1612

20. Imhof P (2012) Computational study of absorption spectra of the photoconvertible fluorescent protein EosFP in different protonation states. J Chem Theor Comput 8(11):4828–4836

21. Jackson S, Sanders J (2009) Green fluorescent protein issue. Chem Soc Rev 38(10):2821

22. Piatkevich KD, Efremenko EN, Verkhusha VV, Varfolomeev DD (2010) Red fluorescent proteins and their properties. Russ Chem Rev 79(3):243–258

23. Piatkevich KD, Hulit J, Subach OM, Wu B, Abdulla A, Segall JE, Verkhusha VV (2010) Monomeric red fluorescent proteins with a large Stokes shift. Proc Natl Acad Sci USA 107(12):5369–5374

24. Piatkevich KD, Malashkevich VN, Almo SC, Verkhusha VV (2010) Engineering ESPT pathways based on structural analysis of LSSmKate red fluorescent proteins with large Stokes shift. J Am Chem Soc 132(31):10762–10770

25. Ortiz-Sánchez JM, Gelabert R, Moreno M, Lluch JM, Anglada JM, Bofill JM (2010) Bipyridyl derivatives as photomemory devices: a comparative electronic-structure study. Chem-Eur J 16(22):6693–6703

26. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, Baker NA (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. Nucl Acids Res 35:W522–W525

27. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucl Acids Res 32:W665–W667

28. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 102(18):3586–3616

29. MacKerell AD, Feig M, Brooks CL (2004) Improved treatment of the protein backbone in empirical force fields. J Am Chem Soc 126(3):698–699

30. MacKerell Jr AD, Brooks B, Brooks III CB, Nilsson L, Roux B, Won Y, Karplus M (1998) CHARMM: The energy function and its parametrization with an overview of the program. Encyclopedia of computational chemistry, P. v.R.Schleyer, N.L Allinger, T. Clark, J. Gasteiger, P.A. Kollman, H.F. Schaefer III, P.R. Schreiner, Chichester vol. 1

31. Brooks BR, Brooks CL, Mackerell AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M (2009) CHARMM: the biomolecular simulation program. J Comput Chem 30(10):1545–1614

32. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) Charmm—a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4(2):187–217

33. Zimmer M (2002) Green fluorescent protein (GFP): applications, structure, and related photophysical behavior. Chem Rev 102(3):759–781

34. Reuter N, Lin H, Thiel W (2002) Green fluorescent proteins: empirical force field for the neutral and deprotonated forms of the chromophore. Molecular dynamics simulation's of the wild type and S65T mutant. J Phys Chem B 106(24):6310–6321

35. Nadal-Ferret M, Gelabert R, Moreno M, Lluch JM (2012). Submitted

36. Murugan NA, Kongsted J, Rinkevicius Z, Agren H (2012) Color modeling of protein optical probes. Phys Chem Chem Phys 14(3):1107–1112

37. 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 J, J. A., 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 NJ, 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 Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2011) Gaussian 09, Revision C.1. Gaussian, Inc, Wallingford CT, 2009

38. Sherwood P, de Vries AH, Guest MF, Schreckenbach G, Catlow CRA, French SA, Sokol AA, Bromley ST, Thiel W, Turner AJ, Billeter S, Terstegen F, Thiel S, Kendrick J, Rogers SC, Casci J, Watson M, King F, Karlsen E, Sjovoll M, Fahmi A, Schafer A, Lennartz C (2003) QUASI: a general purpose implementation of the QM/MM approach and its application to problems in catalysis. J Mol Struc-Theochem 632:1–28

39. Handy NC (2004) The molecular physics lecture 2004: (i) density functional theory, (ii) quantum monte carlo. Mol Phys 102 (23–24):2399–2409

40. Aquino AJA, Lischka H, Hattig C (2005) Excited-state intramolecular proton transfer: a survey of TDDFT and RI-CC2 excited-state potential energy surfaces. J Phys Chem A 109:3201–3208

41. Sobolewski AL, Domcke W (1999) Ab initio potential-energy functions for excited state intramolecular proton transfer: a comparative study of o-hydroxybenzaldehyde, salicylic acid and 7-hydroxy-1-indanone. Phys Chem Chem Phys 1(13):3065–3072

42. Zhang MY, Wang JY, Lin CS, Cheng WD (2011) First-principles study of one- and two-photon absorption of the H-bonding complexes from monomeric red fluorescent proteins with large stokes shifts. J Phys Chem B 115(36):10750–10757

43. Randino C, Moreno M, Gelabert R, Lluch JM (2012) Peek at the potential energy surfaces of the LSSmKate1 and LSSmKate2 proteins. J Phys Chem B 116(49):14302–14310

44. Tawada Y, Tsuneda T, Yanagisawa S, Yanai T, Hirao K (2004) A long-range-corrected time-dependent density functional theory. J Chem Phys 120(18):8425–8433

45. Isborn CM, Götz AW, Clark MA, Walker RC, Martínez TJ (2012) Electronic absorption spectra from MM and ab Initio QM/MM molecular dynamics: environmental effects on the absorption spectrum of photoactive yellow protein. J Chem Theor Comput 8(12):5092–5106

46. Fantacci S, De Angelis F, Sgamellotti A, Marrone A, Re N (2005) Photophysical properties of [Ru(phen)2(dppz)]2+ Intercalated into DNA: an integrated Car− Parrinello and TDDFT study. J Am Chem Soc 127(41):14144–14145

47. De Angelis F, Fantacci S, Sgamellotti A (2007) An integrated computational tool for the study of the optical properties of nanoscale devices: application to solar cells and molecular wires. Theor Chem Acc 117(5–6):1093–1104

48. Eriksen JJ, Olsen JMH, Aidas K, Agren H, Mikkelsen KV, Kongsted J (2011) Computational protocols for prediction of solute NMR relative chemical shifts. A case study of L-Tryptophan in aqueous solution. J Comput Chem 32(13):2853–2864

49. Murugan NA, Aidas K, Kongsted J, Rinkevicius Z, Agren H (2012) NMR Spin–Spin coupling constants in polymethine dyes as polarity indicators. Chem-Eur J 18(37):11677–11684

Dancing multiplicity states supported by a carboxylated group in dicopper structures bonded to O_2

Albert Poater · Luigi Cavallo

Abstract The present study pretends to assign the correct multiplicity state to dinuclear copper complexes when interacting with free molecular oxygen. Recently, the formation of a bridge butterfly μ–η^2:η^2-peroxo dicopper core structure stabilized by the direct interaction of the counterion, a carboxylate group that allows the double bridge linking both metal-centre atoms, was characterized by crystallography. This system was assigned as a diradical singlet with $Ms = 0$. However, after new calculations it has turned out to be triplet ($Ms = 1$) despite the stabilization for this latter multiplicity state is not high. Here, the factors that contribute to make this structure display a multiplicity different with respect to the previously expected diradical singlet are described. In the present theoretical study, the roles of the αSp ligand constraints and the counterion are unravelled. On the other hand, the relative stability between the butterfly μ–η^2:η^2-peroxo structure and the isomeric bis(μ-oxo) species is also on discussion. Despite the relative stabilities of all these either structural or electronic isomeric species are supposed to depend on the computational method, which is a difficulty to reach a definite conclusion about the nature of the active species, all DFT methods using either pure or not pure DFT functionals here reach the same conclusion, favoring the triplet as the ground state for the butterfly μ–η^2:η^2-peroxo dicopper core structure, and the bis(μ-oxo) species when removing the benzoate counterion.

1 Introduction

The binding and activation of O_2 through multiple metal centres play an important role in the catalytic cycle of many metalloenzymes [1–7]. This explains recent efforts to characterize the interaction of O_2 with monocopper [8–10] and dicopper complexes [11–19], the latter of particular interest for being type III copper proteins, with tyrosinase, haemocyanin, and catechol oxidases [20–28] as the prototypical cases of proteins containing dinuclear copper-active sites capable of activating O_2. They often display the μ–η^2:η^2-peroxodicopper(II) (μ–η^2:η^2-peroxo-Cu(II)$_2$) and bis(μ-oxo)dicopper(III) ((μ-O)$_2$–Cu(III)$_2$) species, displayed in Scheme 1. The first crystal structure of a μ–η^2:η^2-peroxo-Cu(II)$_2$ species was first described by Kitajima et al. [29, 30] in 1989, while the bis(μ-oxo)-Cu(III)$_2$ isomer was described by Tolman et al. [31] in 1995.

Recently, a crystallographic study on tyrosinase has revealed significant changes of the Cu…Cu distance along the catalytic cycle for accommodating the entering O_2 molecule [32]. This reinforces the idea that more insight is still required to understand the interaction of O_2 with

A. Poater (✉)
Departament de Química, Institut de Química Computacional, Universitat de Girona, Campus de Montilivi, 17071 Girona, Catalonia, Spain
e-mail: albert.poater@udg.edu

L. Cavallo
KAUST Catalyst Research Center, 4700 King Abdullah University of Science and Technology, Thuwal 23955-6900, Kingdom of Saudi Arabia
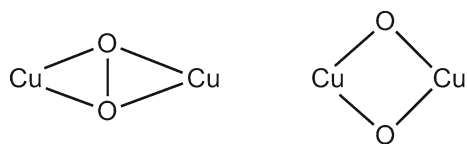
Scheme 1 Schematic representation of peroxo and bis(μ-oxo) dicopper cores

dinuclear Cu^I complexes [28], and indeed a variety of structural motifs for the binding of O$_2$ has been reported [33–39], although the side-on μ–η^2:η^2-peroxo and bis(μ-oxo) isomeric [Cu$_2$O$_2$]$^{2+}$ cores are the most characterized ones experimentally. The transformation between both isomers is facile and their relative stability can be tuned by the electronic and steric behaviour of the ligands, and even by using different organic solvents or counteranions [40–47], low molecular weight biomimetic models have been of great help for understanding the spectroscopic and structural properties of the active site of these proteins. They also show how subtle variations in ligand design strongly affect their reactivity towards O$_2$, and that of the corresponding oxygenate complex towards its intramolecular oxidation or towards the oxidation of an external substrate [5, 44, 48–55].

Theoretical methods have contributed remarkably to understand these systems. For example, for monooxygenase systems [56–60], it has been possible to describe fully the reactivity in aliphatic hydroxylations, as well as for biomimetic models of the tyrosinase [61–70]. However, the debate on the exact role of both geometries for the [Cu$_2$O$_2$]$^{2+}$ cores, that is, the μ–η^2:η^2-peroxo and bis(μ-oxo) isomers, is still open. It is thought that the reactivity probably starts from this point, however, even carefully designed experiments are often unable to clearly characterize all the intermediates involved in this reaction. Therefore, the nature of the active species in this reaction is still a matter of discussion [71, 72]. Furthermore, these isomeric species represent a complicated case for computational studies trying to define the right ground state, due to the small energy difference between the closed-shell and open-shell singlets, and also the triplet states [34, 35, 73–76].

Masuda et al. [47] very recently reported a detailed characterization of the complex [Cu(II)$_2$(αSp)$_2$(μ–η^2:η^2-O$_2$)(Bz$^-$)]$^+$, and the X-ray structure reported in this study has given insight about the important role first of the α-isosparteine (αSp) and second of the counterion, benzoate (Bz$^-$). These two particular factors prevent this structure to adopt a bis(μ-oxo) core. Computations suggested this structure to assume an open-shell singlet multiplicity as the ground state, in agreement with a previous recent study in which Stack et al. [46] described in detail that the axial binding of an anion induces an electronic and not a steric preference for the butterfly μ–η^2:η^2-O$_2$ dicopper core

structure species. However, due to its butterfly shape, this structure displays the triplet multiplicity state close in energy [22]. To verify which is the right multiplicity for the ground state of the structure reported by Masuda et al., some calculations were envisaged, describing the factors that contribute to make this structure display a different multiplicity with respect to the previously expected diradical singlet. Either removing the benzoate or simplifying the structure of the ligand, the role of the counterion explicitly of the ligand has been unravelled. Furthermore, the transformation between the μ–η^2:η^2-peroxo and bis(μ-oxo) isomeric [Cu$_2$O$_2$]$^{2+}$ cores has been analysed by means of density functional theory (DFT) calculations. Although the singlet ground state for bis(μ-oxo) species is clear, a discussion for the μ–η^2:η^2-peroxo core is still open.

2 Theoretical methods

All geometry optimizations, with no symmetry constraints, have been performed with the Gaussian09 package [77], using the B3LYP functional [78–80] and the standard 6-31G(d) basis set [81, 82]. The geometries obtained at the B3LYP/6-31G(d) level were used to perform single-point energy calculations with a larger basis set, the 6-311G(d,p) basis set, and the same functional (B3LYP/6-311G(d,p)// B3LYP/6-31G(d)). For Cu, this basis corresponds to the (14s9p5d)/[9s5p3d] Wachter's basis set [83], with the contraction scheme 611111111/51111/311 supplemented with one f polarization function. In the Gaussian 09 implementation of the Cu basis set, the s and p functions come from Wachter's optimization for the Cu atom in its ^{2}S state, while the d functions come from Wachter's optimization for the Cu atom in its ^{2}D state. Solà et al. [84] reported in 2008 that the Gaussian internal basis set provides more reasonable results for the relative energies among the different analysed states than the basis set with the d functions coming from Wachter's optimization for the Cu atom in its ^{2}D state.

Intrinsic reaction pathways were calculated to confirm that the transition states located connect the expected minima. Analytical Hessians were computed to determine the nature of all the located stationary points and to calculate zero-point energies (ZPEs) and thermodynamic properties at 298 K.

For open-shell states, the geometry optimizations were performed within the broken-symmetry unrestricted methodology, while for the closed-shell singlet states, the restricted formalism was used. Theoretical treatment of diradical singlet species requires multiconfigurational or multireference methods due to strong static electron correlation. Unfortunately, these methods can only be applied to relatively small systems because computationally they

are extremely demanding. As an alternative, the unrestricted UB3LYP method in broken symmetry (BS, using GUESS = MIX) has been used [85]. This method improves the modelling of diradical singlet states at the expense of introducing some spin contamination from higher spin states [86–93]. Although this is not the most appropriate method to treat diradical singlet species, it has been shown that it can be used provided that the overlap between the open-shell orbitals is small (i.e. the unpaired electrons are located in separated atomic centres), which is the case of the systems that show predominant diradical character in this work [92]. In a previous paper, for *p*-benzyne, the application of the sum rule [91] to the energy of the diradical singlet state to remove the spin contamination error does not improve the calculated singlet–triplet energy gap but rather leads to larger errors when compared to the experimental value [85]. For this reason, open-shell singlet energies reported in this work does not contain spin contamination corrections, except for one example computed to corroborate that they are not key here. Anyway, the corresponding singlet energies from unrestricted calculations eliminating spin contamination from the triplet state in the SCF solution are calculated through the sum method [94]. In this approach, the singlet energy is computed as [95–97]

$$E_{\text{singlet}} = \frac{2E_{\langle S_z^2 \rangle=0} - \langle S^2 \rangle E_{\langle S_z^2 \rangle=1}}{2 - \langle S^2 \rangle} \qquad (1)$$

where the triplet energy is computed for the single-determinantal high-spin configuration $S_z = 1$ and $<S^2>$ is the expectation value of the total-spin operator applied to the Kohn–Sham (KS) determinant for the unrestricted $S_z = 0$ calculation.

Solvent effects including contributions of non-electrostatic terms have been estimated in single-point calculations on the gas phase optimized structures, based on the polarizable continuum solvation model (PCM) using CH_2Cl_2 as a solvent, the same solvent used experimentally [98, 99].

The relative free energies reported in this work include energies computed using the B3LYP/6-311G(d,p)//B3LYP/6-31G(d) method together with solvent effects obtained at the B3LYP/6-31G(d) level and zero-point energies, thermal corrections, and entropy effects evaluated at 298 K with the B3LYP/6-31G(d) method in gas phase.

To analyse the electrophilicity of the optimized complexes, we have calculated the electrophilicity index defined by Parr et al. [100] as:

$$\omega = \frac{\mu^2}{2\eta}, \qquad (2)$$

where μ and η are the chemical potential and the molecular hardness, respectively. In the framework of the conceptual

density functional theory (DFT) [101], the chemical potential and molecular hardness for an *N*-electron system with total electronic energy E and external potential are defined as the first and second derivatives of the energy with respect to N at a fixed external potential [102–104]. In numerical applications, μ and η are calculated through the following approximate versions based upon the finite difference approximation and the Koopmans' theorem [105],

$$\mu \cong \frac{1}{2}(\varepsilon_L + \varepsilon_H) \qquad (3)$$

$$\eta \cong \frac{1}{2}(\varepsilon_L + \varepsilon_H), \qquad (4)$$

where ε_H and ε_L are the energies of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO), respectively.

3 Results and discussion

It is worth noting that for all the species, the closed-shell singlet, the diradical singlet, the triplet, and even the quintuplet electronic states have been computed. First, test calculations on quintuplet multiplicity states were always found to be much higher in energy than the other states, and thus they are not included in the discussion. On the other hand, the energy difference between the diradical singlet and the triplet states is lower than 5 kcal mol^{-1} in all cases, thus both multiplicities must be taken into account for all structures except for the bis(μ-oxo) species, which differently favours the closed-shell singlet state. For all the other minima, the diradical singlet is favored and the closed-shell singlet is much higher in energy.

The starting point of this study was the optimization in the triplet state of complex $[Cu(II)_2(\alpha Sp)_2(\mu-\eta^2:\eta^2-O_2)(Bz^-)]^+$ (**1-a**) from the crystallographic data, resulting in a perfect agreement between the DFT structure and the X-ray structure (rmsd = 0.046 Å for distances and 1.3° for angles)[1] [106–112], validating the computational method [62]. For an easy comparison with the paper of Masuda et al. [47], the atom labelling for all species is denoted with the same description they used. This species displays a triplet multiplicity state, being 0.2 kcal mol^{-1} lower in energy with respect to the diradical singlet (**1-b**). This is in opposition to the experimental results of Masuda et al. [47]. Anyway, the preference for the triplet as the most stable

---

[1] Standard deviations for the distances and for the angles, $S_{n-1} = \sqrt{\frac{\sum_{i=1}^{N}(CV-EV)^2}{N-1}}$, where CV means calculated value, EV experimental value (X-ray data), and N is the number of distances or angles taken into account (See supporting information for details).

multiplicity state is corroborated by the fact that removing the possible overstabilization due to spin contamination with Eq. (1), this energetic difference increases slightly to 0.5 kcal mol^{-1}. This also reinforces the hypothesis vide supra that the effect of the spin contamination is negligible [95–97]. Then, in gas phase the difference is 0.3 kcal mol^{-1}. We also tested the stability of this conclusion from the specific correlation–exchange functional used [113–118]. Both GGA and HGGA functionals, with BLYP and BHLYP calculations [119], validate the B3LYP results. For BLYP, the difference is 1.2 kcal mol^{-1}. However, it is necessary to point out that the triplet BLYP structure is quite distorted with respect to the B3LYP one, weakening slightly one Cu–O bond for each copper atom, reaching a pseudo-trans μ–η^1:η^1-peroxo-Cu(II)$_2$ species. Furthermore with the BLYP functional all attempts performed for spin-unrestricted broken-symmetry calculations for the singlet diradical state, starting from unsymmetrical wavefunctions, have reconverged to the restricted solution. This means that, at least at the BLYP level of theory, the singlet is closed-shell. Indeed delocalized states are overstabilized with pure density functionals as a result of a bad cancellation of the self-interaction included in the Coulomb energy by the exchange–correlation functional [120–124]. This overstabilization of delocalized states is partially corrected when hybrid functionals are used, in particular those that incorporate a large percentage of exact Hartree–Fock exchange [125]. However, even with an exchange of 50 %, that is, with BHLYP functional, the diradical singlet does not turn out to be more stable than the triplet species, being just only isoenergetic, thus being less stable than the diradical singlet when removing the possible overstabilization due to spin contamination [95–97]. Furthermore, this BHLYP functional does not reproduce the geometry as well as the B3LYP does, due to an exaggeration of the Hartree–Fock exchange contribution [113–118], whereas B3LYP displays a much lower exchange [126, 127]. To verify the B3LYP results, other DFT functionals as B3LYP*, PBE, OPBE, BP86, BPW91, B3PW91, and M05 were tested. Although in several papers, it is defended that pure DFT functionals describe better the interaction of O$_2$ with dicopper molecules [62, 73–76], neither better insight nor significant qualitative changes were found. Thus, the discussion is basically based on the B3LYP because apart from previous experience, its results here are in better agreement in geometrical terms with respect to the experimental X-ray data (Fig. 1).

The closed-shell singlet species for the μ–η^2:η^2-O$_2$ isomer (1-c) is 16.1 kcal mol^{-1} higher in energy with respect to 1-a. The singlet–triplet crossing is out of scope of this paper and should require high-level computational tools which are not available yet for relatively large inorganic systems. On the other hand, the closed-shell singlet
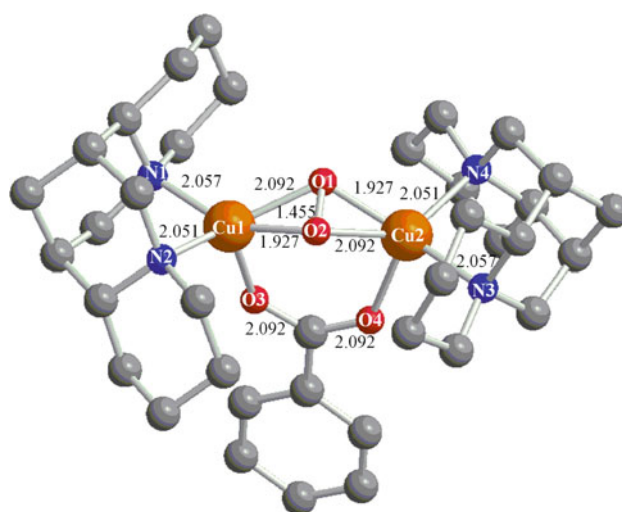

Fig. 1 Computed structure for the [Cu$_2^{II}$(αSp)$_2$(μ–η^2:η^2–O$_2$)(Bz$^-$)]$^+$ complex, **1-a** (distances in Å). For the sake of clarity all H atoms have been omitted

bis(μ-oxo) isomer (1-e) is less stable by 19.0 kcal mol^{-1}, which is not reached anyway due to the fact that requires to overcome a barrier of 25.1 kcal mol^{-1}. This high instability of the bis(μ-oxo) core is in agreement with previous studies, and somewhat a bit exaggerated with B3LYP [37, 62]. The cores of bis(μ-oxo) (1-e) and the transition state (1-d), as well as the closed-shell, the open-shell singlet, and triplet for the μ–η^2:η^2-O$_2$ species are shown in Fig. 2. The structural similarity between species 1-a and 1-b seems obvious according to the negligible energetic difference between them. However, the DFT optimized species 1-a has a Cu$\cdots$Cu distance of 3.256 Å, in perfect agreement with the experimental value (3.265 Å) [47], and 1-a also displays a higher butterfly character with a O2–Cu1–O1–Cu2 dihedral angle of 61.6° for the Cu$_2$O$_2$ moiety, versus a value of 46.3° in 1-c [128, 129]. On the other hand, despite Cu(III) being more stable in a perfect plane [5, 21, 130, 131], this butterfly-shaped core is not completely lost in species 1-e, displaying a dihedral angle of 23.7°. Although higher, this strong bending of the Cu$_2$O$_2$ core in species 1-a is in agreement with respect to other previous studies for similar μ–η^2:η^2-O$_2$ dicopper systems [6, 16, 44, 47, 125, 132, 133]. This increased effect of the bending is due to either the nature of the diaza ligand (αSp), which brings a structural constraint that enforces tetrahedral distortion around the transition metal(II) ions [134–138], or because of the axial coordination of the bridging carboxylate group of benzoate [47]. To discard any effect of the anionic part of the core in complex 1-a reoptimization of species 1-a was done with inclusion of diffuse base functions but no significant differences were found.

At this point, a digression on the relative stability of the μ–η^2:η^2-peroxo and of the bis(μ-oxo) is needed. As
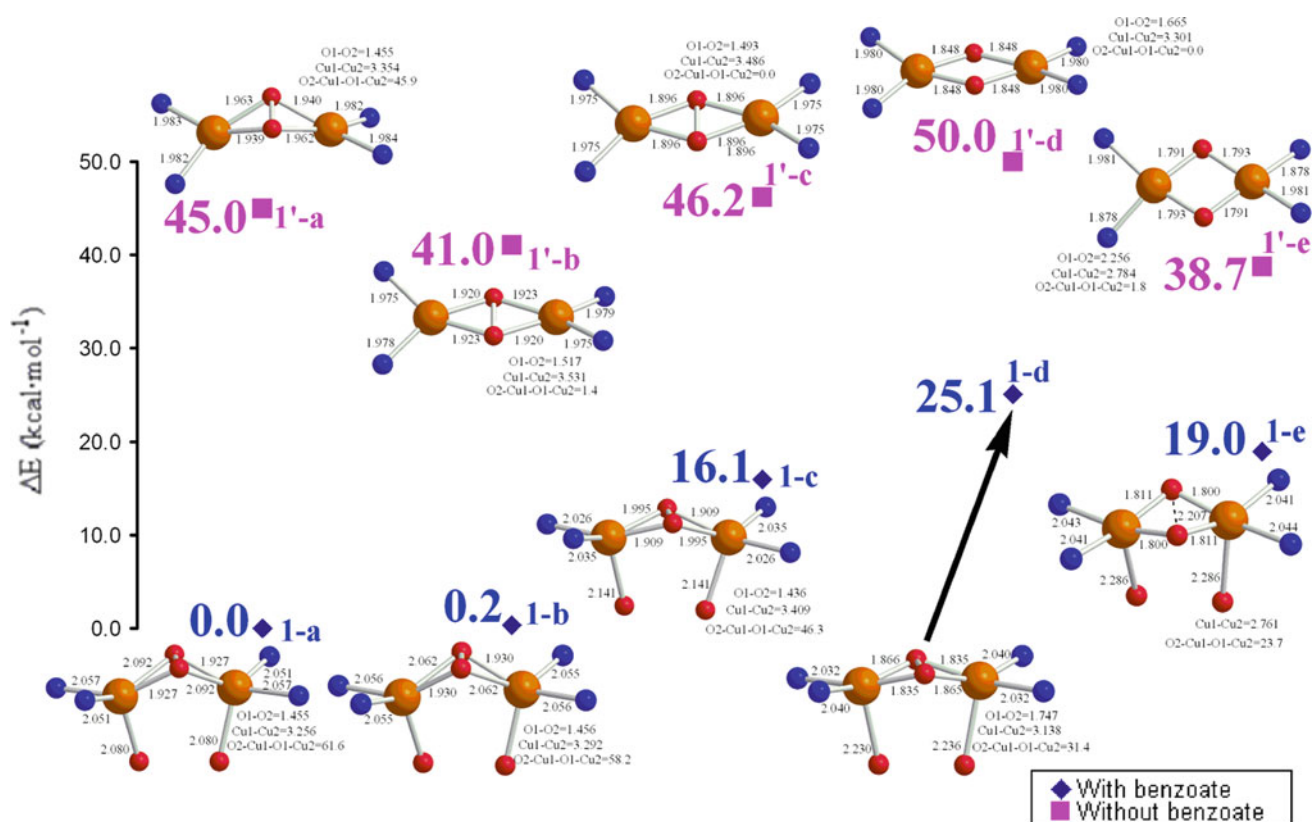
Fig. 2 Computed cores for **1-a–1-e** species with and without the benzoate anion (energies in kcal mol^{-1} relative to species **1-a**, distances in Å and angles in degrees; and all C and H atoms have been omitted)

reported by Cramer [33–35], the equilibrium between the two isomers is artificially displaced towards the peroxo species by HGGA functionals such as the B3LYP functional, due to unbalanced correlation corrections, while GGA functionals like the BLYP functional provide better relative energies [33–35]. Indeed, as stated vide supra, test calculations with the BLYP functional show that the preference for the μ–η^2:η^2-peroxo species is reduced from 19.0 to 1.2 kcal mol^{-1} when switching from the B3LYP functional to the BLYP functional.

The bis(μ-oxo) structure **1-e** displays longer Cu–O bond distances with the carboxylate of the benzoate than **1-b** by about 0.2 Å, however, the carboxylate counterion forces to lose planarity despite the Jahn–Teller effect. Complex μ–η^2:η^2-peroxo **1-a** evolves to the bis(μ-oxo) isomer **1-e** with an energetic destabilization of 19.0 kcal mol^{-1} and the barrier for this conversion requires 25.1 kcal mol^{-1}. The path from both isomers means the cleavage and, consequently, elongation of the O–O bond distance, and electronically the degree of O_2 reduction changes [46]. This transformation is accompanied by a change in the multiplicity from triplet to singlet. The large spatial separation between the two unpaired electrons located one on each copper atom first favours the triplet multiplicity, but the shortening of this distance in the bis(μ-oxo) isomer

favours the cleavage of the O_2 moiety and thus, the closed-shell singlet [73–76].

To separate the effect of the benzoate, the structure **1-a** was optimized without the benzoate anion (**1'-a**). Figure 3 displays the relative stability of the species with the benzoate (**1-a–1-e**) and without the benzoate (**1'-a–1'-e**) with respect to **1-a**. The main difference when extracting the benzoate moiety is that the bis(μ-oxo) becomes the most stable isomeric species by 2.3 kcal mol^{-1} [134] and obviously, with a lower barrier, 9.0 kcal mol^{-1}, for reaching this isomer from the μ–η^2:η^2-O_2 species, in this case diradical singlet as ground state. It is obvious that the barrier **1'-a'**→**1'-e** is lower than in **1-a**→**1-e** because there are no geometrical constraints introduced by the benzoate and, furthermore, the bidentate copper stabilizes more the bis(μ-oxo) species. Only the open-shell singlet structure, **1'-b**, maintains the butterfly-shaped core (see Table 1). Obviously the species that is stabilized most by a bonded benzoate is **1-a** with respect to **1'-a**, by 45.0 kcal mol^{-1}, in comparison with the process **1'-e**→**1-e** that releases only 19.7 kcal mol^{-1}. However, when bonding the entering benzoate moiety to species **1'-e**, the real transition is **1'-e**→**1-a**, releasing 38.7 kcal mol^{-1} [134]. Overall, the formation of a bridged butterfly [Cu_2O_2]$^{2+}$ core in species **1-a** is feasible thanks to the carboxylate group that binds

E_H = -0.305   E_L = -0.199

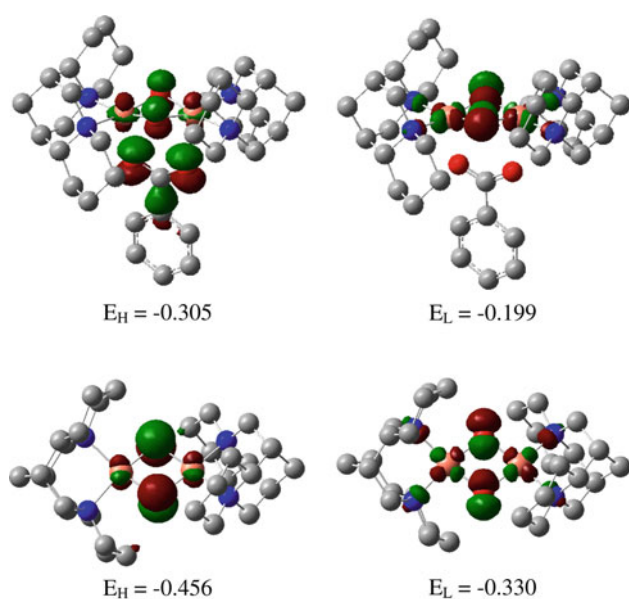E_H = -0.456   E_L = -0.330

Fig. 3 Computed HOMO and LUMO orbitals for species **1-e** and **1'-e** (energies in eV). For the sake of clarity, all H atoms have been omitted

through its oxygen atoms axially with respect to each copper atom. This interaction removes the planarity of the $[Cu_2O_2]^{2+}$ core regardless of the Jahn–Teller effect [139], suggesting a geometrical character for the interaction of the benzoate [46].

On the other hand, it is necessary to point out that the barrier from the peroxo to the bis(μ-oxo) species evolves from 24.9 to 9.0 kcal/mol when removing the benzoate counterion.

At the BLYP level, the difference between the μ–η^2:η^2-peroxo and the bis(μ-oxo) isomers increases to 22.2 kcal mol^{-1}, taking into account that for BLYP the diradical singlet for the first isomer is not stable. Furthermore, it is important to point out here that the calculated energy difference between both species depends on several effects such as the percentage of Hartree–Fock exchange in HGGA functionals, inclusion of solvent effects and the use of different basis sets [116]. For instance, test calculations with the BLYP functional, without including zero-point energies, thermal corrections, and entropy effects, indicate that the bis(μ-oxo) form **1'-e** is favored by 32.9 kcal mol^{-1} when the smaller 6-31G(d) basis set is used in gas phase, which value must be compared to 22.0 kcal mol^{-1} for B3LYP/6-311G(d,p)//B3LYP/6-31G(d) method with solvent effects. Highly correlated methods should be used to solve this issue, but the size of the systems precludes this possibility. Nevertheless, the reliability of B3LYP relative energies has been substantiated by previous studies [61–70, 113–118].

To better describe the importance of the αSp, a complete theoretical characterization of four systems **2–5** in Fig. 1

and Scheme 2 is proposed; studying either the sterical constraints according to the substituents on the N atoms or the rigidity due to the alkyl linker between the N atoms. System **2** removes part of the steric encumbrance on the N atom. System **3** removes the chelation effect. System **4** maintains the chelation effect but reduces deeply the steric encumbrance on the N atoms. Finally, system **5** removes both effects together. For each system, the μ–η^2:η^2-peroxo and the isomeric bis(μ-oxo) species are on discussion, playing with all possible multiplicity states, closed-shell and open-shell singlet or triplet as done for system **1**. Furthermore, for the μ–η^2:η^2-peroxo→bis(μ-oxo) isomerization processes, the barriers have also been studied. Tables 1 and 2 collect information about the most relevant geometrical parameters and the relative energetics, respectively. Again the diradical singlet μ–η^2:η^2-peroxo species is never more stable than the homologous triplet species. In gas phase, the energy difference is 0.3, 0.6, 0.4, 0.7, 0.6 kcal mol^{-1} for species **1**, **2**, **3**, **4**, and **5**, respectively. This reinforces the finding that the triplet is the ground state for species **1-a**. We also checked if the restricted geometry would give a lower energy singlet after removing spin contamination from a broken-symmetry calculation than the structure optimized with broken symmetry. The result is that for the five structures this alternative way drives to a state around 5 kcal mol^{-1} higher in energy with respect to the optimized open-shell singlet structure.

Comparison between species **1-a** and **2-a** reveals that they are geometrically similar but it is necessary to point out the slight differences caused by the simplification of the αSp ligand. According to the identical four Cu–O bond distances, the $[Cu_2O_2]^{2+}$ core in **2-a** displays a perfect symmetry allowing a C_{2v} symmetry group for the whole charged +1 molecule. And the Cu–Cu decreases from 3.256 to 3.048 Å. This shortening helps to understand why the bis(μ-oxo) isomer becomes more stable than the butterfly μ–η^2:η^2-peroxo-Cu(II)$_2$ species for complex **2**. In **2-e**, there is an elongation of the O···O distance with respect to **1-e** by 0.180 Å, as well as a shortening of 0.180 Å of the Cu···Cu distance. Thus, the substituents on the N atoms of the diaza ligand show a great influence in the stability of the different isomeric motifs of the $[Cu_2O_2]^{2+}$ core. For system **3-e**, the sterical hindrance between methyl groups destabilizes this bis(μ-oxo) isomer, displaying an even shorter O···O distance with respect to **1-e** by 0.018 Å, which explains the similar behaviour of complexes **1** and **3**, displaying both a butterfly μ–η^2:η^2-peroxo-Cu(II)$_2$ structure by means of DFT calculations. On the other hand, comparison between complexes **4** and **5** allows the comprehension of the potential effect of the diaza ligand, becoming the bis(μ-oxo) isomer the most stable isomer thanks to the lower repulsion between the H atoms in **5-e**.

Table 1 Geometrical parameters for the computed cores for species **1–5** with the benzoate anion (distances in Å and angles in degrees)

	O–O	Cu–Cu	Cu1–O1	Cu1–O2	Cu2–O1	Cu2–O2	Cu1–O3	Cu2–O4	Cu1–N1	Cu1–N2	Cu2–N3	Cu2–N4	O2–Cu1–O1–Cu2
1-a	1.455	3.256	2.092	1.927	1.927	2.092	2.080	2.080	2.057	2.051	2.057	2.051	61.6
1-b	1.456	3.292	2.062	1.930	1.930	2.062	2.086	2.085	2.055	2.056	2.056	2.055	58.2
1-c	1.436	3.409	1.995	1.909	1.909	1.995	2.141	2.141	2.026	2.035	2.026	2.035	46.3
1-d	1.747	3.138	1.866	1.835	1.835	1.865	2.230	2.236	2.032	2.040	2.032	2.040	31.4
1-e	2.207	2.761	1.811	1.800	1.800	1.811	2.286	2.286	2.043	2.041	2.044	2.041	23.7
2-a	1.476	3.048	1.968	1.968	1.968	1.968	2.061	2.061	2.019	2.019	2.019	2.019	60.3
2-b	1.480	3.152	1.994	1.952	1.952	1.994	2.072	2.072	2.013	2.009	2.013	2.895	58.5
2-c	1.457	3.307	1.921	1.921	1.921	1.921	2.116	2.116	1.997	1.997	1.997	1.997	46.2
2-d	1.671	3.137	1.851	1.851	1.851	1.851	2.174	2.174	1.992	1.992	1.992	2.885	36.1
2-e	2.387	2.577	1.799	1.799	1.799	1.799	2.251	2.251	1.941	1.941	1.941	1.941	24.5
3-a	1.452	3.152	1.984	1.988	1.988	1.984	2.070	2.070	2.089	2.075	2.089	2.075	59.1
3-b	1.419	3.149	1.958	1.952	1.952	1.958	2.087	2.087	2.092	2.083	2.092	2.083	58.2
3-c	1.419	3.149	1.958	1.952	1.952	1.958	2.087	2.087	2.092	2.083	2.092	2.083	45.3
3-d	1.746	3.108	1.843	1.849	1.849	1.843	2.191	2.191	2.073	2.054	2.073	2.054	33.1
3-e	2.189	2.736	1.805	1.802	1.802	1.805	2.092	2.068	2.092	2.068	2.092	2.068	27.2
4-a	1.469	3.005	1.963	1.962	1.963	1.962	2.061	2.061	2.022	2.022	2.022	2.022	61.3
4-b	1.474	3.052	1.988	1.952	1.952	1.988	2.037	2.037	2.030	2.032	2.030	2.032	61.3
4-c	1.450	3.291	1.941	1.895	1.895	1.941	2.127	2.127	1.989	2.002	1.989	2.002	47.8
4-d	1.657	3.126	1.843	1.843	1.843	1.843	2.258	2.258	1.983	1.983	1.983	1.983	36.3
4-e	2.361	2.542	1.787	1.787	1.787	1.787	2.377	2.377	1.934	1.934	1.934	1.934	27.2
5-a	1.461	3.063	2.011	1.935	1.935	2.011	2.038	2.038	2.033	2.052	2.033	2.052	62.7
5-b	1.442	3.158	1.969	1.932	1.932	1.969	2.044	2.044	2.032	2.038	2.032	2.038	57.6
5-c	1.436	3.326	1.975	1.885	1.885	1.975	2.102	2.102	2.003	2.032	2.003	2.032	48.7
5-d	1.678	3.101	1.836	1.836	1.836	1.836	2.229	2.229	1.998	1.998	1.998	1.998	36.0
5-e	2.322	2.565	1.784	1.784	1.784	1.784	2.330	2.330	1.951	1.951	1.951	1.951	27.7
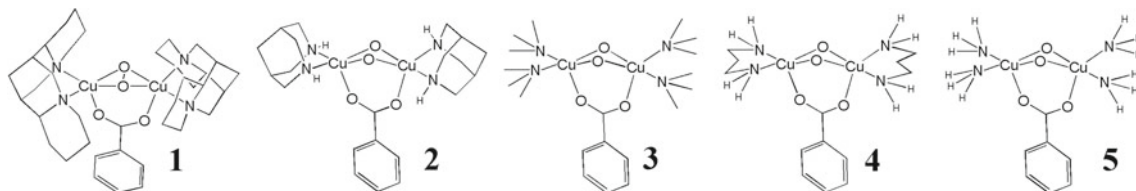
Scheme 2 Schematic representation of the butterfly μ–η^2:η^2-peroxo dicopper core structures **1–5** (for the sake of clarity, all H atoms have been omitted except for the H atoms on the N atoms)

Differently from the recent proposal that the axial bonding on the Cu centres produces just an electronic effect [46], the data in Table 2 reveal that for systems **2**, **4**, and **5,** the bis(μ-oxo) isomer is more stable than any of the μ–η^2:η^2-peroxo-Cu(II)$_2$ species. On the other hand, system **1-e** is a good example that the higher the steric hindrance produced by the diaza ligand the higher the stability of the butterfly μ–η^2:η^2-peroxo-Cu(II)$_2$ species. This means that the external rings of the αSp ligand render difficult the bonding of the entering carboxylate, facilitating the conversion of the bis(μ-oxo) species **1'-e** to the butterfly μ–η^2:η^2-peroxo-Cu(II)$_2$ species **1-a** overcoming a low barrier of 6.1 kcal mol^{-1}. In summary, the diaza amine ligand has an important role, that is, the larger the steric hindrance of the diaza ligand the lower the probability to display the bis(μ-oxo) isomer.

Table 3 collects geometrical data for the structure of **1'–5'**, which are homologous to **1–5** by removal of the benzoate counterion. This increases the molecular symmetry, thus, the Cu2–O1 and Cu2–O2 bond distances are identical in terms of symmetry to the Cu1–O2 and Cu1–O1 ones, respectively, as well as Cu2–N3 and Cu2–N4 with respect to Cu1–N1 and Cu1–N2. Furthermore, according to the O2–Cu1–O1–Cu2 dihedral angle, the bending of the butterfly μ–η^2:η^2-peroxo-Cu(II)$_2$ species decreases. In the absence of the carboxylate group the bis(μ-oxo) isomeric form is remarkably stabilized, in correspondence with all previous works [5, 21, 130, 131], since the absence of a third ligand bonded to each metallic centre results in a shortening of all the Cu–N bond distances. On the other hand, according to the O2–Cu1–O1–Cu2 dihedral angle, the biradical singlet μ–η^2:η^2-peroxo-Cu(II)$_2$ species lose

Table 2 Computed cores for complexes **1–5** (energies in kcal mol^{-1} relative to the corresponding triplet μ–η^2:η^2-O$_2$ structure)

	a	b	c	d	e
1	0.0	0.2	16.1	25.1	19.0
2	0.0	0.1	16.4	20.4	−4.9
3	0.0	0.2	15.6	25.7	20.4
4	0.0	0.4	15.8	17.9	−6.4
5	0.0	0.6	16.4	17.2	−3.2

Table 3 Geometrical parameters for the computed cores for species **1'–5'** with the benzoate anion (distances in Å and angles in degrees)

	O1–O1	Cu1–Cu2	Cu1–O1	Cu1–O2	Cu1–N1	Cu1–N2	O2–Cu1–O1–Cu2
1-a'	1.498	3.354	1.963	1.939	1.983	1.982	45.9
1-b'	1.517	3.531	1.920	1.923	1.975	1.978	1.4
1-c'	1.493	3.486	1.896	1.896	1.975	1.975	0.0
1-d'	1.665	3.301	1.848	1.848	1.980	1.980	0.0
1-e'	2.256	2.784	1.791	1.793	1.981	1.978	1.8
2-a'	1.511	3.277	1.932	1.932	1.963	1.963	47.1
2-b'	1.537	3.469	1.898	1.898	1.955	1.955	4.6
2-c'	1.530	3.414	1.871	1.871	1.952	1.952	0.6
2-d'	1.752	3.244	1.843	1.843	1.954	1.954	0.3
2-e'	2.368	2.666	1.783	1.783	1.916	1.916	0.0
3-a'	1.480	3.376	1.964	1.951	2.014	2.014	46.1
3-b'	1.500	3.529	1.933	1.914	2.006	2.005	13.1
3-c'	1.473	3.507	1.905	1.899	2.007	2.000	2.0
3-d'	1.706	3.261	1.843	1.838	2.022	2.012	1.2
3-e'	2.187	2.843	1.799	1.788	2.036	2.023	1.6
4-a'	1.497	3.280	1.934	1.934	1.973	1.973	47.8
4-b'	1.522	3.467	1.899	1.899	1.966	1.966	11.8
4-c'	1.501	3.438	1.876	1.876	1.961	1.961	1.6
4-d'	1.645	3.284	1.837	1.837	1.963	1.963	1.4
4-e'	2.322	2.695	1.779	1.779	1.931	1.931	1.2
5-a'	1.484	3.299	1.933	1.934	1.992	1.990	47.2
5-b'	1.511	3.479	1.896	1.896	1.983	1.983	0.0
5-c'	1.481	3.453	1.878	1.878	1.978	1.978	0.0
5-d'	1.683	3.238	1.825	1.825	1.979	1.979	0.0
5-e'	2.267	2.730	1.774	1.774	1.950	1.950	0.0

the butterfly shape and become further more stable than the still butterfly-shaped triplet species. Nevertheless, the bis(μ-oxo) isomer is the most stable species for structures **1'–5'**, except for **3'**, for which the permethylation of the amine ligands destabilizes the square planar Cu(III) structure, making species **3'-e** 1.8 kcal mol^{-1} less stable than the isomeric μ–η^2:η^2-peroxo-Cu(II)$_2$ species **3'-b**. The reason is that the N1-Cu1–N2 angle in species **3'-b** is 113.2°, and 106.0° in species **3'-e**. The large angle for the diradical singlet species precludes, or at least reduces some repulsive interactions between the hydrogen atoms of the methyl groups (Table 4).

Binding of the benzoate axially to each Cu centre in structures **1'–5'** is strongly favored, see Table 5. Benzoate binding is stronger for the triplet and diradical singlet μ–η^2:η^2-peroxo-Cu(II)$_2$ species with respect to the (μ-O)$_2$–Cu(III)$_2$. The bending caused by this anionic ligand in part explains this destabilization of the latter isomeric species.

Analysis in the framework of the frontier molecular orbital theory (FMO) allows to describe easily the interaction of complexes **1'–5'** when bonding to O$_2$. The O$_2$ binding requires that the dx^2-y^2 orbital of both Cu atoms perpendicular to the Cu–Cu axis rotates in order to donate two electrons into the π^* orbital set of O$_2$. This interaction means the reduction from O$_2$ to O$_2^{2-}$, that is, peroxide, and the orbital resulting from this interaction corresponds to the LUMO. For **1'-a** there is an interaction of the triplet dioxygen moiety and the singlet ground state of [Cu(I)$_2$(αSp)$_2$]$^{2+}$, with the two-electron transfer and thus completing the intersystem crossing. Differently, in species **1-a**, this two-electron transfer is not fulfilled because of the benzoate counteranion.

Due to the butterfly-shaped structure, the peroxide in species **1-a** is raised above the molecular plane, displaying a triplet multiplicity as ground state. However, the singlet–triplet splitting is feasible due to not a large energy gap between the triplet and the open-shell singlet. As the peroxide is displaced out of the core plane the π_σ^*, that is, the typical LUMO for the side-on peroxo structures, becomes less σ bonding and the π_v^* takes an important role in σ overlap. Furthermore the π^* orbitals on the peroxide bridge become close in energy which allows an orientation along

the Cu–η^2–O$_2^{2-}$ planes. This means that the superexchange coupling between the two Cu(II) atoms is lower and thus the antiferromagnetic interaction is reduced as well. There is a spin-up electron in each orthogonal π^* orbital of the peroxide which correlates with the triplet ground state of O$_2$, stabilizing the triplet ground state by single-centre electron exchange. This effect is higher in **1-a** with respect to that usually calculated for binding of O$_2$ to the deoxy complex. This effect is based on the charge transfer from the copper centres that shifts the unpaired spin density on each copper atom thus reducing the exchange stabilization of the triplet ground state of O$_2$. Then the antiferromagnetic coupling stabilizes the singlet state. This effect is stronger in species **1'-a** in which it is possible to overcome the spin-forbidden O$_2$ binding by the deoxy dicopper complex.

Insights on the interaction of the carboxylate groups and the copper atoms in structures **1–5** and **1'–5'** are achieved by analysis of the HOMO and LUMO. According to Table 6, the influence of the benzoate counterion is a key because both frontier molecular orbitals have been destabilized (see Fig. 3), but being this effect slightly higher for the HOMO. It is necessary to point out that when bonding this anionic species, the copper centres increase their

Table 4 Computed cores for complexes **1'-5'** (energies in kcal·mol^{-1} relative to the corresponding triplet $\mu-\eta^2{:}\eta^2$–O$_2$ structure)

	a	b	c	d	e
1'	0.0	−4.0	1.2	5.0	−6.3
2'	0.0	−6.0	0.6	0.1	−20.8
3'	0.0	−4.6	−1.8	5.6	1.8
4'	0.0	−6.1	0.2	3.2	−15.3
5'	0.0	−5.0	1.2	6.3	−9.0

Table 5 Energy release when bonding benzoate to species **1'-5'** (energies in kcal mol^{-1} relative to the corresponding triplet $\mu-\eta^2{:}\eta^2$-O$_2$ structure)

	a	b	c	d	e
1'	45.0	40.8	30.1	24.9	19.7
2'	36.9	30.8	21.1	16.6	21.0
3'	52.5	47.7	35.1	32.4	33.9
4'	32.3	25.8	16.7	17.6	23.4
5'	31.2	25.6	16.0	20.3	25.4

coordination numbers from 4 to 5 and for the overall charge molecule is +1.

Despite the different multiplicity states that generate species **a–e**, and different ligands in systems **1–5** and also **1'–5'**, the type of orbitals for each studied species is the same. Thus, for the sake of simplicity, we discuss the relative stability of the orbitals from systems **1-e** and **1'-e** (see Fig. 3). The HOMO and LUMO orbitals are the common b$_{1g}$ and b$_{2u}$ orbitals, being the expected b$_{3g}$ orbital as LUMO isoenergetic and only 0.006 eV far in energy with respect to the b$_{2u}$ orbital, in **1-e** and **1'-e**, respectively.

The different overall charge +2 and +1, before and after bonding the benzoate, respectively, easily explains the destabilization of the frontier orbitals after the addition of anionic counterion. The entering benzoate group requires the loose of the planarity of the [Cu$_2$O$_2$]$^{2+}$ core according to the O2-Cu1-O1-Cu2 dihedral angle, going from 2° to 24° for structure **1**. This bending represents a destabilization of both frontier molecular orbitals because of its antibonding character. The gap in energy from the HOMO before and after the bonding of the benzoate is filled basically in part by the π orbitals corresponding to the aromatic ring of the benzoate that do not stabilize the structure because of their non-bonding character (see Fig. 4) [140]. They are placed near in energy with respect to the HOMO. Finally, the higher destabilization of the HOMO with respect to the LUMO when bonding the benzoate group, 0.020 eV see Fig. 3, can be explained by the additional antibonding interaction between benzoate orbitals and the HOMO.

Table 6 Energy of the HOMO (H) and LUMO (L) for complexes **1–5** and **1'–5'** (in eV)

	a		b		e	
	H	L	H	L	H	L
1	−0.291	−0.167	−0.291	−0.160	−0.305	−0.200
2	−0.298	−0.163	−0.297	−0.159	−0.307	−0.201
3	−0.312	−0.182	−0.310	−0.175	−0.315	−0.216
4	−0.308	−0.169	−0.306	−0.164	−0.322	−0.215
5	−0.321	−0.182	−0.317	−0.176	−0.328	−0.236
1'	−0.446	−0.315	−0.435	−0.297	−0.456	−0.330
2'	−0.437	−0.327	−0.455	−0.308	−0.478	−0.346
3'	−0.452	−0.348	−0.469	−0.330	−0.492	−0.366
4'	−0.463	−0.349	−0.480	−0.329	−0.507	−0.377
5'	−0.494	−0.376	−0.508	−0.357	−0.543	−0.415
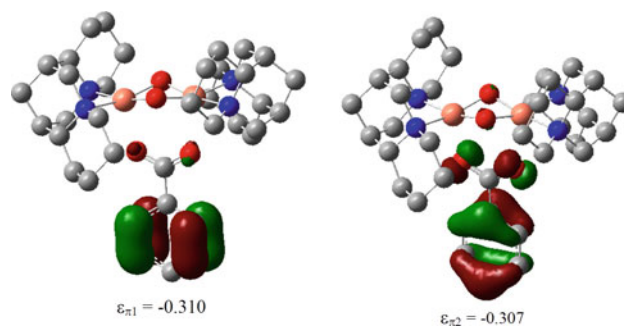
Fig. 4 Computed aromatic π_1 and π_2 orbitals for species **1-e** (energies in eV). For the sake of clarity, all H atoms have been omitted

Repeating the analysis for the butterfly $\mu-\eta^2{:}\eta^2$-peroxo isomer **1-a** indicates that there are qualitative differences (see Fig. 5). For the butterfly-shaped [Cu$_2$O$_2$]$^{2+}$ core, the HOMO is b$_{3g}$ and the LUMO displays a b$_{1g}$ type of orbital, while in species **1-a** was the HOMO. The main difference is that there is not an antibonding contribution of the benzoate in the HOMO, and at the end this factor turns out to be essential because now it is the LUMO the orbital more destabilized when bonding to the benzoate, by 0.021 eV.

Knowing a priori that peroxo systems are predicted as electrophilic species [37, 141], electrophilicity (ω) and chemical hardness calculations (η) of these compounds using the parameter defined by Parr et al. (see Eqs. 2, 4) [100, 142–144] were envisaged. Focusing on the data in Table 7, it is possible to confirm that the butterfly $\mu-\eta^2{:}\eta^2$-peroxo-Cu(II)$_2$ species are less electrophilic than the planar $\mu-\eta^2{:}\eta^2$-peroxo-Cu(II)$_2$ ones in agreement with Solomon et al. [127]. There is the outstanding trend that the bis(μ-oxo) isomer is harder when removing the bonded carboxylate counterion, which helps to understand why there is no any bis(μ-oxo) species

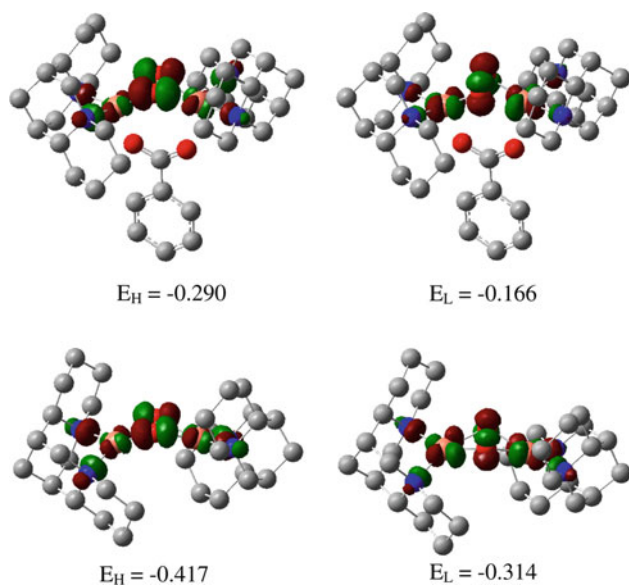$E_H = -0.290$ $E_L = -0.166$

$E_H = -0.417$ $E_L = -0.314$

Fig. 5 Computed HOMO and LUMO orbitals for species **1-a** and **1'-a** (energies in eV). For the sake of clarity, all H atoms have been omitted

with a coordinated axial ligand. On the other hand, there is the general trend that the inclusion of the benzoate decreases greatly the electrophilic character of the molecule, becoming less than the half. However, according to the electrophilic character of peroxo species, the bis(μ-oxo) species is the most electrophilic. It is important to point out that the low value of electrophilicity for the μ–η^2:η^2-peroxo-Cu(II)$_2$ species, either triplet or diradical singlet, is in agreement with the higher reactivity of the bis(μ-oxo) species with external reactants. This hypothesis is reinforced by the higher chemical hardness displayed by these μ–η^2:η^2-peroxo-Cu(II)$_2$ species, thus the HOMO–LUMO gap is larger. Overall, the interaction from a bis(μ-oxo) species the butterfly peroxo one is favored with the presence of a counterion.

4 Conclusions

In conclusion, DFT calculations indicate that for the butterfly μ–η^2:η^2-peroxo species **1-a** studied here, the ground state is the triplet instead of the diradical singlet one, although the difference in stability is small. On the other hand, the presence of the bis(μ-oxo) isomer **1-e** could not be totally ruled out because without the carboxylate counteranion this isomer **1-e'** is found to be the most stable species. Probably this is due to the inclusion of the counterions that guide to a neutral charged complex, that is, SbF$_6^-$ for this cationic [Cu(II)$_2$(αSp)$_2$(μ–η^2:η^2-O$_2$)(Bz$^-$)]$^+$ complex [46].

The role of the carboxylate is a key because its binding axially on the Cu centre induces a strong electronic effect, precluding the O–O bond cleavage. However, the steric effect is relevant as well, since simplifying the diaza amine group the bis(μ-oxo) isomer gains stability, becoming the most stable isomer in complexes **2**, **4**, and **5**.

DFT calculations have revealed correctly the trends in the energy difference between the two isomers upon changes in the diaza ligand despite it is generally assumed that B3LYP calculations somewhat underestimate the stability of the bis-μ oxo isomer [114, 145–148]. It is important to stress that, first, this isomeric species are not far in energy and, second, the relative energies between all these isomers are highly dependent on the method of calculation. For these two reasons, the possibility that diradical multiplicity species is the real species cannot be completely ruled out. However, our conclusion is supported by recent work of Solà et al. [62] that have suggested that B3LYP is extremely accurate for similar complexes. Noteworthy, the key concept highlighted in this work is that the calculations done previously for complex **1-a** led to a non-accurate description of the system [46]. The strategy displayed here suggests that it is convenient to

Table 7 Chemical harness and electrophilicity for complexes **1–5** and **1'–5'** (in eV)

	a		b		c		d		e	
	η	ε	η	ε	η	ε	η	ε	η	ε
1	0.062	1.689	0.065	1.556	0.035	2.872	0.047	2.409	0.053	2.422
2	0.068	1.568	0.069	1.500	0.036	2.782	0.044	2.444	0.053	2.435
3	0.065	1.883	0.068	1.748	0.038	3.103	0.049	2.672	0.049	2.862
4	0.070	1.632	0.071	1.555	0.039	2.847	0.049	2.477	0.054	2.692
5	0.070	1.815	0.070	1.732	0.039	3.204	0.052	2.721	0.046	3.476
1'	0.066	4.414	0.069	3.877	0.047	5.588	0.052	5.231	0.063	4.922
2'	0.055	5.332	0.073	3.964	0.051	5.681	0.063	5.052	0.066	5.138
3'	0.052	6.167	0.069	4.594	0.051	6.254	0.057	5.748	0.063	5.815
4'	0.057	5.771	0.076	4.335	0.055	5.979	0.059	5.742	0.065	5.997
5'	0.059	6.444	0.075	4.960	0.059	6.483	0.065	6.150	0.064	7.153

explore always all possible multiplicity states to predict the right ground state, being a beneficial strategy to characterize better the interaction of dicopper complexes with O_2. New experimental and computational insight is required to reach a better comprehension about the electronic ground state of these butterfly peroxo structures, as well as to describe more accurately the steric and electronic effects of each species. In our laboratory, we are undergoing additional studies with the relatively new M06 [149].

5 Supporting information available

Complete computational methods used and xyz coordinates; and full 3D drawings of all stationary points found. This material is free of charge.

References

1. Karlin KD, Gultneh Y (1987) Prog Inorg Chem 35:219–327
2. Sorrell TN (1989) Tetrahedron 45:3–68
3. Karlin KD, Tyeklar Z (1994) Adv Inorg Biochem 9:123–172
4. Solomon EI, Lowery MD (1993) Science 259:1575–1581
5. Que L Jr, Tolman WB (2002) Angew Chem Int Ed 41:1114–1137
6. Holland PL, Tolman WB (1999) Coord Chem Rev 192:855–869
7. Kitajima N, Moro-oka Y (1994) Chem Rev 94:737–757
8. Itoh S, Taki M, Nakao H, Holland PL, Tolman WB, Que L Jr, Fukuzumi S (2000) Angew Chem Int Ed 39:398–400
9. Balasubramanian R, Rosenzweig AC (2007) Acc Chem Res 40:573–580
10. Rolff M, Tuczek F (2008) Angew Chem Int Ed 47:2344–2347
11. Palavicini S, Granata A, Monzani E, Casella L (2005) J Am Chem Soc 127:18031–18036
12. Battaini G, Monzani E, Perotti A, Para C, Casella L, Santagostini L, Gullotti M, Dillinger R, Näther C, Tuczek F (2003) J Am Chem Soc 125:4185–4198
13. Santagostini L, Gullotti M, Monzani E, Casella L, Dillinger R, Tuczek F (2000) Chem Eur J 6:519–522
14. Karlin KD, Nasir MS, Cohen BI, Cruse RW, Kaderli S, Zuberbühler AD (1994) J Am Chem Soc 116:1324–1336
15. Mahapatra S, Kaderli S, Llobet A, Neuhold Y-M, Palanche T, Halfen JA, Young VG Jr, Kaden TA, Que L Jr, Zuberbuhler AD, Tolman WB (1997) Inorg Chem 36:6343–6356
16. Yamazaki S, Itoh S (2003) J Am Chem Soc 125:13034–13035
17. Granata A, Monzani E, Bubacco L, Casella L (2006) Chem Eur J 12:2504–2514
18. De A, Mandal S, Mukherjee R (2008) J Inorg Biochem 102:1170–1189
19. Stack TDP (2003) Dalton Trans 10:1881–1889
20. Solomon EI, Sundaram UM, Machonkin TE (1996) Chem Rev 96:2563–2606
21. Mirica LM, Ottenwaelder X, Stack TDP (2004) Chem Rev 104:1013–1045
22. Solomon EI, Chen P, Metz M, Lee S-K, Palmer AE (2001) Angew Chem Int Ed 40:4570–4590
23. Karlin KD, Kaderli S, Zuberbühler AD (1997) Acc Chem Res 30:139–147
24. Decker H, Dillinger R, Tuczek F (2000) Angew Chem Int Ed 39:1591–1595
25. Lind T, Siegbahn PEM, Crabtree RH (1999) J Phys Chem B 103:1193–1202
26. Decker H, Schweikardt T, Tuczek F (2006) Angew Chem Int Ed 45:4546–4550
27. Siegbahn PEM (2003) J Biol Inorg Chem 8:567–576
28. Lewis EA, Tolman WB (2004) Chem Rev 114:1047–1076
29. Kitajima N, Fujisawa K, Morooka Y (1989) J Am Chem Soc 111:8976–8978
30. Kitajima N, Fujisawa K, Fujimoto C, Morooka Y, Hashimoto S, Kitagawa T, Torinmi K, Tatsumi K, Nakamura A (1992) J Am Chem Soc 114:1277–1291
31. Mahapatra S, Halfen JA, Wilinson EC, Pan G, Cramer CJ, Que L Jr, Tolman WB (1995) J Am Chem Soc 117:8865–8866
32. Matoba Y, Kumagai T, Yamamoto A, Yoshitsu H, Sugiyama M (2006) J Biol Chem 281:8981–8990
33. Lewin JL, Heppner DE, Cramer CJ (2007) J Biol Inorg Chem 12:1221–1234
34. Cramer CJ, Wloch M, Piecuch P, Puzzarini C, Gagliardi L (2006) J Phys Chem A 110:1991–2004
35. Cramer CJ, Wloch M, Piecuch P, Puzzarini C, Gagliardi L (2007) Erratum ibid. J Phys Chem A 111:4871
36. Rode MF, Werner H-J (2005) Theor Chem Acc 114:309–317
37. Costas M, Ribas X, Poater A, López-Valbuena JM, Xifra R, Company A, Duran M, Solà M, Llobet A, Corbella M, Usón MA, Mahía J, Solans X, Shan X, Benet-Buchholz J (2006) Inorg Chem 45:3569–3581
38. Menif R, Martell AE, Squattrito PJ, Clearfield A (1990) Inorg Chem 29:4723–4729
39. Malmqvist PA, Pierloot K, Shahi ARM, Cramer CJ, Gagliardi L (2008) J Chem Phys 128:204109–204110
40. Halfen JA, Mahapatra S, Wilkinson EC, Kaderli S, Young VG Jr, Que L Jr, Zuberbühler AD, Tolman WB (1996) Science 271:1397–1400
41. Mahapatra S, Halfen JA, Tolman WB (1996) J Am Chem Soc 118:11575–11586
42. Osako T, Tachi Y, Taki M, Fukuzumi S, Itoh S (2001) Inorg Chem 40:6604–6609
43. Hatcher LQ, Vance MA, Narducci Sarjeant AA, Solomon EI, Karlin KD (2006) Inorg Chem 45:3004–3013
44. Mahadevan V, Hodgson KO, Stack TDP (1999) J Am Chem Soc 121:5583–5584
45. Kodera M, Kajita Y, Tachi Y, Katayama K, Kano K, Hirota S, Fujinami S, Suzuki M (2004) Angew Chem Int Ed 43:334–337
46. Ottenwaelder X, Rudd DJ, Corbett MC, Hodgson KO, Hedman B, Stack TDP (2006) J Am Chem Soc 128:9268–9269
47. Funahashi Y, Nishikawa T, Wasada-Tsutsui Y, Kajita Y, Yamaguchi S, Arii H, Ozawa T, Jitsukawa K, Tosha T, Hirota S, Kitagawa T, Masuda H (2008) J Am Chem Soc 130:16444–16445
48. Zhang CX, Liang H-C, Kim E-I, Shearer J, Helton ME, Kim E, Kaderli S, Incarvito CD, Zuberbühler AD, Rheingold AL, Karlin KD (2003) J Am Chem Soc 125:634–635
49. Tolman WB (1997) Acc Chem Res 30:227–237
50. Mahadevan V, Henson MJ, Solomon EI, Stack TDP (2000) J Am Chem Soc 122:10249–10250
51. Taki M, Itoh S, Fukuzumi S (2001) J Am Chem Soc 123:6203–6204

52. Taki M, Teramae S, Nagatomo S, Tachi Y, Kitagawa T, Itoh S, Fukuzumi S (2002) J Am Chem Soc 124:6367–6377
53. Utz D, Kisslinger S, Heinemann FW, Hampel F, Schindler S (2011) Eur J Inorg Chem 255–267
54. Garcia-Bosch I, Company A, Frisch JR, Torrent-Sucarrat M, Cardellach M, Gamba I, Güell M, Casella L, Que L Jr, Ribas X, Luis JM, Costas M (2010) Angew Chem Int Ed 49:2406–2409
55. Zapata-Rivera J, Caballol R, Calzado CJ (2011) Phys Chem Chem Phys 13:20241–20247
56. Yoshizawa K, Shiota Y (2006) J Am Chem Soc 128:9873–9881
57. Yoshizawa K, Suzuki A, Shiota Y, Yamabe T (2000) Bull Chem Soc Jpn 73:815–827
58. Poater A, Cavallo L (2009) Inorg Chem 48:4062–4066
59. de la Lande A, Parisel O, Gérard H, Moliner V, Reinaud O (2008) Chem Eur J 14:6465–6473
60. Yoshizawa K, Kihara N, Kamachi T, Shiota Y (2006) Inorg Chem 45:3034–3041
61. Spuhler P, Holthausen MC (2003) Angew Chem Int Ed 42:5961–5965
62. Poater A, Ribas X, Llobet A, Cavallo L, Solà M (2008) J Am Chem Soc 130:17710–17717
63. Sander O, Henß A, Näther C, Würtele C, Holthausen MC, Schindler S, Tuczek F (2008) Chem Eur J 14:9714–9729
64. Herres-Pawlis S, Verma P, Haase R, Kang P, Lyons CT, Wasinger EC, Flörke U, Henkel G, Stack TDP (2009) J Am Chem Soc 131:1154–1169
65. Lam BMT, Halfen JA, Young VG Jr, Hagadorn JR, Holland PL, Lledós A, Cucurull-Sánchez L, Novoa JJ, Álvarez S, Tolman WB (2000) Inorg Chem 39:4059–4072
66. Naka N, Kondo Y, Usui S, Hashimoto Y, Uchiyama M (2007) Adv Synth Catal 349:595–600
67. Itoh K, Hayashi H, Furutachi H, Matsumoto T, Nagatomo S, Tosha T, Tereda S, Fujinami S, Suzuki M, Kitagawa T (2005) J Am Chem Soc 127:5212–5223
68. Inoue T, Shiota Y, Yoshizawa K (2008) J Am Chem Soc 130:16890–16897
69. Poater A (2009) J Phys Chem A 113:9030–9040
70. Herres-Pawlis S, Verma P, Haase R, Kang P, Lyons CT, Wasinger EC, Flrke U, Henkel G, Stack TDP (2009) J Am Chem Soc 131:1154–1169
71. Mirica LM, Vance M, Jackson-Rudd D, Hedman B, Hodgson KO, Solomon EI, Stack TDP (2005) Science 308:1890–1892
72. Company A, Palavicini S, Garcia-Bosch I, Mas-Ballesté R, Que L Jr, Rybak-Akimova EV, Casella L, Ribas X, Costas M (2008) Chem Eur J 14:3535–3538
73. Gherman BF, Cramer CJ (2008) Coord Chem Rev 253:723–753
74. Hong S, Huber SM, Gagliardi L, Cramer CJ, Tolman WB (2007) J Am Chem Soc 129:14190–14192
75. Cramer CJ, Kinal A, Wloch M, Piecuch P, Gagliardi L (2006) J Phys Chem A 110:11557–11568
76. Bernardi F, Bottoni A, Casadio R, Fariselli P, Rigo A (1996) Inorg Chem 35:5207–5212
77. 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 Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09, Revision A.1. Gaussian Inc., Wallingford CT
78. Becke AD (1993) J Chem Phys 98:5648–5652
79. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785–789
80. Stevens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) J Phys Chem 98:11623–11627
81. Hehre WJ, Ditchfield R, Pople JA (1972) J Chem Phys 56:2257–2261
82. Hehre WJ, Radom L, PVR Schleyer, Pople JA (1986) Ab initio molecular orbital theory. Wiley, New York
83. Wachters AJH (1970) J Chem Phys 52:1033–1036
84. Güell M, Luis JM, Rodríguez-Santiago L, Sodupe M, Solà M (2008) J Phys Chem A 113:1308–1317
85. Caballol R, Castell O, Illas F, Moreira IDPR, Malrieu JP (1997) J Phys Chem A 101:7860–7866
86. Winkler M (2005) J Phys Chem A 109:1240–1246
87. Lindh R, Bernhardsson A, Schütz M (1999) J Phys Chem A 103:9913–9920
88. Cramer CJ (1999) J Chem Soc Perkin Trans 2:2273–2283
89. Kikuchi A, Ito H, Abe J (2005) J Phys Chem B 109:19448–19453
90. Borden WT (1998) Diradicals. In: PvR Schleyer, Allinger NL, Clark T, Gasteiger J, Kollman PA, Schaeffer HF III (eds) The encyclopedia of computational chemistry. Wiley, Chichester, pp 708–722
91. Poater J, Bickelhaupt FM, Solà M (2007) J Phys Chem A 111:5063–5070
92. Gräfenstein J, Kraka E, Filatov M, Cremer D (2002) Int J Mol Sci 3:360–394
93. Borden WT, Davidson ER (1996) Acc Chem Res 29:67–75
94. Ziegler T, Rauk A, Baerends EJ (1977) Theor Chim Acta 43:261–271
95. Yamaguchi K, Jensen F, Dorigo A, Houk KN (1988) Chem Phys Lett 149:537–542
96. Lim MH, Worthington SE, Dulles FJ, Cramer CJ (1996) In: Laird BB, Ross RB, Ziegler T (eds) Chemical applications of density functional theory, vol 629. Washington, DC, American Chemical Society, p 402
97. Isobe H, Takano Y, Kitagawa W, Kawakami T, Yamanaka S, Yamaguchi K, Houk KN (2002) Mol Phys 100:717–727
98. Barone V, Cossi M (1998) J Phys Chem A 102:1995–2001
99. Tomasi J, Persico M (1994) Chem Rev 94:2027–2094
100. Parr RG, von Szentpaly L, Liu S (1999) J Am Chem Soc 121:1922–1924
101. Geerlings P, De Proft F, Langenaeker W (2003) Chem Rev 103:1793–1873
102. Parr RG, Yang W (1989) Density Functional Theory of Atoms and Molecules. Oxford University Press, New York
103. Parr RG, Donnelly RA, Levy M, Palke WE (1978) J Chem Phys 68:3801–3807
104. Parr RG, Pearson RG (1983) J Am Chem Soc 105:7512–7516
105. Koopmans T (1934) Physica 1:104–113
106. Sala X, Poater A, von Zelewsky A, Parella T, Fontrodona X, Romero I, Solà M, Rodríguez M, Llobet A (2008) Inorg Chem 47:8016–8024
107. Mola J, Rodríguez M, Romero I, Llobet A, Parella T, Poater A, Solà M, Benet-Buchholz J (2006) Inorg Chem 45:10520–10529
108. Sala X, Plantalech E, Romero I, Rodríguez M, Llobet A, Poater A, Duran M, Solà M, Jansat S, Gómez M, Parella T, Stoeckli-Evans H, Benet-Buchholz J (2006) Chem Eur J 12:2798–2807
109. Manzini S, Urbina-Blanco CA, Poater A, Slawin AMZ, Cavallo L, Nolan SP (2011) Angew Chem Int Ed 51:1042–1045
110. Poater A, Moradell S, Pinilla E, Poater J, Solà M, Martínez MA, Llobet A (2006) Dalton Trans 1188–1196
111. Duran J, Polo A, Real J, Benet-Buchholz J, Poater A, Solà M (2003) Eur J Inorg Chem 4147–4151

112. Rich J, Rodríguez M, Romero I, Fontrodona X, van Leeuwen PWNM, Freixa Z, Sala X, Poater A, Solà M (2013) Eur J Inorg Chem. doi:10.1002/ejic.201201154
113. Siegbahn PEM (2003) J Biol Inorg Chem 8:577–585
114. Güell M, Siegbahn PEM (2007) J Biol Inorg Chem 12:1251–1264
115. Poater A, Cavallo L (2009) Inorg Chem 48:2340–2342
116. Güell M, Luis JM, Solà M, Siegbahn PEM (2009) J Biol Inorg Chem 14:229–242
117. Ribas X, Calle C, Poater A, Casitas A, Gómez L, Xifra R, Parella T, Benet-Buchholz J, Schweiger A, Mitrikas G, Solà M, Llobet A, Stack TDP (2010) J Am Chem Soc 132:12299–12306
118. Casitas A, Poater A, Solà M, Stahl SS, Costas M, Ribas X (2010) Dalton Trans 39:10458–10463
119. Becke AD (1993) J Chem Phys 98:1372–1377
120. Braïda B, Hiberty PC, Savin A (1998) J Phys Chem A 102:7872–7877
121. Sodupe M, Bertran J, Rodríguez-Santiago L, Baerends EJ (1999) J Phys Chem A 103:166–170
122. Chermette H, Ciofini I, Mariotti F, Daul C (2001) J Chem Phys 115:11068–11079
123. Grüning M, Gritsenko OV, van Gisbergen SJA, Baerends EJ (2001) J Phys Chem A 105:9211–9218
124. Poater J, Solà M, Rimola A, Rodríguez-Santiago L, Sodupe M (2004) J Chem Phys A 108:6072–6078
125. Metz M, Solomon EI (2001) J Am Chem Soc 123:4938–4950
126. Reiher M, Salomon O, Hess BA (2001) Theor Chem Acc 107:48–55
127. Salomon O, Reiher M, Hess BA (2002) J Chem Phys 117:4729–4737
128. Blackburn NJ, Strange RW, Farooq A, Haka MS, Karlin KD (1988) J Am Chem Soc 110:4263–4272
129. Pidcock E, Obias HV, Abe M, Liang HC, Karlin KD, Solomon EI (1999) J Am Chem Soc 121:1299–1308
130. Company A, Gómez L, Mas-Ballesté R, Korendovych IV, Ribas X, Poater A, Parella T, Fontrodona X, Benet-Buchholz J, Solà M, Que L Jr, Rybak-Akimova EV, Costas M (2007) Inorg Chem 46:4997–5012
131. Company A, Gómez L, Mas-Ballesté R, Korendovych IV, Ribas X, Poater A, Parella T, Fontrodona X, Benet-Buchholz J, Solà M, Que L Jr, Rybak-Akimova EV, Costas M (2006) Inorg Chem 45:5239–5241
132. Kitajima N, Fujisawa K, Moro-oka Y (1989) J Am Chem Soc 111:8975–8976
133. Kodera M, Katayama K, Tachi Y, Kano K, Hirota S, Fujinami S, Suzuki M (1999) J Am Chem Soc 121:11006–11007
134. Funahashi Y, Nakaya K, Hirota S, Yamauchi O (2000) Chem Lett 1172–1173
135. Okumura T, Hayami S, Ozawa T, Funahashi Y, Maeda Y, Masuda H (2007) Chem Lett 96–97
136. Boschmann E, Weinstock LM, Carmack M (1974) Inorg Chem 13:1297–1300
137. Choi S-N, Bereman RD, Wasson JR (1975) J Inorg Nucl Chem 37:2087–2090
138. Kuroda R, Mason SF (1979) J Chem Soc, Dalton Trans 1979:727–730
139. Lever ABP (1984) Inorganic Electronic Spectroscopy, 2nd edn. Elsevier, Amsterdam-Oxford-New York-Tokyo
140. Xifra R, Ribas X, Llobet A, Poater A, Duran M, Solà M, Stack TDP, Benet-Buchholz J, Donnadieu B, Mahía J, Parella T (2005) Chem Eur J 11:5146–5156
141. Hatcher LQ, Karlin KD (2004) J Biol Inorg Chem 9:669–683
142. Poater A, Saliner AG, Cavallo L, Poch M, Solà M, Worth AP (2012) Curr Med Chem 19:5219–5225
143. Poater A, Saliner AG, Solà M, Cavallo L, Worth AP (2010) Expert Opin Drug Deliv 7:295–305
144. Poater A, Saliner AG, Carbó-Dorca R, Poater J, Solà M, Cavallo L, Worth AP (2009) J Comput Chem 30:275–284
145. Bérces A (1997) Inorg Chem 35:4831–4837
146. Flock M, Pierloot K (1999) J Phys Chem A 103:95–102
147. Rode MF, Werner H-J (2005) Theor Chem Acc 114:309–317
148. Aullón G, Gorun SM, Álvarez S (2006) Inorg Chem 45:3594–3601
149. Zhao Y, Truhlar DG (2008) Theor Chem Acc 120:215–241

Theoretical study of the benzoquinone–tetrathiafulvalene–benzoquinone triad in neutral and oxidized/reduced states

Joaquín Calbo · Juan Aragó · Enrique Ortí

Abstract This work presents a thorough DFT quantum-chemical characterization of the structural and electronic properties of the benzoquinone–tetrathiafulvalene–benzoquinone (Q–TTF–Q) triad in its neutral and oxidized/reduced states. The minimum-energy structure of the triad in its neutral state corresponds to a C_{2v} boat conformation, although the planar D_{2h} structure is also found to be a minimum-energy structure practically degenerate with the boat conformation. Upon oxidation, electrons are mainly extracted from the TTF backbone and the boat-shaped TTF structure evolves toward a planar D_{2h} and to a twisted D_2 conformation in the cation and dication, respectively. The theoretical characterization of the triad in its radical anion and dianion states is performed with several hybrid (B3LYP, PBE0 and BHHLYP), meta-hybrid (M06-2X) and long-range-corrected (CAM-B3LYP, ωB97X and ωB97XD) functionals. In gas phase, BHHLYP, M06-2X, CAM-B3LYP, ωB97X and ωB97XD functionals correctly predict the localized Q⁻–TTF–Q structure as the most stable structure, whereas the hybrid B3LYP and PBE0 functionals lead to a delocalized (Q–TTF–Q)⁻ structure. The incorporation of solvent effects is crucial to provide a qualitatively correct description of the anion species where the localized solutions are found to be minimum-energy structures for all the functionals evaluated. The (Q–TTF–Q)²⁻ dianion is predicted to be a biradicaloid species in its electronic ground state. Calculations predict that both Q⁻ → Q⁰ and TTF → Q⁰ charge-transfer electronic transitions are associated with the low-energy, broad absorption band experimentally observed for the localized Q⁻–TTF–Q anion.

Keywords DFT calculations · Mixed-valence compounds · Tetrathiafulvalene derivates · Acceptor–donor–acceptor triad

J. Calbo · J. Aragó · E. Ortí (✉)
Instituto de Ciencia Molecular, Universidad de Valencia, 46980 Paterna, Spain
e-mail: enrique.orti@uv.es

1 Introduction

Since the discovery of the first organic metal [1], the charge-transfer complex formed by the electron-donor tetrathiafulvalene (TTF) and the electron-acceptor tetracyano-*p*-quinodimethane (TCNQ), TTF and its derivatives have aroused a great interest as building blocks for low-dimensional organic conductors and superconductors [2, 3]. TTF derivatives have been also used as donor units covalently linked to acceptor compounds through molecular bridges to promote efficient photoinduced intramolecular electron transfer and to gain more insight into the fundamental electron-transfer process at the molecular scale [4]. In contrast, the behavior of the TTF-conjugated unit as a molecular wire between redox centers has been scarcely explored. In this context, an acceptor–donor–acceptor (A–D–A) fused triad, Q–TTF–Q, where the donor moiety is TTF and the acceptor unit is *p*-benzoquinone (Q), was synthesized (Fig. 1) [5, 6]. The acceptor *p*-benzoquinone fragments were incorporated because they are particularly useful organic redox centers and can form persistent radical anions upon reduction. The TTF bridge allows for the electronic coupling between the electron-acceptor Q units

giving rise to an organic mixed-valence (MV) system in the anion state. The electrochemical and spectroscopic characterization of Q–TTF–Q revealed that the triad in its anion state behaves as a class II MV compound, according to the Robin-Day classification [7], where the extra electron is mainly located on one of the benzoquinone moieties [5]. Additionally, the study of the intramolecular electron-transfer process demonstrated that the TTF unit is able to act efficiently as an excellent π-conjugated bridge to assist the electron transfer between the Q redox centers [5].

MV compounds are of great relevance since they are excellent models for the fundamental understanding of intramolecular electron-transfer phenomena. Although the most common MV compounds involve metal atoms, such as the most prominent Creutz-Taube ion, an increasing number of purely organic MV compounds are being reported [8, 9]. Among these singular compounds, the Q–TTF–Q triad plays a fundamental role because it has been used as a model to obtain new MV compounds incorporating stronger and more sophisticated acceptor units [10]. The versatility of organic synthesis to create novel compounds with the desired properties opens the way toward the design of simpler MV compounds which allow us to gain a deeper knowledge about the mixed-valency phenomenon, specially for MV systems lying in the class II/III borderline [8, 9].

In this contribution, a thorough theoretical characterization of the Q–TTF–Q triad is performed by using quantum-chemical calculations. Emphasis is made on the evolution of the structural and electronic properties of the Q–TTF–Q triad upon oxidation/reduction. Special attention will be paid to the quantum-chemical characterization of anion MV species, which turns out to be a difficult task. For these species, unrestricted Hartree–Fock (UHF) calculations tend to provide structural and electron-density symmetry breaking and therefore lead to localized charge and spin densities even in those systems where the extra charge should be delocalized [11]. An accurate description of the MV species would require high-level ab initio wave function methods able to adequately deal with the electron correlation effect. Such methods, however, are computationally very demanding and prohibitive for medium-size molecular systems, and density functional theory (DFT) methods, which partially account for the electron

correlation effect, may be seen as a practical alternative. DFT functionals with high portions of Hartree–Fock (HF) exchange should be employed to reduce the self-interaction error and to avoid an excessive electron delocalization in molecular systems [12, 13]. Solvent effects can also play a pivotal role in the localized/delocalized theoretical description of the MV molecular system. The performance of different density functionals and the effect of the solvent are therefore discussed herein.

2 Computational details

DFT calculations were carried out using the Gaussian 09 program package (version A.02) [14]. Geometry optimizations of both the neutral molecules and the oxidized (cation and dication) and reduced (anion and dianion) species were first performed with Becke's three-parameter B3LYP exchange-functional [15, 16] and the 6-31G** basis set [17]. Calculations were done using the C_{2v} symmetry point group for the neutral species and the delocalized anion, and the C_s-symmetry for the localized anions. The cation and dication species were optimized within the D_{2h} and D_2 symmetries, respectively. The anion and the cation of Q–TTF–Q were calculated as open-shell doublet systems using the unrestricted UB3LYP approach. The localized anion was also computed as an open-shell doublet system within the unrestricted broken-symmetry (BS) approximation (UB3LYP(BS)). The dication was computed as a closed-shell system, whereas the dianion was treated both as a closed-shell and an open-shell system to disentangle its electronic nature. The reduced species were also optimized with the hybrid (PBE0 [18, 19] and BHHLYP [15]), meta-hybrid (M06-2X [20]) and the long-range-corrected (CAM-B3LYP [21], ωB97X [22] and ωB97XD [23]) functionals. To take into account the interaction with the solvent, both neutral and charged species were also optimized in the presence of the solvent, dimethylsulfoxide (DMSO), within the self-consistent reaction field (SCRF) theory using the polarized continuum model (PCM) approach [24, 25]. The PCM model considers the solvent as a continuous medium with a dielectric constant ε and represents the solute by means of a cavity built with a number of interlaced spheres [24, 25]. Recently, it has been demonstrated that polarizable continuum models provide a reasonable description concerning the localized/delocalized nature of organic mixed-valence systems [11, 26, 27]. In particular, simple polarizable continuum models are able to stabilize the localized MV structures, where the charge excess is asymmetrically distributed between the redox centers, in comparison with the delocalized structures, where the charge is symmetrically distributed in the redox centers. Atomic charges discussed in the text correspond to
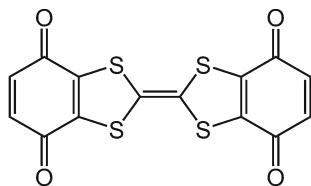

Fig. 1 Chemical structure of the Q–TTF–Q triad

those obtained within the Natural Population Analysis (NPA) algorithm [28]. Molecular orbitals were plotted using Chemcraft 1.6 [29].

Vertical electronic excitation energies were first computed for the localized anion species at the B3LYP/6-31+G** level using the time-dependent DFT (TDDFT) approach [30–32] and the optimized ground-state molecular geometries. However, the B3LYP functional usually underestimates the energy of the charge-transfer (CT) excited states. This shortcoming of standard global hybrid functionals has been reported for donor–acceptor compounds for which there is a negligible overlap between the molecular orbitals involved in the excitation [33, 34]. Long-range functionals, such as CAM-B3LYP and the ωB97 family, have been recommended to overcome this important deficiency [33, 35]. In this sense, vertical excitation energies using the hybrid, meta-hybrid and long-range-corrected functionals mentioned above were also evaluated. The complete active space method combined with a second-order perturbation approach (CASSCF/CASPT2) was employed to help in the interpretation of the TDDFT results. CASSCF/CASPT2 calculations using the (3, 3) active space (3 orbitals and 3 electrons) and the cc-pVDZ basis set were carried out using the MOLCAS 7.6 program package [36].

3 Results and discussion
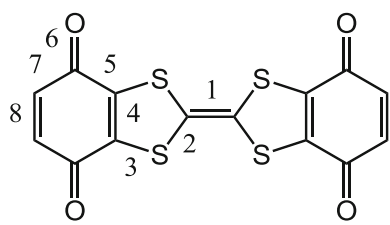
3.1 Structural and electronic properties of the neutral Q–TTF–Q triad

To investigate the effect of fusing the p-benzoquinone units on the molecular and electronic structure of the TTF backbone, geometry optimizations of the Q–TTF–Q triad and the constituting molecules (p-benzoquinone and TTF) were performed at the B3LYP/6-31G** level both in gas phase and in the presence of the solvent. Table 1 compares the optimized values calculated for selected bond lengths of Q–TTF–Q with those obtained for TTF and p-benzoquinone. The main differences are found for the bonds connecting the donor and the acceptor moieties. The carbon–sulfur (CS) bonds 2 and 3 of the TTF fragment undergo a notably shortening and lengthening, respectively, in comparison with isolated TTF. The carbon–carbon (CC) double bond shared by the TTF and Q units (bond 4) elongates from 1.337 and 1.343 (in TTF and Q, respectively) to 1.356 Å in the Q–TTF–Q triad. Fusion also affects to the Q units and the CC single bonds 5 shrink from 1.487 in Q to 1.480 Å in the triad. These structural changes are mainly ascribed to the electron-density transfer that takes place in the triad from the central TTF moiety, which accumulates a positive charge of +0.44e, to the

benzoquinone units, which bear a negative charge of −0.22e. Geometry optimization in the presence of solvent (DMSO) provides practically identical bond distances and net charges than those computed in gas phase. The theoretical trends calculated in passing from the isolated molecules (TTF and Q) to the Q–TTF–Q triad, especially those predicted for the CC double bonds 4, are in general agreement with those observed from experimental X-ray diffraction data (see Table 1) [37, 38].

The minimum-energy geometry computed for the Q–TTF–Q triad in gas phase exhibits the C_{2v} boat conformation shown in Fig. 2 with a folding angle θ along the S···S axes of 9.8°, which is similar to that previously reported both experimentally [39] and theoretically [40] for isolated TTF. The folding angle is indeed almost identical to that calculated for isolated TTF (9.6°). The incorporation of solvent effects by the PCM model confirms the boat conformation of Q–TTF–Q as an energy minimum with a slightly larger dihedral angle (14.0°). Fig. S1 in the Supporting Information displays the boat-to-planar energy profile computed for Q–TTF–Q by varying simultaneously the two folding angles θ. The energy profile suggests that there are at least two stabilization/destabilization effects to be considered: The steric hindrance produced by the sulfur atoms, which is minimized in the boat-like structure, and the π-delocalization along the whole molecule, which

Table 1 Selected B3LYP/6-31G**-optimized bond lengths (Å) calculated for Q–TTF–Q and for the TTF and p-benzoquinone (Q) molecules in gas phase

Bond	Q–TTF–Q[a]	TTF[b]	Q[c]
1	1.349 (1.340)	1.350 (1.337)	–
2	1.792 (1.769)	1.788 (1.755)	–
3	1.753 (1.739)	1.763 (1.731)	–
4	1.356 (1.349)	1.337 (1.309)	1.343 (1.322)
5	1.480 (1.471)	–	1.487 (1.477)
6	1.224 (1.227)	–	1.225 (1.222)
7	1.489 (1.462)	–	1.487 (1.477)
8	1.342 (1.342)	–	1.343 (1.322)

X-ray diffraction data are given in parentheses

[a] X-ray data from Ref. [6]

[b] X-ray data from Ref. [37]

[c] X-ray data from Ref. [38]

Fig. 2 Minimum-energy structure (C_{2v}, boat conformation) computed for Q–TTF–Q at the B3LYP/6-31G** level in gas phase. Sulfur and oxygen atoms are in *yellow* and *red*, respectively
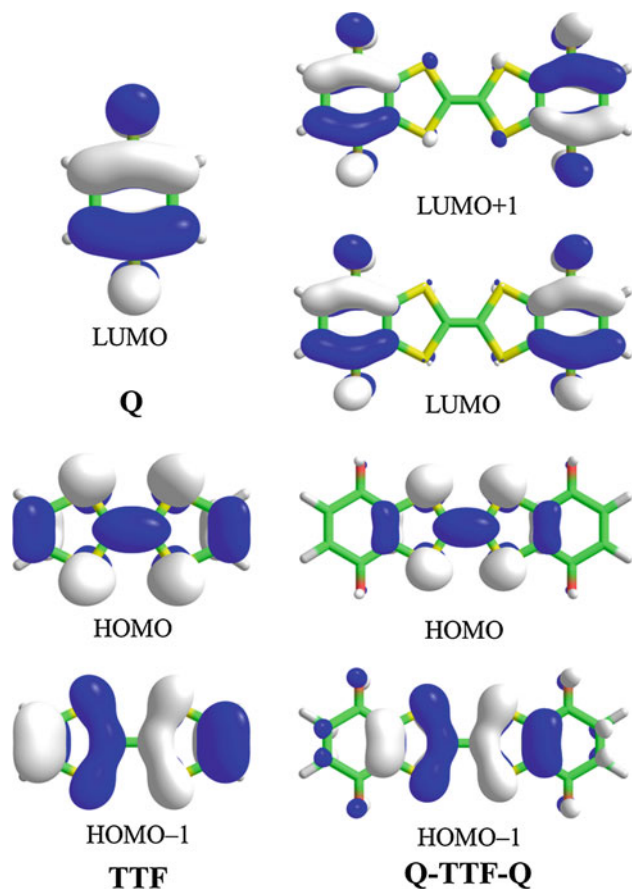

Fig. 3 Electron-density contours (0.03 e bohr^{-3}) calculated for the HOMOs and LUMOs of the Q–TTF–Q triad and the TTF and Q molecules at the B3LYP/6-31G** level

stabilizes the planar conformer. The planar D_{2h} structure is actually predicted as a minimum in the molecular potential energy surface and is found to be higher in energy than the boat conformation by only 0.01 kcal/mol. Both structures are separated by almost negligible barriers of 0.05–0.06 kcal/mol. These small barriers support the fully planar X-ray structure experimentally observed in the solid state where packing forces determine the planarization of the molecule [6].

Figure 3 displays the topology of the frontier MOs computed for the Q–TTF–Q triad and the TTF and Q fragments. In the triad, the highest occupied molecular orbital (HOMO) and the HOMO–1 spread over the electron-donor TTF unit and match exactly with the HOMO and HOMO–1 of the TTF molecule. The lowest

unoccupied molecular orbital (LUMO) and the LUMO+1 are completely localized on the electron-acceptor Q units and clearly resemble the LUMO of the *p*-benzoquinone molecule. They correspond to the in-phase and out-of-phase combination of the LUMO of the Q units, respectively (Fig. 3). An aspect to be kept in mind is that the delocalized description of the LUMO and LUMO+1 over the two identical Q fragments shown in Fig. 3 is perfectly compatible with a localized description where the LUMO and LUMO+1 would be degenerate and localized over different Q moieties. Although the localized/delocalized notation is not relevant for the neutral A–D–A triad, it will be used below for the electronic description of the anion since it gives rise to MV species. In oxidation/reduction terms, the distribution for the MOs clearly suggests that first oxidations will occur on the TTF unit, whereas first reductions will take place on the acceptor moieties.

3.2 Oxidized species

To analyse the effect of oxidation on the molecular structure of the Q–TTF–Q triad, geometry optimizations of the radical cation and dication species were performed at the B3LYP/6-31G** level in gas phase and in the presence of the solvent (DMSO). Table 2 summarizes the optimized bond lengths computed for Q–TTF–Q in its neutral, radical cation and dication states. The topology of the HOMO of the A–D–A triad shown in Fig. 3 helps in the interpretation of the structural changes that Q–TTF–Q undergoes upon oxidation. The oxidation process involves the extraction of electrons from the HOMO, thus weakening bonding and antibonding interactions localized on CC and CS bonds, respectively. As a consequence, CC double bonds lengthen and CS single bonds shorten and a more aromatic molecular structure with a lower bond length alternation is found upon oxidation. For example, the length of the central CC bond (bond 1) largely increases from 1.349 in the neutral state to 1.398 and 1.457 Å in the cation and dication states, respectively. On the contrary, the CS bonds 2 and 3 undergo a remarkable shortening, more pronounced for bond number 2 that shrinks by 0.081 Å. The CC bond 4 of Q–TTF–Q also experiences a lengthening on going from the neutral (1.354 Å) to the cation (1.356 Å) and dication (1.365 Å) species. The relative lengthening is significantly smaller than that computed for isolated TTF (from 1.337 to 1.346 and 1.360 Å, respectively) owing to the fusion of the

Table 2 Selected B3LYP/6-31G**-optimized bond lengths (Å) calculated in DMSO for the Q–TTF–Q triad in its neutral, radical cation and dication states

Bond[a]	Neutral	Cation	Dication
1	1.349 (1.349)	1.398 (1.399)	1.457 (1.457)
2	1.791 (1.792)	1.750 (1.753)	1.710 (1.720)
3	1.755 (1.753)	1.738 (1.741)	1.720 (1.725)
4	1.354 (1.356)	1.356 (1.355)	1.365 (1.364)
5	1.479 (1.480)	1.489 (1.494)	1.500 (1.510)
6	1.226 (1.224)	1.222 (1.218)	1.218 (1.213)
7	1.487 (1.489)	1.483 (1.484)	1.479 (1.480)
8	1.342 (1.342)	1.344 (1.346)	1.347 (1.351)

Bond distances in gas phase are given within parentheses

[a] See Table 1 for bond numbering

benzoquinone rings. Although the main structural changes take place in the TTF backbone, in agreement with the HOMO picture, the Q fragments are also affected by oxidation; bonds 5 conspicuously elongate from 1.479 (neutral) to 1.489 (cation) and 1.500 Å (dication), whereas bonds 6 and 7 slightly shrink. The structural changes in the Q fragments alleviate the stress caused by the more pronounced changes affecting the donor TTF unit.

The oxidation process has also a significant impact on the conformation of the A-D-A triad. A planarization of the boat-shaped TTF is found for the cation species (D_{2h} symmetry) as a consequence of the decrease in the electron density of the TTF moiety and especially of the S atoms. Figure 4 shows the NPA charge distribution calculated for Q–TTF–Q in neutral, radical cation and dication states. As expected, the acceptor Q moieties concentrate the electron density and present negative net charges in the neutral state that are compensated by the positive charge accumulated by the donor TTF fragment. Upon oxidation, electronic density is mainly extracted from this fragment, which augments its charge from +0.42 in the neutral state to +1.22 in the cation and to +1.98 in the dication. This implies the extraction of 80 % (neutral → cation) and 76 % (cation → dication) of the electron density from the TTF backbone. As mentioned above, oxidation causes an important lengthening of the central CC bond (bond 1) in TTF, which endows with certain flexibility to the central part of the molecule and allows the internal rotation of the TTF rings to reduce the steric hindrance between the sulfur atoms. In the minimum-energy structure computed for the dication state, the TTF rings are indeed twisted by an S–C–C–S dihedral angle of 30° in gas phase that increases to 40° in the presence of the solvent. These results are in concordance with those reported in the literature for TTF [39], and evidence that the nature of TTF is maintained in the triad and that electrons are mainly taken out from the TTF skeleton.
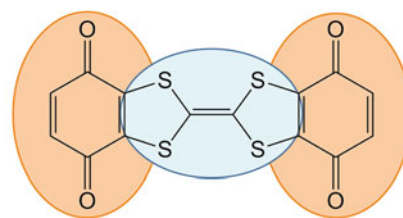
Fig. 4 Accumulated NPA net charges computed for neutral, radical cation and dication species of the Q–TTF–Q triad at the B3LYP/6-31G** level in DMSO. Charges related to the four carbon atoms connecting the TTF and Q units are divided equally between the fused fragments

3.3 Reduced species

The theoretical characterization of the Q–TTF–Q triad in its radical anion and dianion states is much more complex due to the possibility to form localized (Q⁻–TTF–Q) or delocalized [(Q–TTF–Q)⁻] anions (mixed-valence systems) and dianions with a biradicaloid character. To explore the electronic and structural properties of the reduced species, geometry optimizations of the radical anions and dianions were performed by using different hybrid (B3LYP, PBE0 and BHHLYP), meta-hybrid (M06-2X) and long-range-corrected (CAM-B3LYP, ωB97X and ωB97XD) density functionals in gas phase and in DMSO. The calculations were carried out with the 6-31+G** basis set because it incorporates diffuse functions, which are more appropriate for the treatment of anion species.

Table 3 collects the energy difference between the localized and delocalized structures computed for the Q–TTF–Q radical anion at different DFT levels both in gas phase and in DMSO. In gas phase, the functionals used within the unrestricted UDFT approach predict a delocalized C_{2v}-symmetry molecular structure where the charge excess is equally distributed between the Q moieties (Q–TTF–Q)⁻. Reoptimizations using the unrestricted broken-symmetry (BS) approximation were employed to find a possible localized structure lower in energy than the delocalized C_{2v}-symmetry solution. The UB3LYP(BS) and UPBE0(BS) methods converge to the same delocalized molecular and electronic structure as the unrestricted approach (Table 3). This theoretical description contrasts with the experimental characterization by electrochemistry and ESR spectroscopy that suggests a localized Q⁻–TTF–Q anion (class II mixed-valence system) where the extra electron is localized on one of the Q rings [5]. The poor description provided by the hybrid B3LYP and PBE0 functionals for the Q–TTF–Q anion in gas phase has been previously reported and was attributed to the

Table 3 Energy difference (kcal/mol) between the delocalized (C_{2v}) and localized (C_s) structures calculated for the Q–TTF–Q anion in gas phase and in DMSO

Functional	ΔE_{gas} ($C_{2v} - C_s$)	ΔE_{DMSO} ($C_{2v} - C_s$)
B3LYP	–[a]	1.71
PBE0	–[a]	3.16
BHHLYP	3.02	11.28
M06-2X	2.51	10.62
CAM-B3LYP	3.54	11.72
ωB97X	10.03	18.46
ωB97XD	6.48	14.50

[a] B3LYP and PBE0 are unable to predict stable Q⁻–TTF–Q localized structures in gas phase

self-interaction error which favors the charge-delocalized structures [41, 42]. In contrast, the hybrid BHHLYP and the meta-hybrid M06-2X functionals, with a large portion of HF exchange, as well as the long-range-corrected CAM-B3LYP, ωB97X and ωB97XD functionals are able to obtain localized C_s-symmetry structures more stable than the delocalized C_{2v}-symmetry ones. The localized and delocalized nature of these anion radicals is illustrated in Fig. 5 by the unpaired-electron spin density and NPA charge distributions calculated for the respective anions. Whereas localized Q⁻–TTF–Q structures (C_s-symmetry) are characterized by concentrating the extra electron in one of the Q moieties, the charge in the delocalized (Q–TTF–Q)⁻ structures (C_{2v}-symmetry) is equally distributed among the acceptor units.

Among the functionals assessed, the BHHLYP, M06-2X and CAM-B3LYP functionals predict an energy barrier for the interconversion of the localized anions (Q⁻–TTF–Q and Q–TTF–Q⁻) through the delocalized (Q–TTF–Q)⁻ structure of ∼3 kcal/mol (Table 3). When using the ωB97X and ωB97XD functionals, the barrier increases to 10.03 and 6.48 kcal/mol, respectively. These values are closer to the activation energy (7.96 kcal/mol) estimated for the intramolecular electron-transfer process from ESR experiments [6]. It should be clarified that the difference between the ωB97X and ωB97XD functionals is not only the dispersion term, which is expected to play no significant role in the estimation of the energy barrier, but also the parameter defining the short- and long-range separation (ω), which has a different value (0.3 and 0.2 for ωB97X and ωB97XD, respectively). This ω parameter is responsible for the different estimates provided by the ωB97X and ωB97XD functionals for the energy difference between localized and delocalized structures.

The incorporation of solvent effects by means of the PCM model in the optimization process has a strong impact on the relative stability of the localized and delocalized structures of the Q–TTF–Q anion (Table 3). In the
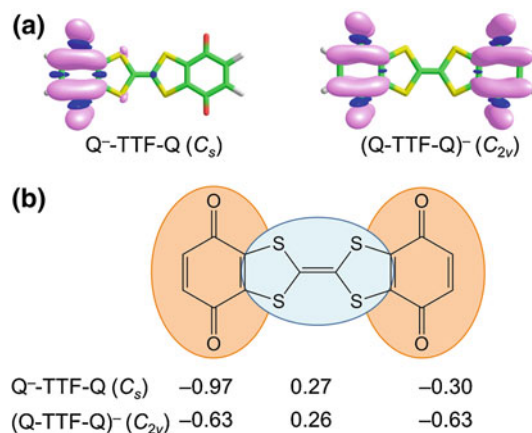
Fig. 5 a Unpaired-electron spin density distribution and **b** NPA charge distribution computed for the localized Q⁻–TTF–Q anion (C_s-symmetry) and the delocalized (Q–TTF–Q)⁻ anion (C_{2v}-symmetry) at the UBHHLYP(BS)/6-31+G** and UBHHLYP/6-31+G** levels, respectively. Charges related to the four carbon atoms connecting the TTF and Q units are divided equally between the fused fragments

presence of solvent, the localized Q⁻–TTF–Q solutions are found to be minimum-energy structures with all the functionals evaluated including B3LYP and PBE0. However, the interconversion energy barrier provided by the B3LYP and PBE0 functionals (1.71 and 3.16 kcal/mol, respectively) is significantly lower than that computed with the other functionals and underestimates the experimental value. The solvent clearly stabilizes the localized structures (Q⁻–TTF–Q and Q–TTF–Q⁻) and increases the energy barrier through the delocalized (Q–TTF–Q)⁻ structure by an almost constant amount of ∼8 kcal/mol when comparing with the values computed in gas phase (Table 3). The BHHLYP and M06-2X functionals and the long-range CAM-B3LYP functional provide energy barriers around 11 kcal/mol that overestimate the experimental activation energy (7.96 kcal/mol). Higher energy barriers are predicted with the ωB97X (18.46 kcal/mol) and ωB97XD (14.50 kcal/mol) functionals.

In contrast to the Q–TTF–Q anion, the closed-shell minimum-energy structure optimized for the dianion within the restricted DFT approach corresponds to a C_{2v}-symmetry (Q–TTF–Q)²⁻ structure. However, the restricted wave function shows a restricted → unrestricted instability. The biradical character of the (Q–TTF–Q)²⁻ dianion was investigated by reoptimizing its structure using the UDFT(BS)/6-31+G** approach. The triplet state was also optimized at the UDFT/6-31+G** level. Table 4 collects the energy differences computed between the closed-shell and open-shell (biradicaloid) singlet solutions and between the closed-shell singlet and triplet solutions. In general, the unrestricted open-shell solutions are found to be lower in energy than the closed-shell singlet and the biradicaloid singlet and triplet solutions are predicted to be virtually

Table 4 Energy differences (kcal/mol) between the closed-shell singlet (CS) and open-shell singlet (OS) states and between the closed-shell singlet and open-shell triplet (T) states of the Q–TTF–Q dianion calculated in gas phase (values in DMSO are given within parentheses)
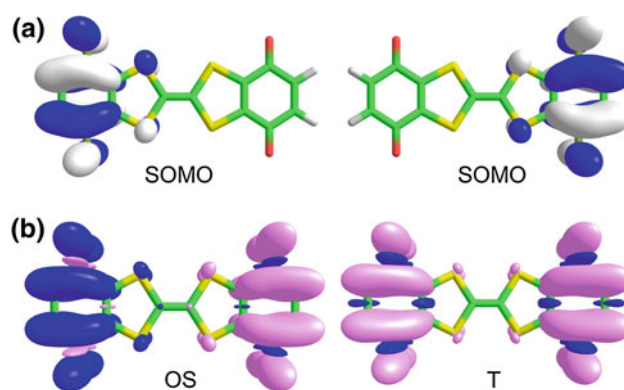
Functional	ΔE (CS–OS)	ΔE (CS–T)
B3LYP	17.52 (17.09)	17.50 (17.07)
PBE0	21.05 (20.51)	21.04 (20.49)
BHHLYP	37.66 (36.92)	37.65 (36.91)
M06-2X	34.76 (33.79)	34.75 (34.13)
CAM-B3LYP	37.78 (37.24)	37.77 (37.23)
ωB97X	50.90 (49.43)	50.88 (50.16)
ωB97XD	43.36 (41.65)	43.35 (42.33)

Fig. 6 a Electron-density contours (0.03 e bohr^{-3}) calculated at the UB3LYP(BS)/6-31+G** level for the singly occupied molecular orbitals of the Q–TTF–Q dianion and **b** unpaired-electron spin density distribution computed for the open-shell singlet (OS) and triplet (T) states

degenerate. The solvent plays no role in the relative stability of the restricted and unrestricted solutions. The dianion of the Q–TTF–Q triad is therefore predicted to be a biradicaloid species in its electronic ground state.

Investigation of the frontier molecular orbitals gives additional insight into the nature of the open-shell ground state predicted for (Q–TTF–Q)$^{2-}$. Figure 6a displays the two singly occupied molecular orbitals (SOMOs) calculated for (Q–TTF–Q)$^{2-}$ at the UDFT(BS)/6-31+G** level. They are essentially confined on the Q fragments with a small contribution on the S atoms of the TTF unit and almost no overlap between them. The (Q–TTF–Q)$^{2-}$ dianion can be therefore classified as a disjoint biradical [43–45]. The spin densities calculated for the open-shell singlet and triplet solutions (Fig. 6b) clearly show that the unpaired electrons reside on different parts of the molecule.
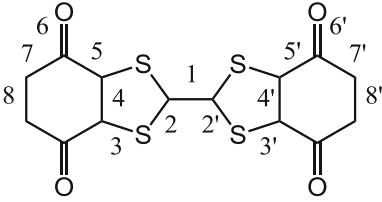
Table 5 gathers the B3LYP/6-31+G**-optimized bond distances computed for the neutral, radical anion (localized/delocalized) and dianion (open-shell singlet) species of the Q–TTF–Q triad in the presence of DMSO. As seen in Table 5, the most important structural changes affect the *p*-benzoquinone rings because, as mentioned above, the extra electrons are mainly injected in the Q units. The single and double CC bonds of the *p*-benzoquinone rings shrink and lengthen, respectively, and the Q rings become more aromatic upon reduction. This aromatization is accompanied by a remarkable lengthening of the double carbon–oxygen bonds (bonds 6 and 6′). The adjacent carbon–sulfur bonds (bonds 3 and 3′) of the fused TTF unit also undergo a sizeable lengthening. Table 5 clearly illustrates that in the localized Q$^-$–TTF–Q anion the structural changes are concentrated on one of the Q moieties in good agreement with the unpair-electron spin density and charge distributions shown in Fig. 5. Similar structural trends are obtained when using the other functionals employed in this work (Tables S1 and S2 in the Supporting Information).

The localized Q$^-$–TTF–Q anion exhibits boat-like and fully planar minimum-energy conformations as it was

found for the neutral Q–TTF–Q triad (Fig. S1). At the B3LYP/6-31+G** level, the energy difference between both conformations is also negligible (0.01 kcal/mol), but now favors the planar conformation. In the boat conformation, the folding θ angle shows different values for the Q$^-$ (12.4°) and Q^0 (16.5°) fragments that are similar to that calculated for the neutral molecule (13.2°). The difference between the folding angles is an additional structural signature which corroborates that the extra electron enters in only one Q moiety.

We finally investigate the photoinduced electron transfer between the negatively charged and neutral benzoquinone rings of the localized Q$^-$–TTF–Q anion. This photoinduced transfer was experimentally studied, but the difficulties found in solubilizing the Q$^-$–TTF–Q mixed-valence species prevented a quantitative study of the electron-transfer process by Vis/NIR spectroelectrochemical methods [5]. The deconvolution of the Vis/NIR spectra at different reduction steps revealed a broad absorption band centered around 1,300 nm (0.95 eV) which was associated to the Q$^-$–TTF–Q anion. To gain more insights into the nature of the absorption experimentally observed, the lowest-energy doublet excited states of Q$^-$–TTF–Q were calculated using the TDDFT approach and the previously optimized ground-state geometries. Three doublet excited electronic states (D$_n$) are found below 1.5 eV when calculating with the B3LYP, PBE0, M06-2X, CAM-B3LYP, ωB97X and ωB97XD functionals and the 6-31+G** basis set in DMSO (Table 6). At the B3LYP level, the D$_0 \rightarrow$ D$_1$ excitation can be described by a one-electron promotion from the SOMO to the LUMO, which are located on different Q moieties. The D$_0 \rightarrow$ D$_1$ excitation therefore corresponds to the Q$^- \rightarrow$ Q^0 electron transfer and is computed at 0.42 eV with a small oscillator strength ($f = 0.0086$). In the same energy range, the D$_0 \rightarrow$ D$_3$

Table 5 Selected B3LYP/6-31+G**-optimized bond distances (Å) computed in DMSO for the Q–TTF–Q triad in its neutral, radical anion and dianion states

Bond	Neutral	Localized anion (C_s)	Delocalized anion (C_{2v})	Dianion (OS)
1	1.351	1.352	1.351	1.349
2	1.790	1.784	1.787	1.789
2′	1.790	1.792	1.787	1.789
3	1.753	1.771	1.762	1.770
3′	1.753	1.754	1.762	1.770
4	1.359	1.375	1.368	1.376
4′	1.359	1.359	1.368	1.376
5	1.481	1.448	1.461	1.448
5′	1.481	1.478	1.461	1.448
6	1.226	1.272	1.250	1.264
6′	1.226	1.230	1.250	1.264
7	1.489	1.449	1.468	1.456
7′	1.489	1.487	1.468	1.456
8	1.343	1.373	1.357	1.370
8′	1.343	1.343	1.357	1.370

Table 6 Vertical excitation energies (ΔE) and oscillator strengths (f) computed in DMSO for the localized Q^-–TTF–Q structure at different TDDFT levels using the 6-31+G** basis set

Method	Excitation	Description	ΔE (eV)	f
B3LYP	$D_0 \to D_1$	$Q^- \to Q^0$	0.42	0.0086
	$D_0 \to D_2$	$TTF \to Q^0$	0.43	0.0000
	$D_0 \to D_3$	$TTF \to Q^0$	0.64	0.0026
PBE0	$D_0 \to D_1$	$TTF \to Q^0$	0.51	0.0000
	$D_0 \to D_2$	$Q^- \to Q^0$	0.68	0.0067
	$D_0 \to D_3$	$TTF \to Q^0$	0.79	0.0038
M06-2X	$D_0 \to D_1$	$TTF \to Q^0$	0.59	0.0000
	$D_0 \to D_2$	$Q^- \to Q^0$	0.72	0.0061
	$D_0 \to D_3$	$TTF \to Q^0$	0.86	0.0042
CAM-B3LYP	$D_0 \to D_1$	$TTF \to Q^0$	0.44	0.0000
	$D_0 \to D_2$	$Q^- \to Q^0$	1.05	0.0061
	$D_0 \to D_3$	$TTF \to Q^0$	1.06	0.0082
ωB97X	$D_0 \to D_1$	$TTF \to Q^0$	0.55	0.0000
	$D_0 \to D_2$	$Q^- \to Q^0$	0.72	0.0059
	$D_0 \to D_3$	$TTF \to Q^0$	0.82	0.0041
ωB97XD	$D_0 \to D_1$	$TTF \to Q^0$	0.57	0.0000
	$D_0 \to D_2$	$Q^- \to Q^0$	0.71	0.0061
	$D_0 \to D_3$	$TTF \to Q^0$	0.85	0.0041
CASPT2 (3,3)[a]	$D_0 \to D_1$	$Q^- \to Q^0$	0.53	0.0049
	$D_0 \to D_2$	$TTF \to Q^0$	0.79	0.0073
	$D_0 \to D_3$	$TTF \to Q^0$	0.87	0.0024

[a] Single-point CASPT2 calculation performed in gas phase on the B3LYP/6-31+G** optimized geometry using the cc-pVDZ basis set

excitation is calculated at 0.64 eV ($f = 0.0026$) and involves an electron transfer from the TTF moiety to the neutral Q^0 fragment. The $D_0 \to D_2$ excitation has no intensity ($f = 0.0000$). The B3LYP functional therefore suggests that two electronic transitions ($Q^- \to Q^0$ and $TTF \to Q^0$) contribute to the broad band experimentally observed around 0.95 eV [5], but underestimates the energy of this band. It is well known that B3LYP tends to underestimate significantly the energy of the CT excitations [46]. The PBE0 functional provides a similar description with the $Q^- \to Q^0$ ($D_0 \to D_2$) and $TTF \to Q^0$ ($D_0 \to D_3$) transitions now computed at 0.68 ($f = 0.0067$) and 0.79 eV ($f = 0.0038$), respectively. The long-range-corrected CAM-B3LYP functional predicts excitation energies of 1.05 ($f = 0.0061$) and 1.06 eV ($f = 0.0082$), respectively, which are closer to the experimental absorption energy. The meta-hybrid (M06-2X) and the long-range-corrected (ωB97X and ωB97XD) functionals provide a similar description of the lowest-energy excited states, with excitation energies for the weak $Q^- \to Q^0$ ($D_0 \to D_2$) and $TTF \to Q^0$ ($D_0 \to D_3$) transitions comprised between the values predicted by PBE0 and CAM-B3LYP. Single-point gas phase CASSCF/CASPT2 calculations performed on the B3LYP/6-31+G** optimized geometry using the (3,3) active space and the

cc-pVDZ basis set support the presence of both the $Q^- \to Q^0$ and $TTF \to Q^0$ transitions at similar energies: 0.53 ($f = 0.0049$) and 0.79 eV ($f = 0.0073$). Although the (3,3) active space is very reduced, it includes the molecular orbitals and the electrons that play a major role in the $Q^- \to Q^0$ and $TTF \to Q^0$ transitions, that is, the HOMO localized on the TTF unit and the SOMO and LUMO localized on the Q^- and Q^0 units, respectively.

4 Conclusions

The acceptor–donor–acceptor Q–TTF–Q triad has been theoretically characterized by using DFT quantum-chemical calculations. Emphasis has been made on the evolution of the structural and electronic properties upon oxidation/reduction. The minimum-energy structure of the Q–TTF–Q triad in its neutral state is a C_{2v} boat conformation, although the planar D_{2h} structure is also predicted as a minimum in the molecular potential energy surface. Both structures are separated by almost negligible barriers in accord with the fully planar structure experimentally observed in the solid state. The donor TTF and acceptor Q

units preserve their electronic structure in the neutral system, for which a charge transfer takes place from the central TTF fragment to the external Q units. Upon oxidation, electrons are mainly extracted from the TTF moiety. The cation species has a fully planar D_{2h} structure, whereas the dication shows a twisted conformation as a consequence of the elongation of the central CC bond. The structural and electron-donor properties of the TTF unit are therefore maintained in the triad.

The radical anion and dianion states of the Q–TTF–Q triad were investigated by using different hybrid (B3LYP, PBE0 and BHHLYP), meta-hybrid (M06-2X) and long-range-corrected (CAM-B3LYP, ωB97X and ωB97XD) density functionals. In gas phase, BHHLYP, M06-2X and the long-range CAM-B3LYP, ωB97X and ωB97XD functionals predict that the Q–TTF–Q triad behaves as a class II mixed-valence system and form localized Q^-–TTF–Q anions where the unpaired electron is mainly located over one acceptor Q moiety. In contrast, the hybrid B3LYP and PBE0 functionals are unable to predict stable Q^-–TTF–Q structures and lead to a delocalized $(Q$–TTF–$Q)^-$ structure, where the extra electron is equally distributed among the two acceptor moieties. The incorporation of solvent effects stabilizes the localized Q^-–TTF–Q solutions, and these solutions are found to be minimum-energy structures with all the functionals evaluated. The dianion of the Q–TTF–Q triad is predicted to be a biradicaloid species, for which the unpaired electrons mainly reside on the external Q units. TDDFT and CASSCF(3,3)/ CASPT2 calculations suggest that the broad absorption band experimentally observed for the localized Q^-–TTF–Q anion is associated with two charge-transfer electronic transitions: $Q^- \rightarrow Q^0$ and $TTF \rightarrow Q^0$.

References

1. Ferraris J, Walatka V, Perlstei JH, Cowan DO (1973) J Am Chem Soc 95:948–949
2. Williams JM, Ferraro JR, Thorn RJ, Carlson KD, Geiser U, Wang HH, Kini AM, Whangbo MH (1992) Organic superconductors (including fullerenes). synthesis, structure, properties, and theory. Prentice Hall, Englewoods Cliffs
3. Organic Conductors: Fundamentals and Applications (1994). Marcel Dekker, New York
4. Bendikov M, Wudl F, Perepichka DF (2004) Chem Rev 104:4891–4945
5. Gautier N, Dumur F, Lloveras V, Vidal-Gancedo J, Veciana J, Rovira C, Hudhomme P (2003) Angew Chem Int Ed 42:2765–2768
6. Dumur F, Gautier N, Gallego-Planas N, Sahin Y, Levillain E, Mercier N, Hudhomme P, Masino M, Girlando A, Lloveras V, Vidal-Gancedo J, Veciana J, Rovira C (2004) J Org Chem 69:2164–2177
7. Robin MB, Day P (1968) Adv Inorg Chem Radiochem 10:247–422
8. Heckmann A, Lambert C (2012) Angew Chem Int Ed 51:326–392
9. Hankache J, Wenger OS (2011) Chem Rev 111:5138–5178
10. Oton F, Lloveras V, Mas-Torrent M, Vidal-Gancedo J, Veciana J, Rovira C (2011) Angew Chem Int Ed 50:10902–10906
11. Kaupp M, Renz M, Parthey M, Stolte M, Wurthner F, Lambert C (2011) Phys Chem Chem Phys 13:16973–16986
12. Cohen AJ, Mori-Sanchez P, Yang WT (2008) Science 321:792–794
13. Mori-Sánchez P, Cohen AJ, Yang W (2008) Phys Rev Lett 100:146401
14. 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, 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, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09, Revision A.02 Gaussian, Inc., Wallingford CT
15. Becke AD (1993) J Chem Phys 98:5648–5652
16. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785–789
17. Francl MM, Pietro WJ, Hehre WJ, Binkley JS, Gordon MS, Defrees DJ, Pople JA (1982) J Chem Phys 77:3654–3665
18. Perdew JP, Burke K, Ernzerhof M (1996) Phys Rev Lett 77:3865–3868
19. Adamo C, Barone V (1999) J Chem Phys 110:6158–6170
20. Zhao Y, Truhlar D (2008) Theor Chem Acc 120:215–241
21. Yanai T, Tew DP, Handy NC (2004) Chem Phys Lett 393:51
22. Chai J-D, Head-Gordon M (2008) J Chem Phys 128:084106
23. Chai J-D, Head-Gordon M (2008) Phys Chem Chem Phys 10:6615–6620
24. Tomasi J, Mennucci B, Cammi R (2005) Chem Rev 105:2999–3094
25. Tomasi J, Persico M (1994) Chem Rev 94:2027–2094
26. Renz M, Theilacker K, Lambert C, Kaupp M (2009) J Am Chem Soc 131:16292–16302
27. Renz M, Kess M, Diedenhofen M, Klamt A, Kaupp M (2012) J Chem Theory Comput 8:4189–4203
28. Reed AE, Curtiss LA, Weinhold F (1988) Chem Rev 88:899–926
29. http://www.chemcraftprog.com
30. Jamorski C, Casida ME, Salahub DR (1996) J Chem Phys 104:5134–5147
31. Petersilka M, Gossmann UJ, Gross EKU (1996) Phys Rev Lett 76:1212–1215
32. Casida ME, Jamorski C, Casida KC, Salahub DR (1998) J Chem Phys 108:4439–4449
33. Peach MJG, Benfield P, Helgaker T, Tozer DJ (2008) J Chem Phys 128:044118
34. Wiggins P, Williams JAG, Tozer DJ (2009) J Chem Phys 131:091101
35. Jacquemin D, Perpète E, Ciofini I, Adamo C (2011) Theor Chem Acc 128:127–136
36. Aquilante F, De Vico L, Ferré N, Ghigo G, Malmqvist P-å, Neogrády P, Pedersen TB, Pitoňák M, Reiher M, Roos BO, Serrano-Andrés L, Urban M, Urban M, Veryazov V (2010) J Comput Chem 31:224–247

37. Ellern A, Bernstein J, Becker JY, Zamir S, Shahal L, Cohen S (1994) Chem Mater 6:1378–1385
38. Trotter J (1960) Acta Cryst 13:86–95
39. Hargittai I, Brunvoll J, Kolonits M, Khodorkovsky V (1994) J Mol Struct 317:273–277
40. Viruela R, Viruela PM, Pou-Amérigo R, Ortí E (1999) Synth Met 103:1991–1992
41. Wu Q, Voorhis TV (2006) J Chem Phys 125:164105
42. Vydrov OA, Scuseria GE (2006) J Chem Phys 125:234109
43. Borden WT, Davidson ER (1977) J Am Chem Soc 99:4587–4594
44. Borden WT, Iwamura H, Berson JA (1994) Acc Chem Res 27:109–116
45. Filatov M, Shaik S (1999) J Phys Chem A 103:8885–8889
46. Autschbach J (2009) Chem Phys Chem 10:1–5

Structures and energetics of organosilanes in the gaseous phase: a computational study

**Ryusuke Futamura · Miguel Jorge ·
José R. B. Gomes**

Abstract The gas-phase conformations and stabilities of neutral and anionic organosilanes with structure ((HO)$_3$Si-organic linker-Si(OH)$_3$), where the organic linker is benzene, ethene, or ethane, were studied using density functional theory. The calculations were performed at the B3LYP/6-311+G(2d,2p) level of theory and show that the cis-bis(trihydroxysilyl)-ethene and gauche-bis(trihydroxysilyl)-ethane species are more stable than their trans and anti-counterparts, respectively. The local geometries of the organic and inorganic fragments in these hybrid compounds are similar to those found in the case of pure silicate compounds or in the parent organic molecules. The calculated enthalpies of deprotonation for these species suggest an acid–base behavior for 1,4-bis(trihydroxysilyl)-benzene species that is intermediate of those calculated for the silicate monomer and for the silicate dimer, while for the cis-bis(trihydroxysilyl)-ethene and gauche-bis(trihydroxysilyl)-ethane, an acid–base behavior that is intermediate of those calculated for small and for large pure silicates. It was also found that the calculated charges of the Si atoms are almost independent of the type of carbon atom to which they are bonded and that the charge localized on the organic moiety is always negative, even for the neutral species. This information is valuable for the development of molecular force fields for simulating systems involving organosilicates.

Keywords Density functional theory · Periodic mesoporous organosilicas · Organosilicates · Atomic point charges · Enthalpies of deprotonation

R. Futamura · J. R. B. Gomes (✉)
CICECO, Department of Chemistry, University of Aveiro,
Campus Universitário de Santiago, 3810-193 Aveiro, Portugal
e-mail: jrgomes@ua.pt

Present Address:
R. Futamura
Research Center for Exotic Nanocarbons (JST), Shinshu
University, 4-17-1 Wakasato, Nagano 380-8553, Japan

M. Jorge
LSRE, Laboratory of Separation and Reaction Engineering,
Associate Laboratory LSRE/LCM, Faculdade de Engenharia,
Universidade do Porto, Rua Dr. Roberto Frias,
4200-465 Porto, Portugal

1 Introduction

Nanoporous materials have been attracting the attention of researchers for several decades due to their intrinsic characteristics, such as very high surface areas, which make these materials ideal for many applications in nanotechnology and nanoscience [1–3]. Porous materials can adsorb and store large volumes of gases [4], which can be effectively removed under adequate conditions, for example, by changing the temperature. Therefore, these materials are ideal candidates for capturing greenhouse gases [5], or for storage of alternative energies to fossil fuels [6]. Porous materials are also finding application in various industrial processes, acting as water purifiers [7] or catalysts [8], and are being used in the construction of electrodes possessing optimal pore sizes for sophisticated high-performing capacitors [9] or in basic research of fluid behavior confined in nanospaces [10–12].

Over several decades, activated carbons and periodic mesoporous silica (PMS) have been employed as ideal material structures with nanospaces. The pores in these materials have defined slit or cylindrical shapes but show some microscopic disordering. Other more interesting mesoporous hybrid materials have been introduced recently. In fact, three groups independently reported a new type of mesoporous material, which was synthesized using micelle template methods as in the synthesis of PMS but using organosilicates (organic linker was ethane or ethene) instead of silicates as silica sources [13–15]. These new materials were named periodic mesoporous organosilicas (PMO). Just a few years later, Inagaki et al. [16] reported the synthesis of an ordered mesoporous organosilica hybrid material, where the organic linker was benzene, having a hexagonal array of mesopores with a lattice constant of 52.5 Å and crystal-like pore walls that exhibit structural periodicity with a spacing of 7.6 Å along the channel direction. These PMOs have two remarkable features which ordinary PMSs do not possess, that is, molecular order periodicity within the pore walls and hybrid organic–inorganic structure, alternating hydrophilic and hydrophobic layers, where each individual organic group is covalently bonded to two or more silicon atoms [17]. Although the detailed microscopic structure and the formation process of PMOs are not yet fully understood, the molecular ordering of the PMO pore walls was confirmed by large-angle XRD measurements [16]. It was also anticipated that the inorganic parts of PMOs supply strong framework structures and the organic parts are responsible for unique features of the materials, which as the field of PMO progresses, are increasingly responsible for new applications, for example, in lighting, solar energy, displays, and electrically conductive PMO matrices [17]. Thus, PMOs are still promising materials for modern society with an enormous potential mainly due to the versatility of their chemical compositions. The novel functional PMOs often contain large and complex precursors, and their synthesis and self-assembly may pose a significant challenge despite the knowledge acquired so far. Therefore, the understanding of the PMO formation process and structural details is crucial for creating new types of materials.

Some studies have been carried out recently aiming at elucidating the structure and properties of PMO materials. Commotti et al. [18] reported a 2D NMR study of PMOs synthesized from a benzenesilicate (1,4-bis(triethoxysilyl)benzene) and quantified for the first time the amounts of the different Si-condensed species, which were found to be 64.9 %, 32.3 %, and 2.8 % for the T^3 (fully condensed Si), T^2 (Si unit containing one OH group), and T^1 (Si unit containing two OH groups) species, respectively. Attributing negligible T^1 species to defects in the structure, and based on the $\sim 2:1$ ratio between T^3 and T^2 species, they suggested that the inorganic layer in the channel walls is constituted by alternate six- and four-member rings of a repeating unit of $(T^3\text{-}T^2\text{-}T^3)_n$ along each wall face rather than a regular sequence of six-member silica rings with alternate T^3 and T^2 species, that is, 1:1 ratio between T^3 and T^2 species of a repeating unit of $(T^3\text{-}T^2)_n$ along each wall face [18]. Very recently, Martinez et al. [19] reported the results of a periodic density functional theory (DFT) study on the interactions of CO, CO_2, and CH_4 molecules with the internal walls of the two channel wall models described above, that is, $(T^3\text{-}T^2\text{-}T^3)_n$ and $(T^3\text{-}T^2)_n$ repeating units. These gaseous molecules interact preferentially with the silanol groups, with calculated adsorption energies of 14.3 kJ/mol, 13.3 kJ/mol, and 9.2 kJ/mol for CO, CO_2, and CH_4, respectively, where dispersion was found to account for about one half of the interaction energy. Martinez and Pacchioni introduced some modifications in their models (e.g., inorganic part modified with Ti–OH, Si–O$^-$, and Al$^-$–OH groups or organic part modified with pyridine or biphenylene groups) and found that the interaction energies were identical or lower than in the case of the original phenylene (organic linker is benzene) species. Soichi et al. studied the Si–C bond stability under acidic and basic conditions of organosilane molecules during polycondensation to organosilica hybrids with NMR and density functional theory (DFT) at the B3LYP//6-311++G(d,p) level of theory [20], in order to explain why these bonds in some organosilane precursors are unexpectedly cleaved under such harsh conditions. For example, Si–C bond breakage was found to occur in the case of thiophene-bridged organosilane in basic conditions while almost no Si–C bond rupture was found in the cases where the organic linkers were derived from ethane, ethylene, or benzene [20].

Previously, we have reported a comprehensive computational work devoted to the understanding of the early stages of PMS synthesis. To reach our purposes, detailed atomistic molecular dynamics (MD) simulations were performed for several different aqueous solutions containing a surfactant molecule (decyltrimethylammonium), neutral and charged silicate monomers and oligomers, and bromide counterions, which were used to model realistic synthesis solutions with concentrations of different silicate species obeying experimental compositions [21–23]. The potential used for the description of the different silicate species was based on DFT calculations [24]. We have found that anionic silicates interact more strongly with the ammonium heads of the cationic surfactants than the bromide ions, and this was found to be related with an increase of the micelle sizes and with the enhancement of silica condensation at the micelle surface (cooperative templating mechanism). The simulations led to the proposal of a mechanism that validates several previous experimental and modeling studies and answers many controversial issues regarding the synthesis of mesoporous silicas. It would be interesting to adapt

the DFT + MD technique used for PMS to hybrid PMO materials, aiming at elucidating the initial stages of the formation mechanism of the latter. To accomplish such task, it is necessary to extend the potential that was employed before in the study of the early stages of the synthesis of PMS [24]. Hence, DFT calculations were performed to optimize the structures of the organosilicate monomers (neutral and single negative charged silanes with ethane, ethane, or benzene linkers), considered as models of the precursor species employed in the synthesis of PMO materials. These calculations are also important in the context of modeling chemical reactions involving organosilicate species.

The full set of computational details is given in Sect. 2, while in Sect. 3 the DFT results are analyzed and discussed, that is, the calculated partial atomic charges and the optimized bond lengths and angles used to develop the new potential for atomistic MD simulations of the PMO synthesis. Finally, the most important conclusions are summarized in Sect. 4.

2 Computational details

The structures of the organosilane species ((HO)$_3$Si-organic linker-Si(OH)$_3$), namely 1,4-bis(trihydroxysilyl)-benzene, bis(trihydroxysilyl)-ethene, and bis(trihydroxysilyl)-ethane, were optimized in the gas phase by means of density functional theory. The DFT calculations were performed by means of the hybrid B3LYP method, which is based on the ideas of Becke [25] and considers a mixture of Hartree–Fock and DFT exchange using the B3 functional, together with the LYP correlation functional due to Lee et al. [26] as implemented in the Gaussian 03 computational code [27]. The atomic electronic density was described by means of the 6-311+G(2d,2p) basis set.

This exchange–correlation functional and basis set were used for computational consistency with the structures optimized previously for silicate monomers as well as for linear and branched oligomers [24]. In fact, in that work, we have considered the B3LYP functional and some other computational approaches for calculating the enthalpy of the reactions H$_2$SiO → HSiO$^-$ + H$^+$ and [SiOH]$^+$ → SiO + H$^+$ but none of the others emerged as a clear and more appropriate choice than B3LYP (c.f. electronic annex A1 in ref. 24). The DFT calculations were also performed on the following negatively charged species: [(HO)$_3$Si-organic linker-Si(OH)$_2$O]$^-$ and [O(HO)$_2$Si-organic linker-Si(OH)$_2$O]$^{2-}$. The refinement of the potential needs the consideration of charged silane species since high pH conditions, pH = ∼11, are used in the experimental synthesis of the PMO materials. Under these alkaline conditions, most (organo) silicates are in their deprotonated anionic form [28].

Due to a large number of conformational degrees of freedom (e.g., cis and trans conformations are possible for the ethene organic linker), the structural optimization of the global minimum for each of the neutral and negatively charged configurations followed the computational strategy below: First of all, at least 10 different initial configurations for each organosilane were designed using information from our previous DFT work on neutral and charged linear, branched, and cyclic silicates [24] and optimized using very tight criteria (int = ultrafine and opt = very-tight keywords in the Gaussian code) by B3LYP but employing the smaller 6-31G(d) basis set. Please notice that several initial geometries lead to the same final state after optimization with the 6-31G(d) basis set. Then, the most stable configurations (∼5) optimized with the latter approach were further re-optimized at the B3LYP/6-311+G(2d,2p) level of theory. All optimized configurations were confirmed to be potential energy minima with vibrational frequency calculation after the re-optimization. Finally, the most stable configurations were used for subsequent atomic point charge calculations, which considered the Mulliken [29], CHelpG (charges from electrostatic potential grid method) [30], and NPA [31] (natural population analysis) approaches.

In order to facilitate the reading, we have considered the following notation for the silica fragments: The neutral monomer Si(OH)$_4$ is denoted by SN; the ionic monomer [Si(OH)$_3$O]$^-$ is denoted by SI (not to be confused with the silicon atom, Si), and the different 1,4-bis(trihydroxysilyl)-benzene molecules are denoted as SNBSN, SIBSN, and SIBSI, respectively, the neutral, the single, and the double negative charged species. Similar notation was used for the organosilanes based on ethane and ethylene, that is, SNE3SN, SIE3SN, and SIE3SI and SNE2SN, SIE2SN, and SIE2SI, respectively. With regard to atom type, these will be denoted as follows: Si is a silicon atom, Oh is a hydroxyl oxygen atom, Oc is a deprotonated oxygen atom (negative charge), Ho is a hydroxyl hydrogen atom, H is a hydrogen atom in an organic group, and C is a carbon atom. For the latter two atom types, subscript A refers to a phenyl group, subscript sp2 refers to an ethylene group, and subscript sp3 refers to an ethane group. Furthermore, subscripts N and I are used to denote whether atom types belong to a neutral or to an anionic fragment, respectively.

3 Results and discussion

3.1 Optimized organosilane structures

Views of the optimized structures for the neutral and negatively charged organosilicates are shown in Figs. 1, 2, and 3 for organic linkers benzene, ethene, and ethane,

respectively. Full geometrical details for these molecules are given as supplementary information (Tables S1-S3).

Let us begin with the description of the optimized geometry for 1,4-bis(trihydroxysilyl)-benzene and the anions obtained by single deprotonation at just one or at both silyl groups. In the following discussion, the organosilane molecular structure will be divided into three fragments, that is, SN or SI units and organic linker. The internal O–Si–O angles in SNBSN are 114.5°, 107.5°, and 103.9° (Fig. 1a and Table S1 in the Supporting Information). These geometrical parameters optimized in the gaseous phase are in quite good agreement with the experimental results from X-ray powder structural measurements by Cerveau et al. [32], respectively, 116.4(4)°, 112.9(4)°, and 103.4(4)°. As it can be seen, the SN fragments do not possess local C3 symmetry either in the gas or crystal phases. The calculated C–C–C angles of the aromatic ring differ by up to 2.2° from the values in benzene (120°). The Si–O bond lengths are in the range 1.643–1.652 Å, the Si–C bond lengths are 1.861 Å, the HC–CH bond lengths are 1.390 Å, and the SiC–CH bond distances are 1.400 Å. All these results are in quite good agreement with the X-ray powder diffraction results [33].

However, it is found that the Si–C bonds are shorter in the gas than in the crystal phase while the Si–O bond lengths are included in the range defined by the X-ray values.

The optimized structure for the SIBSN anion is shown in Fig. 1b. As expected, the most notable difference between the structural parameters optimized for the neutral SNBSN species and this anion is observed for the SI fragment. The Si–Oc distance in the SI fragment is shortened by 0.08–0.09 Å and the Si–Oh distances in the SI fragment are elongated by ∼0.05 Å with respect to Si–Oh bonds in SNBSN molecule. The lengthening of Si–Oh bonds occurs as a response to the shorter Si–Oc bond, due to stereo-electronic effects caused by orbital overlap since the electronic cloud of the Oc atom is larger, and closer to the Si atom, than that of Oh. A similar lengthening (mean value is 0.07 Å [23]) was found for the pure silicates (monomers and oligomers) optimized with the same computational approach. The Si–Oh bonds in the SN fragment of the SIBSN species are in the range 1.651–1.663 Å (Table S1), that is, they are practically unchanged when compared with the SN fragments in SNBSN molecule. The concentration of negative charge in the SI fragment of the SIBSN species induces the elongation of the Si–C bond

Fig. 1 Side and top views of the optimized configurations for a SNBSN, b SIBSN, and c SIBSI molecules. Bond lengths are in Å, and bond angles are in degrees
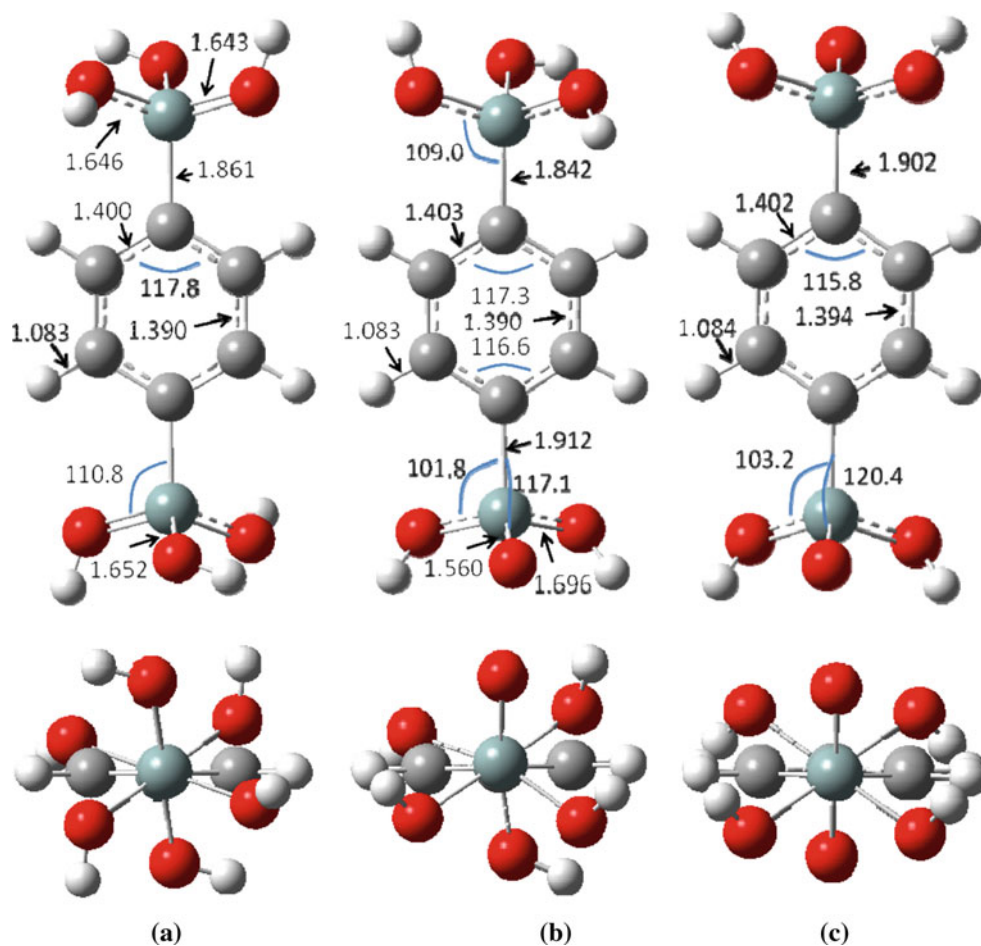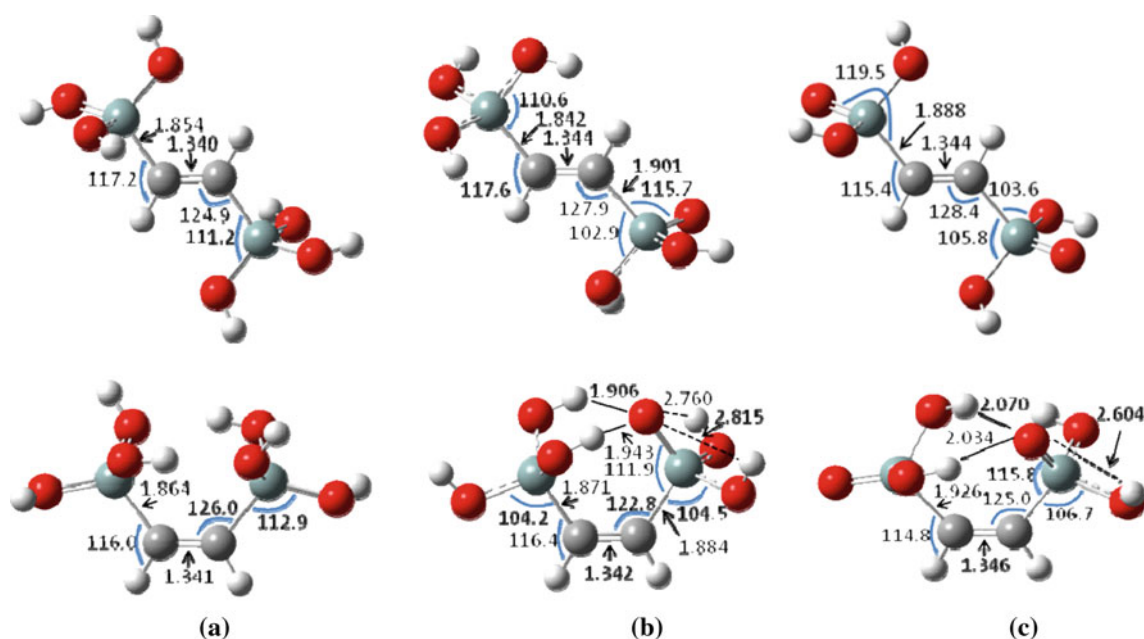
Fig. 2 Optimized trans (*top*) and cis (*bottom*) configurations for **a** SNE2SN, **b** SIE2SN, and **c** SIE2SI molecules. Bond lengths are in Å, and bond angles are in degrees
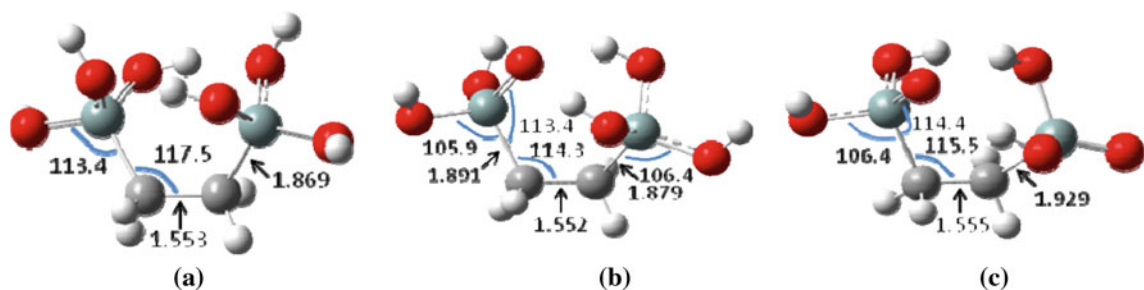
Fig. 3 Optimized structures for **a** SNE3SN, **b** SIE3SN, and **c** SIE3SI molecules. Bond lengths are in Å, and bond angles are in degrees

when compared with the Si–C bonds in SNBSN, and an opposite but small variation is found for the Si–C bond in the SN fragment of SIBSN when compared with the neutral molecule. These variations are a consequence of charge transfer as discussed below.

In the case of the double-charged ion, that is, SIBSI, Fig. 1c, the changes in the SI fragment bond lengths and angles with respect to the SN fragment in the neutral molecule are similar to those described above for the SI fragment of the SIBSN anion. Hence, the Si–Oc distances are shortened while the Si–Oh and Si–C bonds are elongated with respect to Si–Oh and Si–C bonds in SNBSN. Contrasting to what was found for the silica oligomers [24], the rigidity of the aromatic ring in these organosilanes prevents the formation of hydrogen bond contacts between the H(Oh) and Oc atoms in different SI fragments. Thus, the molecular structures of benzenesilicates are unfolded. As it can be seen in the bottom row of Fig. 1, the benzenesilicates show eclipsed configuration of the SN or SI

moieties along the Si–C-(...)-C–Si direction as in the crystallographic results by Cerveau et al. [33] which is expected to facilitate intermolecular hydrogen bonding in condensed phases. This is very relevant for understanding the superior ordering and the periodic arrangement of the structures of PMO materials.

As said above, the synthesis of the type of mesoporous materials considered in this work is usually carried out at high pH, and therefore, the solutions used in the experimental work are found to present relative proportions of anionic species higher than of neutral species. It is interesting to compare the deprotonation enthalpies of the organosilicates with those calculated previously for linear and branched silicates [24]. The enthalpies of deprotonation are listed in Table 1 and were calculated using the energetic data compiled in Table S4 of the Supporting Information. When analyzing Table 1, it is important to notice that these deprotonation enthalpies were calculated in the gaseous phase. To obtain values for the enthalpies in

Table 1 Calculated enthalpies of deprotonation at $T = 298.15$ K for the organosilicates considered in this work

Reaction	ΔH_{dep} (kJ/mol)	$\Delta H_{dep}/n$ (kJ/mol)[a]
SNBSN → SIBSN + H$^+$	1448.9	1448.9
SNBSN → SIBSI + 2H$^+$	3131.0	1565.5
trans-SNE2SN → trans-SIE2SN + H$^+$	1449.3	1449.3
trans-SNE2SN → trans-SIE2SI + 2H$^+$	3186.4	1593.2
cis-SNE2SN → cis-SIE2SN + H$^+$	1387.7	1387.7
cis-SNE2SN → cis-SIE2SI + 2H$^+$	3191.8	1595.9
anti-SNE3SN → anti-SIE3SN + H$^+$	1459.5	1459.5
anti-SNE3SN → anti-SIE3SI + 2H$^+$	3194.4	1597.2
gauche-SNE3SN → gauche-SIE3SN + H$^+$	1398.4	1398.4
gauche-SNE3SN → gauche-SIE3SI + 2H$^+$	3203.3	1601.7

[a] Enthalpy of deprotonation divided by the number of H$^+$ species involved in the reaction

solution, solvent effects would need to be considered using a realistic approach, which is beyond the scope of this work. Nevertheless, it is instructive to compare qualitative trends between different organosilicate species.

It is found that the enthalpy required to remove a proton from SNBSN to yield SIBSN is 25 kJ/mol lower than that required to remove a proton from a silicate monomer, SN, via the Si(OH)$_4$ → [SiO(OH)$_3$]$^-$ + H$^+$ reaction, but is 46 kJ/mol higher than the deprotonation enthalpy calculated for the SN2 dimer and $> \sim 100$ kJ/mol than the deprotonation enthalpies for larger linear and branched silicates [24]. Thus, it seems that the SNBSN compound has an acid–base behavior intermediate of those found for SN and SN2 species. The mean value calculated for single deprotonation at both silicates is intermediate of those computed before for SN2 yielding SISI and for SN3 yielding SISNSI species.

The optimized geometries for the trans and cis configurations of bis(trihydroxysilyl)ethene are reported in Fig. 2a. The most striking difference between trans and cis conformers is that in the latter, internal hydrogen bonding between H and Oh atoms is possible. This makes the cis isomer more stable than the trans species ($\Delta H = 9.9$ kJ/mol), where such H-bond interactions are not possible. This energy difference becomes larger in the case of the anions (Fig. 2b, c) since stronger hydrogen bonds are formed between Oc and H atoms in different silica fragments, and hence, these species are much more folded than the parent neutral SNE2SN species. The trans to cis enthalpy gain is 71.5 kJ/mol and 4.6 kJ/mol for SIE2SN and SIE2SI species, respectively. The very large difference between the stability of cis and trans SIE2SN species is due to the formation of two very strong internal hydrogen bonds in the former and concomitant cooperative effect introduced

by two other but weaker hydrogen bonds (dashed lines in Fig. 2b).

As it is shown in Table 1, the first enthalpy of deprotonation calculated for the trans-SNE2SN species is similar to that calculated for the SNBSN compound (difference of only 0.4 kJ/mol) but that calculated for the cis-SNE2SN compound is lower by ~ 60 kJ/mol. Once again, this difference is due to the presence of a strong internal hydrogen bond in cis-SIE2SN that is not verified in the cases of trans-SIE2SN and SIBSN. The calculated second enthalpies of deprotonation are similar for cis-SIE2SN and trans-SIE2SN, both resulting in the corresponding SIE2SI derivative, but are significantly larger than the first enthalpies of deprotonation calculated for cis-SNE2SN and trans-SNE2SN and than the second enthalpy of deprotonation calculated for the silicate trimer, SN3, yielding the SISNSI species [24].

Finally, from comparison of data in Figs. 1, 2 and in Tables S1 and S2, it is found that the geometrical parameters for the three ethenesilicates are much the same as those calculated for the benzenesilicates.

Views of the optimized geometries for the ethanesilicates and selected geometrical parameters are shown in Fig. 3 (full details are supplied in Table S3). Both in the neutral and in the anionic species, all the ethane derivatives present gauche-type configuration. Due to the formation of internal hydrogen bonds, the gauche configurations are much more stable than the anti configurations (not shown). The calculated enthalpies of deprotonation for the most and for the least stable ethanesilicates are close to their ethene counterparts suggesting that the stability of the silicates species in the gas phase is controlled by the number and nature of the internal hydrogen bonds in these compounds.

In the supplementary information (Tables S5-S10) are given the averages and standard deviations for each bond length and angle calculated over all the gaseous molecules considered above. For the discussion, we will focus separately on bond lengths and bond angles for three different regions, that is, silica groups (Tables S5 and S8, respectively), organic moieties (Tables S6 and S9, respectively), and the hybrid structure in the region between the organic and inorganic parts (Tables S7 and S10, respectively).

The analysis of the calculated parameters for the silica groups shows that the differences between the Si–O and O–H bond lengths and O–Si–O and Si–O–H angles in organosilicates and the mean values calculated for the different silicates are smaller than the differences between different silicates, for example, between linear, branched, or cyclic silicates [23, 24]. Thus, the geometrical differences in these bond lengths and bond angles due to the presence of the organic linkers are in the uncertainty intervals defined for the same properties in the silicates. This suggests that the parameters used in our previous MD simulations for such

bonds and angles are transferable to the species considered in this work. Similarly, the averages for the bond lengths and bond angles in the organic moieties of the several neutral and anionic organosilicates studied here are almost unchanged when compared with the same parameters in the unsubstituted parent organic molecules, that is, benzene, ethene, and ethane. Moreover, it is also evident that the structures of the organic part are negligibly affected by the charges of neighboring silica fragments. In fact, C–C and C–H bond lengths in the case of the neutral 1,4-bis(tri-hydroxysilyl)-benzene (Fig. 1) and 1,4-bis(trihydroxysi-lyl)-ethane (Fig. 3) species are identical to those in their respective anions, that is, differences in the bond lengths are lower than 0.005 Å. Similar findings are observed for the C=C and C–H bonds in neutral and charged 1,4-bis(trihydroxysilyl)-ethene (Fig. 2). This suggests that the parameters for the organic moieties are transferable from neutral to anion or to dianion species.

New parameters for the hybrid region between the organic and inorganic moieties had to be developed here. The analysis of average values in Tables S7 and S10 shows that the Si–C bond length and the C–Si–O angle are almost the same independent of the organic linker and that the Si–C bond lengths are elongated by ~ 0.04 Å upon deprotonation in adjacent silicate fragments. In charged silicate fragments, there are two types of C–Si–O angles, that is, C–Si–Oc and C–Si–Oh$_I$, while just a single type is found for neutral fragments, that is, C–Si–Oh$_N$. The C–Si–Oc angle is larger than the C–Si–Oh$_N$ angle by 5°–7° but the C–Si–Oh$_I$ angle is smaller than the C–Si–Oh$_N$ angle by 7°–9°. The variations found in the charged silicates are caused by the repulsion between the Oc atom and Oh or C atoms directly bonded to the same Si, which causes the elongation of bond lengths and the closing of angles in all charged organosilanes when compared with the parent neutral compounds. Finally, the angles C–C–Si and H–C–Si depend on the organic group, but do not depend on the type of Si atom (neutral or ionic).

3.2 Atomic point charges

The partial atomic charges were computed with three different standard methods, namely Mulliken, CHelpG, and NPA approaches. For further explanation about the differences and meaning of these methods, the reader is referred to ref [33]. However, it is meaningful to describe these features briefly in order to understand the discussion. Mulliken and NPA are molecular orbital-based methods, and NPA charges are preferred when the chemical nature of atoms is discussed. This is because the NPA charges are much more chemically reasonable when compared with the Mulliken charges; in Mulliken analysis, the overlap

population is evenly divided between the two neighboring atoms without considering their nature and electronegativity while NPA employs explicitly orthogonalized (natural) atomic orbitals and thus solves the overlap population problem [34]. The NPA method is almost basis set independent but the charges are unsuited for use when electrostatic properties are of interest (e.g., development of force fields for simulations in aqueous solutions) since their values are usually too positive or too negative when compared with charges obtained from methods which fit charges to match the electrostatic potential [35]. The CHelpG scheme fits the charges to the electrostatic potential calculated in a regularly spaced grid of points around the molecule under the constraint that the dipole moment is preserved. For this reason, CHelpG charges are suitable for use as point charges in molecular mechanics force fields for modeling in aqueous environment [36], although some erratic behaviors of CHelpG charges in the description of the chemical nature of silica compounds were reported [37], and were found to be superior to charges obtained from molecular orbital-based methods for obtaining multipole moments [35]. For these historical or practical importances, extensive tables with charges calculated with these three methods for each atomic type (average values) in the different gaseous molecules considered in this work are supplied as Supporting Information (Tables S11 to S14). Other methods for calculating point charges could be used but they were found to give unrealistic behaviors (e.g., Bader charges predicted a strongly ionic character for several all-silica crystal polymorphs) or predict charges that are too low for being compatible with other force fields (e.g., charges calculated with the Hirschfeld method) [38].

The following discussion will mainly consider the NPA charges which are more chemically intuitive for molecules in the gaseous phase than those calculated with the CHelpG approach. As expected from similar calculations performed for gaseous silicates [24], highly positive charges are calculated for silicon, negative charges are found on the oxygen atoms and small positive charges on the hydrogen atoms. Furthermore, small negative charges are calculated here for the carbon atoms. Charges for Oh$_N$, Oh$_I$, Ho$_N$, Ho$_I$, and Oc atom types in the organosilanes are close to those calculated for silicates [24]. On the other hand, Si charges connected to C atoms are lower than the previous values [24]. It is worth to point out here that there are some differences depending on the method considered for calculating the atomic charges.

Since we could regard the organic groups as bridging moieties, it is very interesting to compare the partial charges calculated for organosilanes with those for the silicate dimers calculated in our previous work [24] where

two silica (SN or SI) fragments are bridged by an oxygen atom. We can qualitatively mention that the bridging parts are always negatively charged, although the sign of the total charge in the silica fragments depends on the degree of deprotonation in the fragment. In the case of the silicates [24], the total charge of an SN fragment was $+0.632e$ in the case of the neutral dimer (SN2 species or, in the present notation, SNOSN) and $+0.535e$ in the case of an SNSI dimer (present notation is SNOSI). The total charge of an SI fragment was $-0.253e$ in the case of SNSI and $-0.347e$ in the case of the SI2 (present notation is SIOSI) molecule. The charges calculated for the bridging oxygens were $-1.265e$, $-1.282e$, and $-1.307e$ for SN2, SNSI, and SI2 species, respectively.

The calculated charges for the different fragments in the benzenesilicates are shown in Fig. 4. Clearly, the charges of the bridging parts in the benzenesilicates are less negative than those calculated for the bridging oxygen atom in the pure silicates. Thus, the organic linker is found to visibly affect the charges of the silica fragments, especially the atomic charges of the silicon atoms, which become less positive than silicon atoms in the silicate dimers. Similar findings are observed for cis-ethene, trans-ethene, and ethane organosilanes (Supporting Information). As it can be seen in Fig. 4, the total negative charge of the organic moiety is mainly due to the atomic charges of the C atoms bonded directly to the silicate fragments. Interestingly, the charges in the Si atoms are almost independent of the type of carbon atom to which it is bonded. As suggested in the previous section, the elongated Si–C bonds are caused by the negative charges on the deprotonated silicate fragments as it can be seen in Fig. 4, which cause electrostatic repulsion.

3.3 Potential parameters

The full set of potential parameters for the organosilicates is given in Table 2. The charges, bond lengths, and bond angles were obtained by averaging the CHelpG and geometric data calculated for the several different configurations of charged and neutral silicates and organosilicates, following our previous approach for silicates [21]. Some of the charges had to be slightly adjusted to ensure that each molecule retained its nominal charge. It is worth noticing that, despite the different approaches used, our values for the point charges of neutral silicates, shown in Table 2, are not very far from those proposed previously by Pereira et al. [39] based on calculations using a different DFT functional and a numerical basis sets.

Interestingly, the parameters regarding the silicate and the organic moieties in the organosilicates are identical to those obtained before for silicates (linear, branched, and cyclic) and to those available for the parent organic molecules (benzene, ethylene, and ethane), respectively. In fact, the most important changes refer to the parameters involving the linking of the organic and inorganic parts, that is, the Si–C bond length and the C–Si-O, C–C-Si, and H–C-Si bond angles. Fortunately, in terms of force field transferability, some of these parameters were found to be almost independent of which organic linker is considered, that is, the parameters are much the same if the linker is ethane, ethylene, or benzene. Importantly, it was found that the C–C-Si and H–C-Si bond angles depend on the organic linker but do not depend on the global charge (neutral, anion, or dianion) of the organosilicate species while the Si–C bond length and the C–Si-O bond angle do not

Fig. 4 Total charges in organic and silica fragments for **a** SNBSN, **b** SIBSN, and **c** SIBSI molecules. Atomic charges for C atoms bonded to Si atoms are also indicated
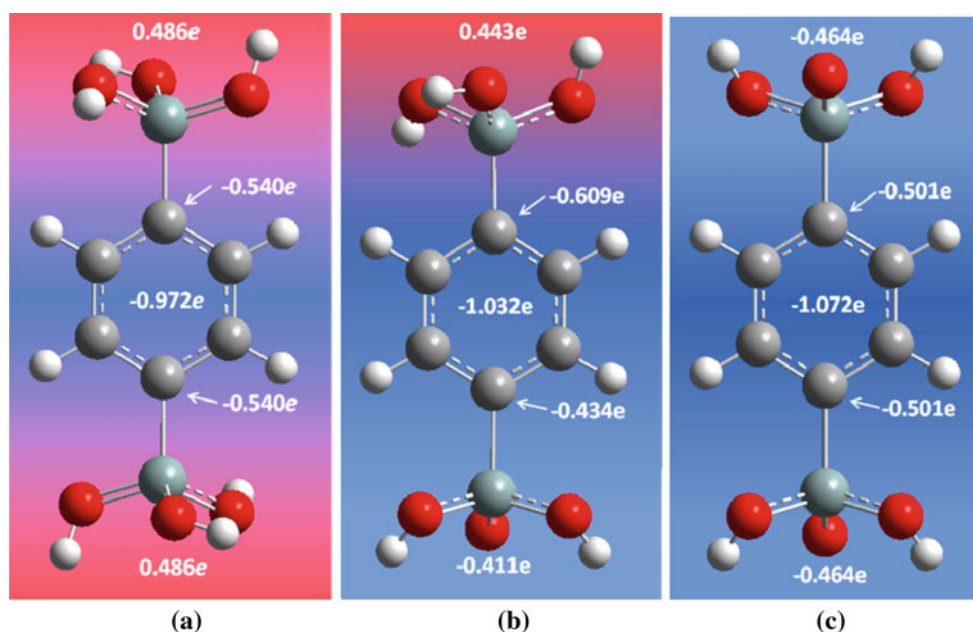
Table 2 Geometric parameters and point charges for neutral and anionic organosilicates

Atom	Charge (a.u.)	Bond	Length (nm)	Angle	Value (°)
Si_N	1.206	Si_N-O	0.165	Si_N-Oh_N-Ho_N	114.4
Si_I	1.332	Si_I-O	0.169	Si_I-Oh_I-Ho_I	109.8
Oh_N	−0.762	Si_I-O_c	0.158	Oh_N-Si_N-Oh_N	109.6
Oh_I	−0.839	Oh-Ho	0.097	Oh_I-Si_I-Oh_I	104.9
Oc	−1.046	Si_N-C	0.186	Oh_I-Si_I-O_c	114.2
H_{sp^3}	0.060	Si_I-C	0.190	Oh_N-Si_N-C	111.0
H_{sp^2}	0.115	C_{sp^3}—C_{sp^3}	0.153	Oh_I-Si_I-C	103.7
Ho_N	0.405	C_{sp^3}—H_{sp^3}	0.109	O_c-Si_I-C	116.4
Ho_I	0.376	C_{sp^2}—C_{sp^2}	0.134	Si—C_{sp^3}—C_{sp^3}	115.4
H_A(SN)	0.115	C_{sp^2}—H_{sp^2}	0.108	Si—C_{sp^2}—C_{sp^2}	126.0
H_A(SI)	0.115	C_A-C_A	0.140	Si-C_A-C_A	120.0
C_{sp^3}(SN)	−0.255	C_A-H_A	0.108	Si—C_{sp^3}—H_{sp^3}	109.9
C_{sp^3}(SI)	−0.480			Si—C_{sp^2}—H_{sp^2}	117.6
C_{sp^2}(SN)	−0.250			C_{sp^3}—C_{sp^3}—H_{sp^3}	110.7
C_{sp^2}(SI)	−0.475			H_{sp^3}—C_{sp^3}—H_{sp^3}	107.8
C_A(SN)	−0.135			C_{sp^2}—C_{sp^2}—H_{sp^2}	120.0
C_A(SI)	−0.360			C_A-C_A-C_A	120.0
C_A(H,SN)	−0.115			C_A-C_A-H_A	120.0
C_A(H,SI)	−0.115				

depend on the organic linker but depend on the global charge of the organosilicate.

In conclusion, the potential parameters optimized in this work for molecular simulations involving organosilanes in solution can be combined with those developed before for silicates possessing different geometries, for example, linear, branched, cyclic, and different degrees of deprotonation.

4 Conclusions

This paper presents the results of a density functional theory study, with a considerably large basis set, on the geometries, energies, and point charges of neutral, singly deprotonated, and doubly deprotonated organosilicate molecules (1,4-bis(trihydroxysilyl)-benzene, bis(trihydroxysilyl)-ethene, and bis(trihydroxysilyl)-ethane) in the gaseous phase. Several possible starting geometries were considered to assure that the most stable conformer was duly characterized for each of these species. The bond lengths and angles for the neutral and anionic organosilanes were discussed in detail. The analysis of the optimized structures for the organosilanes shows that the local geometries of the silica fragments of the organosilicate are similar to those found in the pure silicate compounds and that the local geometries of the organic part of the organosilicates are similar to those found in the parent organic

molecules. Since our aim is to develop a new potential for classical molecular dynamics simulations of organosilicates in solution, these similarities are very important for transferability of the potential.

From the calculated enthalpies of deprotonation, it is concluded that the 1,4-bis(trihydroxysilyl)-benzene species has an acid–base behavior that is intermediate of those calculated for the SN (silicate monomer) and for the SN2 (silicate dimer). The acid–base behavior of the 1,4-bis(trihydroxysilyl)-benzene is similar to those found for the least stable silanes with ethene or ethane as the organic linker, that is, it is similar to the behaviors of trans-bis(trihydroxysilyl)-ethene and anti-bis(trihydroxysilyl)-ethane compounds, respectively. For the most stable conformers of the latter organosilanes, that is, cis-bis(trihydroxysilyl)-ethene and gauche-bis(trihydroxysilyl)-ethane, the calculated enthalpies of deprotonation are lower than those calculated for the SN and SN2 species, suggesting a behavior that is intermediate of those found for small (up to 2 Si atoms) and larger pure silicates.

The partial atomic charges of the organosilicates were calculated with three different approaches. From a natural population analysis, it is found that the calculated charges of the Si atoms are almost independent of the type of carbon atom to which they are bonded but the charges on the neutral silicate fragments are less positive and the charges on the anionic silicate fragments are more negative when compared with those in the pure silicates, either the neutral SN2 or the singly deprotonated SNSI species. Finally, since the silicate fragments are separated by the organic linker in the case of the organosilanes, the charge variation in the SN unit on going from SNXSN to SNXSI and the charge variation in the SI unit on going from SIXSN to SIXSI are less important than the variations noticed for the pure silicate species.

References

1. Iijima S (1991) Nature 354:56
2. Li H, Eddaoudi M, O'Keeffe M, Yaghi MO (1999) Nature 402:276
3. Beck JS, Vartuli JC, Roth WJ, Leonowicz ME, Kresge CT, Schmitt KD, Chu CTW, Olson DH, Sheppard EW, McCullen SB (1992) J Am Chem Soc 114:10834
4. Gregg SJ, Sing KSW (1981) Adsorption, surface area and porosity. Academic, London

5. Zelenak V, Halamova D, Gaberova L, Bloch E, Llewellyn P (2008) Micropor Mesopor Mater 116:358
6. Luo J, Xu H, Liu Y, Zhao Y, Daemen LL, Brown C, Timofeeva TV, Ma S, Zhou HC (2008) J Am Chem Soc 130:9626
7. Park SJ, Jang YS (2003) J Coll Int Sci 261:238
8. Sayari A (1996) Chem Mater 8:1840
9. Chmiola J, Yushin G, Gogotsi Y, Porter C, Simon P, Taberna PL (2006) Science 313:1760
10. Kaneko K, Fukuzaki N, Kakei K, Suzuki T, Ozeki S (1989) Langmuir 5:960
11. Iiyama T, Nishikawa K, Suzuki T, Otowa T, Hijiriyama M, Nojima Y, Kaneko K (1997) J Phys Chem B 101:3037
12. Striolo A, Gubbins KE, Gruszkiewicz MS, Cole DR, Simonson JM, Chialvo AA, Cummings PT, Burchell TD, More KL (2005) Langmuir 21:9457
13. Inagaki S, Guan S, Fukushima Y, Ohsuna T, Terasaki O (1999) J Am Chem Soc 121:9611
14. Melde BJ, Holland BT, Blanford CF, Stein A (1999) Chem Mater 11:3302
15. Asefa T, MacLachlan MJ, Coombs N, Ozin GA (1999) Nature 402:867
16. Inagaki S, Guan S, Ohsuna T, Terasaki O (2002) Nature 416:304
17. Wang W, Lofgreen JE, Ozin GA (2010) Small 6:2634
18. Comotti A, Bracco S, Valsesia P, Ferretti L, Sozzani P (2007) J Am Chem Soc 129:8566
19. Martinez U, Pacchioni G (2010) Micropor Mesopor Mater 129:62
20. Shirai S, Goto Y, Mizoshita N, Ohashi M, Tani T, Shimada T, Hyodo S, Inagaki S (2010) J Phys Chem A 114:6047
21. Jorge M, Gomes JRB, Cordeiro MNDS, Seaton NA (2007) J Am Chem Soc 129:15414
22. Jorge M (2008) Langmuir 24:5714
23. Jorge M, Gomes JRB, Cordeiro MNDS, Seaton NA (2009) J Phys Chem B 113:708
24. Gomes JRB, Cordeiro MNDS, Jorge M (2008) Geochim Cosmochim Acta 72:4421
25. Becke AD (1993) J Chem Phys 98:5648
26. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785
27. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery Jr JA, Vreven T, 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, Bakken V, 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 (2004) Gaussian 03, Revision C.01, Gaussian, Inc., Wallingford
28. Šefčík J, McCormick AV (1997) AIChE J 43:2773
29. Mulliken RS (1955) J Chem Phys 23:1833
30. Breneman CM, Wiberg KB (1990) J Comput Chem 11:361
31. Reed AE, Weinstock RB, Weinhold FA (1985) J Chem Phys 83:735
32. Cerveau G, Corriu RJP, Dabiens B, Bideau JL (2000) Angew Chem Int Ed 39:4533
33. Jensen F (2007) Introduction to computational chemistry. Wiley, Chichester
34. Guerra CF, Handgraaf J-W, Baerends EJ, Bickelhaupt FM (2003) J Comput Chem 25:189
35. Sigfridsson E, Ryde U (1998) J Comput Chem 19:377
36. Carlson HA, Nguyen TB, Orozco M, Jorgensen WL (1993) J Comput Chem 14:1240
37. Murashov V (2003) J Mol Struct 650:141
38. Zwijnenburg MA, Bromley ST, van Alsenoy C, Maschmeyer T (2002) J Phys Chem A 106:12376
39. Pereira JCG, Catlow CRA, Price GD (2002) J Phys Chem A 106:130

REGULAR ARTICLE

Analysis of the origin of lateral interactions in the adsorption of small organic molecules on oxide surfaces

José J. Plata · Veronica Collico · Antonio M. Márquez · Javier Fdez Sanz

Abstract A decomposition scheme is proposed to analyze the physical contributions to the decrease in the binding energy of chemisorbed species with increasing coverage. This scheme is applied to the acetaldehyde–TiO_2 (110) rutile system as a model for other small organic molecule—oxide surface systems. Different density functional theory (DFT) functionals have been employed at both low-medium and high coverages to understand how the different theoretical descriptions of the various terms influence the adsorbate–surface interaction. At low coverages, it is found that the localized adsorbate to surface electron donation is the fundamental physical process that influences the adsorbate–surface interaction. This results shows that while it is usually assumed that only pairwise adsorbate–adsorbate interactions influence the adsorption energy, the progressive modification of the surface properties (surface reduction in this case) may also play a significative role. The DFT+U functional results, in this case, in the best agreement with the experimental binding energy, and the inclusion of the dispersive forces results in largely overestimated adsorption energies. At higher coverages, the pure GGA and GGA+U functionals overestimate the repulsive terms and the computed binding energy is well below the experimental data. The inclusion of the dispersive forces is required to correctly reproduce the experimental results. The contributions of the different physical terms are also analyzed.

Keywords DFT · Van der Waals · Adsorption · Lateral interactions · TiO_2 · Acetaldehyde

1 Introduction

The interaction of atoms and molecules with surfaces is ubiquitous, since all materials interact with their environment via their surfaces. For this reason, the adsorption of small organic molecules on the surface of solids is one of the central and more important topics in heterogeneous catalysis, photocatalysis, self-assembled layers, dye sensitized solar cells, and sensors [1–10].

It is widely known that the properties of adsorbed species are not independent of coverage [11]. The strength of the adsorbate-surface bond can be perturbed by the presence of neighboring species. Adsorbate–adsorbate interactions can be purely electrostatic or chemical in origin. Dipole–dipole coupling or charge–dipole interactions belongs to the first group while the decreasing in the orbital overlap between the adsorbate and the surface as the coverage increases is included in the second group. Such adsorbate–adsorbate or "lateral interactions" are responsible of the non-ideal behavior of the adsorption process with increasing surface coverage. Its effects are usually quantified as a perturbation of the adsorption energy of the adsorbate in absence of lateral interactions (i.e., at the limit of zero coverage). While the importance of such interactions is widely recognized [11], a detailed physical understanding of the different factors that influence the binding of adsorbates with increasing surface coverage is a

J. J. Plata (✉) · V. Collico · A. M. Márquez · J. F. Sanz
Departamento de Química Física, Universidad de Sevilla,
41012 Sevilla, Spain
e-mail: jplata@us.es

Springer

topic that remains to be addressed. Here, we propose a theoretical scheme that allows to analyze the observed decrease in the adsorbate to surface binding energy into different physical contributions. Different density functional theory (DFT) models are employed to quantify the influence of the theoretical description in these physical contributions at low- and high-surface coverages.

As a benchmark system, we have used the adsorption of acetaldehyde on rutile TiO$_2$ (110) surface. Titanium dioxide has received considerable attention in the literature due to its utility as a catalyst and as a photocatalyst. In particular, the rutile TiO$_2$ (110) surface has become a prototype system in surface science studies of metal oxide surfaces [12–14]. The (110) surface of rutile TiO$_2$ is composed of alternating rows of twofold coordinated bridging O$_b$ atoms and channels that expose both fivefold coordinated Ti (Ti$_{5c}$) and in-plane threefold coordinated O atoms (see Fig. 1). Moreover, the surface chemistry and photochemistry of organic compounds on the TiO$_2$ surface is a topic of recurrent interest in the literature. In particular, acetaldehyde photodecomposition is of singular importance because it is commonly found as a volatile contaminant, aside from being present either as reactant, intermediate, or product in many catalytic processes [15–19]. The chemistry and photochemistry of acetaldehyde on the rutile TiO$_2$ (110) surface has been recently studied from a experimental point of view by Henderson [20].

The rest of this paper was organized as follows. In Sect. 2, the methodology and the surface model is described. In Sect. 3, the partition model, that allow to analyze the physical contributions to the inter-adsorbate lateral interactions, is presented. In Sect. 4, we present and discuss the

results obtained. Finally, concluding remarks were given in Sect. 5.

2 Computational details

In order to model the extended nature of these surfaces, periodic three-dimensional DFT calculations were carried out using the VASP 5.2 code [21–23] with the projector augmented wave (PAW) method [24, 25]. In these calculations, the energy was obtained using the generalized gradient approximation (GGA) implementation of DFT proposed by Perdew et al. [26], and the electronic states were expanded using plane wave basis set with a cutoff of 400 eV. Forces on the ions were calculated through the Hellmann–Feyman theorem as the partial derivatives of free energy with respect to the atomic coordinates, including the Harris-Foulkes correction to forces [27]. This calculation of the force allows a geometry optimization using the conjugated gradient scheme. Iterative relaxation of the atomic positions was stopped when the forces on the atoms were less than 0.02 eV/Å. In order to analyze the charge transfer upon adsorption, the total charge of the acetaldehyde molecule was determined using the algorithm introduced by Henkelman et al. [28, 29] for the evaluation of the Bader charges [30].

As gradient corrected density functional theory is known to assign a delocalized character to the excess electrons induced on the TiO$_2$ surface by defects like O vacancies or adsorbates [31, 32], the GGA+U formalism was also employed. The Hubbard U term was added to the plain GGA functional using the rotationally invariant approach proposed by Dudarev et al. [33], in which the Coulomb U and exchange J parameters are combined into a single parameter $U_{\text{eff}} = U - J$. Recent work by Deskins et al. [34] has shown how the main effect of the use of this approach is to determine the location of gap states within the gap, and that reasonable gap states occur for values of U_{eff} between 3.3 and 4.1 eV. However, the trends are not strongly affected by the particular choice of the U_{eff} within this range. To be consistent with previous work on our group, a U_{eff} parameter of 4.5 eV was chosen, as it quantitatively reproduces the experimental position of the Ti 3d levels observed in the valence photoemission spectra of Ce/TiO$_2$(110) [35].

In order to understand how dispersion forces modify the description of lateral interactions, the van der Waals density functional (vdW-DF) derived by Dion et al. [36] was employed. Recently, there has been an enormous progress in the treatment of dispersion forces in DFT [37]. Adsorption on solid surfaces is an area where great steps forward have been made in this respect, but there are still challenges for dispersion–based DFT methods at present.
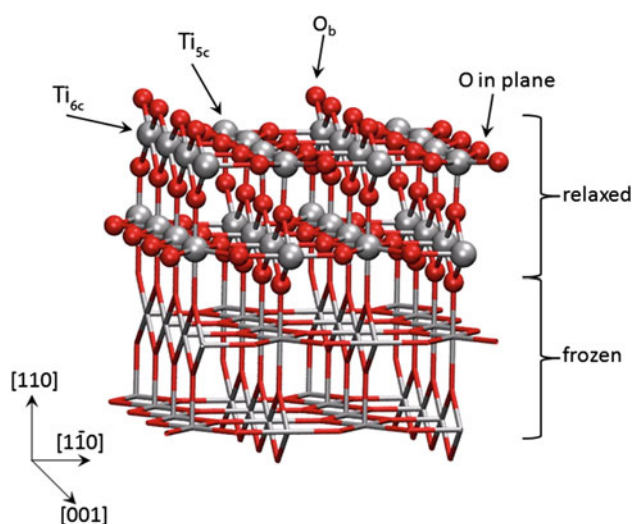

Fig. 1 4 × 2 supercell model of a TiO$_2$ (110) surface with four O-Ti-O layers; the two *lower layers* are kept frozen and the two *upper layers* are fully optimized. Atom colors *red* O, *gray* Ti

The functional employed in this work (optB86b-vdW) approximately accounts for dispersion interactions, it is a solution which balances computational efficiency and accuracy [38] and it is known to describe accurately structural properties of both gas-phase clusters and bulk materials [39].

It is known [40] that vacancy formation energies based on slab calculations show an oscillating behavior with the number of layers and that at least six TiO$_2$ layers are required to obtain fully converged values. In previous works, we did not find big differences between four and six layers for acetone adsorption [41] so in the present work, we used supercell models with four TiO$_2$ tri-layers to represent the TiO$_2$ (110) surface and explore the energetics and geometries of adsorption of acetaldehyde on this surface. Each slab was separated by a vacuum of 13 Å, considered enough to avoid interaction between the slabs [40]. In all cases, the optimized lattice parameters for the bulk used were $a = 4.616$ Å, $c = 2.974$ Å, and $u = 0.304$ Å. The a and c parameters were kept fixed during the surface atomic positions relaxation. Multiples of the unit cell along the [001] direction give cells of the n × 1 type and j doubling along the [110] direction gives n × 2 cells. The surface cell that were used in the present work was 4 × 2 size, where adsorption of a single acetaldehyde molecule represents a coverage of $\theta = 0.125$. The calculations for the supercell model where computed at the Γ point of the Brillouin zone. Adsorption energy for the incoming acetaldehyde molecule is computed as

$$E_{\text{ads}}(\theta_2) = E_{\text{total}}(\theta_2) - E_{\text{total}}(\theta_1) - E_{\text{acetal}} \qquad (1)$$

where $E_{\text{total}}(\theta_i)$, $i = 1, 2$, are the total energies of model systems with $N-1$ and N acetaldehyde molecules, respectively, and E_{acetal} is the total energy of an isolated acetaldehyde molecule. With this definition, negative adsorption energies represent bound states stable with respect to desorption, and a direct comparison with experimental desorption energies derived from thermal programmed desorption (TPD) experiments is possible.

3 Partition model

This section introduces a simple model to analyze the coverage dependence in the adsorption energies into physically meaningful contributions. The main objective is to devise a simple model that allow us to understand the importance of the different physical interactions, their relation to the level of theory used, and their role at different degrees of surface coverage. Detailed experimental TPD data available for acetone and acetaldehyde on the (110) surface of rutile TiO$_2$ shows that binding energy, even a low coverages (0.01–0.50 ML), decrease with

increasing surface loading [20, 42]. Previous theoretical work has related this effect to the bonding mechanism of adsorption of those carbonyl compounds to the rutile surface [41, 43]. The adsorption of organic carbonyl compounds at the TiO$_2$ rutile (110) surface takes place through interaction of the CO dipole moment with the electric field of the Ti$_{5c}$ surface cations. For small coverages, a tiny charge transfer from the organic molecule adlayer to the surface takes place, reducing the acidity of the remaining Ti$_{5c}$ sites and, thus, the binding energy of incoming molecules. However, the role of the long-range dipole–dipole and short-range steric repulsive interactions was not analyzed. Taking into account these considerations, the decrease in the adsorption energy with increasing coverage can be expressed as

$$E_{\text{ads}}(\theta_2) = E_{\text{ads}}(\theta_1) + f_{s+d} + f_e \qquad (2)$$

where $E_{\text{ads}}(\theta_i)$ is the adsorption energy at coverage θ_i (with $\theta_2 > \theta_1$), and f_{s+d} and f_e are contributions for steric and dispersion forces (f_{s+d}) and electronic (adsorbate to surface charge transfer, f_e) physical contributions above described. The f_{s+d} contribution can be estimated as the energy difference per molecule between two isolated acetaldehyde monolayers whose geometries are identical to those in surface + adsorbate systems with coverages θ_i

$$f_{s+d} = E_{\text{ML}}(\theta_2) - E_{\text{ML}}(\theta_1) \qquad (3)$$

It should be taken into account that the electronic density on the acetaldehyde molecules is changed upon adsorption, thus the above approximation holds while the adsorbate to surface charge transfer is not too high, as it is our case. The contribution of the adsorbate–surface charge transfer is then easily obtained as

$$f_e = E_{\text{ads}}(\theta_2) - E_{\text{ads}}(\theta_1) - f_{s+d} \qquad (4)$$

4 Results and discussion

4.1 Low-medium coverages

A single acetaldehyde molecule on our 4 × 2 supercell model provides the lower coverage explored, 0.125 ML. The pure GGA functional produces a binding energy of −0.83 eV, slightly lower than the experimental estimated value of −0.93 eV (see Table 1). Addition of a second acetaldehyde molecule results in a coverage of 0.250 ML. The computed binding energy of this second acetaldehyde molecule is reduced (−0.64 to −0.74 eV) with respect to the first adsorbed molecule, in agreement with the observed reduction in the experimental TPD data (−0.81 eV). The reduction in binding energy can be ascribed to steric and dipole–dipole repulsive interactions between the adsorbed molecules but also to the adsorbate to surface charge

Table 1 Computed (GGA, GGA+U, and optB86b-vdW) and experimental acetaldehyde adsorption energies (E_{ads}) at different surface coverages

θ (ML)	E_{ads} (eV)			
	GGA	GGA+U	+vdW	Exp. [20]
0.125	−0.83	−1.02	−1.23	−0.93
0.250				
(a)	−0.64/−0.69	−0.86/−0.88	−1.16/−1.21	−0.81
(b)	−0.70/−0.74	−0.91/−0.96	−1.19/−1.23	
1.000	−0.20	−0.37	−0.82	−0.71

Acetaldehyde molecules are adsorbed in the same (a) or different (b) channels of the TiO$_2$(110) surface
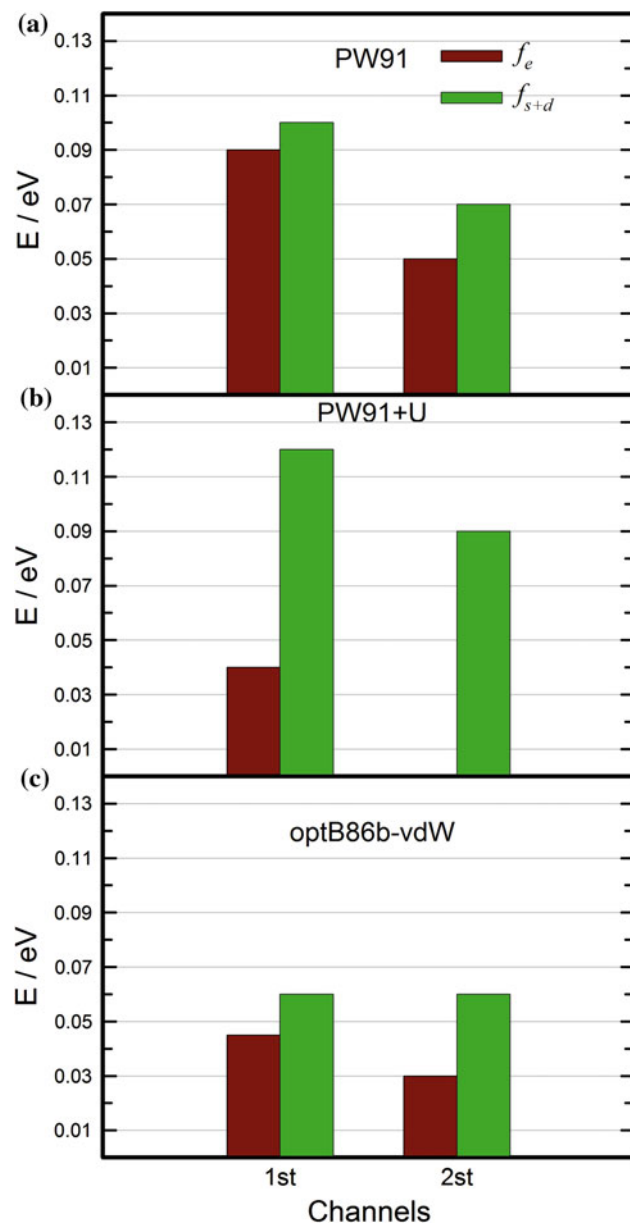
Fig. 2 Histogram for inter-molecular lateral interaction terms at low coverage ($\theta = 0.250$ ML) using different functionals. *Brown columns* electronic contribution. *Green columns* steric and dispersion contributions

transfer that takes place upon acetaldehyde adsorption. Using the above proposed scheme, we can quantify how these factors contribute to the observed reduction in the strength of the binding between the acetaldehyde molecule and the TiO$_2$ surface. As shown in Fig. 2, repulsive interactions are larger when the incoming molecule adsorbs in the already occupied channel and both contributions are of the same order and larger in this case. Clearly, the distance between molecules is shorter and the dipole–dipole and steric repulsions must be higher in this situation. The contribution of the adsorbate-surface charge transfer is, also, significantly larger showing how, upon acetaldehyde adsorption, the small charge transfer that takes place delocalizes over the nearby Ti$_{5c}$ cations, reducing their acidity and, thus, their ability to bind incoming acetaldehyde molecules.

Gradient corrected density functional theory has known limitations in the description of electronic defects related to inherent deficiencies in the functionals, mainly the insufficient cancelation of the self-interaction energy [44]. As result, the electron density transferred by the incoming acetaldehyde molecules will be excessively delocalized over the surface and reduce the acidity of Ti cations in an unrealistic way. In consequence, f_e could be increased in an unrealistic way. The use of hybrid exchange-correlation functionals or the DFT+U approach, which adds an adjustable correction to enhance electron localization, allows to overcome these deficiencies.

At a coverage of 0.125 ML, the GGA+U computed adsorption energy is −1.02 eV which is a bit higher than the experimental value of −0.81 eV. The DFT+U approximation enhances the ionicity of the oxide and the charge transfer between the acetaldehyde molecule and the surface. Both effects reinforce the chemisorption bond of the acetaldehyde molecule to the surface resulting in the observed increase in the computed binding energy. At a 0.250 ML coverage, the GGA+U binding energies (see Table 1) are also higher than those obtained with the pure GGA functional, but in this case they are closer to the experimental value. The energetic decomposition of the

lateral interaction shows a qualitatively different picture of the inter-molecule interaction at this GGA+U theory level. While the f_e and f_{s+d} contributions where similar when the pure GGA functional was used, the inclusion of the U parameter drastically reduces the importance of the adsorbate to surface charge transfer (see Fig. 2b). The DFT+U approximation forces the localization of the electrons donated by the acetaldehyde molecule in the Ti cation in which the molecule is adsorbed and little electron density is delocalized to nearby Ti_{5c} cations. In fact, the f_e contribution is reduced by a 50 % factor when the second acetaldehyde molecule gets adsorbed in the same channel than the first one and no f_e contribution is obtained when adsorption takes place in the empty channel.

Dispersion forces are long-range attractive forces that are known to be important in order to correctly describe the stability of a wide range of systems, including the adsorption of molecules to solid surfaces [39, 45]. However, the description of these forces, purely quantum mechanical in origin, pose a challenge to classical DFT theory as, in most implementations, the exchange–correlation potential does not include contributions from unshared electronic distributions [44]. The optB86b-vdW functional has been chosen to evaluate the effect of such dispersion forces in the present problem. At the lower coverage employed (0.125 ML), the computed adsorption energy is now -1.23 eV, which is a 35 % higher than the experimental value. At a 0.250 ML coverage, the computed adsorption energies (see Table 1) are still largely overestimated but they show the same decreasing tendency as the experimental data. This overestimation of the adsorption energy in vdW-DF has been observed previously by other authors and has been adscribed to a double counting of some short-range exchange-correlation effects [46]. The decomposition of the inter-molecule lateral interactions obtained with our proposed scheme is shown in Fig. 2c. The adsorbate to surface charge transfer term (f_e) decreases significatively when compared with the pure GGA result. This can be related to an increase in the Ti-OC bond distance from 2.16 Å (GGA) to 2.19 Å that will reduce the orbital overlap and, thus, the amount of charge transfer that takes place. Another observation is that, as in the pure GGA functional, the transferred electron density delocalizes over all nearby Ti_{5c} cations, including those in the secondary channel. The dipole–dipole and steric term (f_{s+d}) shows also a significative reduction, as could be expected from the inclusion in the functional of the attractive London dispersion forces. The decrease is more significative when the two acetaldehyde molecules are adsorbed in the same channel as expected from the nature of these dispersion forces [44].

4.2 Full coverage

In the previous section, we have seen how the inclusion of the Hubbard U term and the London dispersion forces can modify the values of adsorption energy and the description and analysis of lateral interactions at low and medium coverages. Nevertheless, the consideration of these effects should be even most important at high coverages when the partial reduction of the surface and dispersion forces are bigger. For this reason, the interaction of acetaldehyde molecule with the TiO_2 surface has been examined at a coverage of $\theta = 1$ ML. To adequately compare our results with the experimental values, the acetaldehyde molecules have been placed alternating their α hydrogen to either side of the Ti_{5c} channels, in order to produce the most stable configuration. (see Fig. 3).

Thermal programmed desorption data revealed that the peak that appears at 245 K corresponds to acetaldehyde desorption at $\theta = 1$ ML [20]. Redhead analysis [47] of these peaks assuming a typical pre-exponential factor of
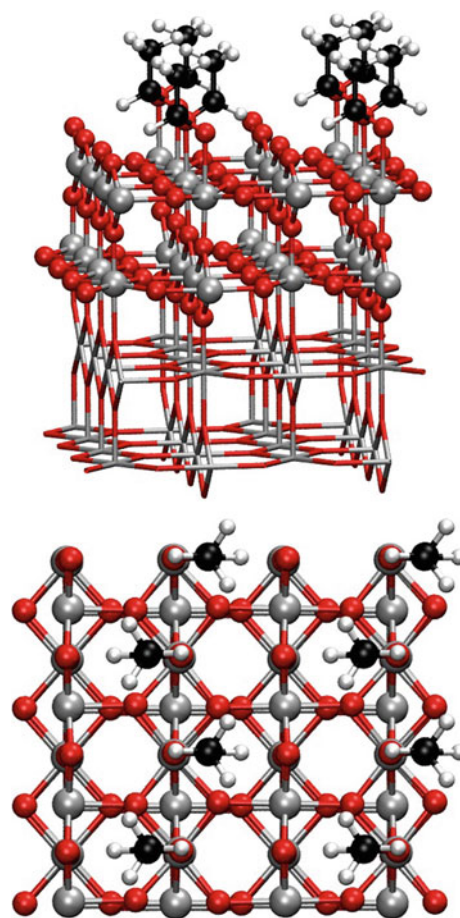
Fig. 3 Side and top views of and acetaldehyde monolayer on TiO_2 (110) surface at $\theta = 1$ ML. Atom colors *red* O, *gray* Ti, *black* C
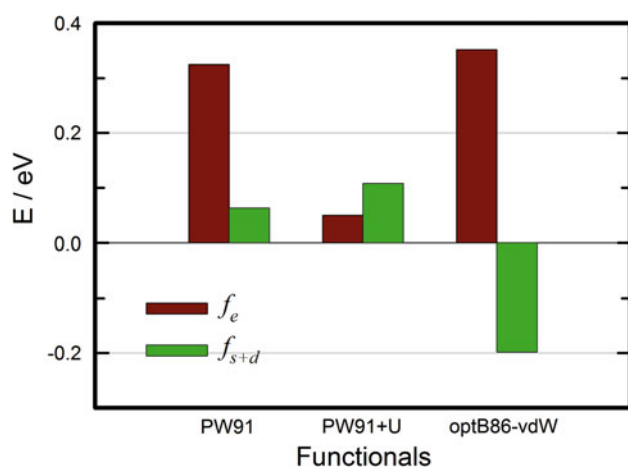
Fig. 4 Histogram for inter-molecular lateral interaction terms at full coverage ($\theta = 1.0$ ML)

10^{13} produces an estimated experimental adsorption energy of -0.71 eV. Table 1 show the calculated adsorption energies for an acetaldehyde coverage of 1 ML with the different functionals used in this work. Both the pure GGA and the GGA+U functionals produce adsorption energies that are well below the experimental data. This is not surprising as the short-range repulsive forces must dominate when the surface is highly crowded. However, the attractive dispersion forces must compensate these repulsive terms, but they are not included in these two functionals. The decomposition of the inter-molecule interaction as proposed in this paper, shown in Fig. 4, reveals that the adsorbate to surface charge transfer term (f_e) is dominant in the case of the GGA functional. The surface is highly reduced and the electrostatic interaction between the CO dipole moment and the Ti_{5c} charge is reduced. This term is largely reduced to a magnitude similar to that found at lower coverages when the GGA+U functional is used. The charge donated from the acetaldehyde overlayer to the Ti_{5c} cations is localized, and the free surface cations are only slightly affected. The f_{s+d} term has a smaller magnitude, compared to low-coverage results, when the pure GGA functional is used. This is probably related to an increased distance of 2.33 Å between the Ti cations and the acetaldehyde overlayer compared to 2.16 Å at 0.125 ML coverage that will reduce the repulsive steric interactions between the acetaldehyde molecules and the protruding oxygen atoms. In contrast to these results, the optB86b-vdW functional produces a binding energy of -0.82 eV that overestimates the experimental value by only a 15 % (this is probably within the experimental error bar). The decomposition of the inter-molecule lateral interactions (see Fig. 4) shows that the term related to the surface reduction is similar in magnitude to the data produced by the pure GGA functional. As the surface is highly

crowded and this functional does not include an electron localization term, the degree of surface reduction must be similar in both cases. In fact, the Ti-OC bond length is slightly larger in this case (2.36 Å) reflecting the slightly larger f_e term. Not surprisingly the steric term f_{s+d} is now negative: the high density of acetaldehyde molecules at the surface increases the dispersion attractive term and this term overcompensates the repulsive dipole–dipole and steric repulsive interactions.

5 Conclusions

The physical origin of the so-called lateral interactions, responsible of the decrease in the adsorbate-surface binding energy with increasing coverage has been examined in the acetaldehyde–TiO_2 (110) rutile system. A decomposition scheme is proposed that allows to understand how the progressive reduction in the oxide surface and the inter-molecule dipole–dipole and steric repulsive forces influence the binding energy at low-medium and high coverages. Three different DFT functional have been used: pure GGA, GGA+U, and the optB86b-vdW that includes a treatment of the attractive dispersion forces. The results show that at low-medium coverages the pure GGA and the optB86b-vdW functionals over delocalize the electron density donated from the adsorbate adlayer to the surface, resulting in an artificially reduced binding energy. The optB86b-vdW functional overbinds the acetaldehyde molecules, a result already observed and that has been ascribed to a double counting of some exchange-correlation terms. The inclusion of the U Hubbard term in the GGA+U functional overcomes the excessive electron delocalization observed in the pure GGA functional, and this theory level seems to be the most adequate at these surface coverages. At higher coverages, both GGA and GGA+U functional largely overestimate the repulsion forces, resulting in binding energies too low compared to experimental values. The decomposition of the inter-molecule interactions, as proposed in this paper, reveals that the adsorbate to surface charge transfer term (f_e) is dominant in the pure GGA functional case. The surface is being reduced by the already adsorbed acetaldehyde molecules, and the incoming molecules find cationic sites with reduced acidity, resulting in the observed lower binding energies. At these high loadings, the inclusion of the dispersion forces is found to be essential to adequately describe the highly crowded acetaldehyde overlayer. The computed optB86b-vdW binding energy overestimates the experimental data by a small amount (15 %) and is probably within the experimental error bar. The decomposition of the inter-molecule lateral interactions shows that although the adsorbate to surface charge transfer term is still important,

it is largely compensated by the attractive dispersion contribution introduced in this functional.

References

1. Mallat T, Baiker A (2004) Chem Rev 104:3037–3058
2. Hashmi ASK, Hutchings GJ (2006) Angew Chem Int Ed 45:7896–7936
3. Hoffmann MR, Martin ST, Choi WY, Bahnemann DW (1995) Chem Rev 95:69–96
4. Kubacka A, Bachiller-Baeza B, Colón G, Fernández-García M (2009) J Phys Chem C 113:8553–8555
5. Schreiber F (2000) Prog Surf Sci 65:151–256
6. Smith RK, Lewis PA, Weiss PS (2004) Prog Surf Sci 75:1–68
7. O'Regan B, Grätzel M (1991) Nature 353:737–740
8. Bach U, Lupo D, Comte P, Moser JE, Weissortel F, Salbeck J, Spreitzer H, Grätzel M (1998) Nature 395:583–585
9. Hagleitner C, Hierlemann A, Lange D, Kummer A, Kerness N, Brand O, Baltes H (2001) Nature 414:293–296
10. Li J, Lu YJ, Ye Q, Cinke M, Han J, Meyyappan M (2003) Nano Lett 3:929–933
11. Chorkendor I, Niemantsverdriet J (2003) Concepts of modern catalysis and kinetics. Wiley-VCH, Weinheim
12. Diebold U (2003) Surf Sci Rep 48:53–229
13. Linsebigler AL, Lu GQ, Yates JT (1995) Chem Rev 95:735–758
14. Thompson TL, Yates JT (2006) Chem Rev 106:4428–4453
15. Sauer ML, Ollis DF (1996) J Catal 158:570–582
16. Falconer JL, Magrini-Bair KA (1998) J Catal 179:171–178
17. Fujiwara N, Friedrich KA, Stimming U (1999) J Elec Chem 472:120–125
18. Asahi R, Morikawa T, Ohwaki T, Aoki K, Taga Y (2001) Science 293:269–271
19. Deluga GA, Salge JR, Schmidt LD, Verykios XE (2004) Science 303:993–997
20. Zehr RT, Henderson MA (2008) Surf Sci 602:2238–2249
21. Kresse G, Furthmuller J (1996) Phys Rev B 54:11169–11186
22. Kresse G, Furthmuller J (1996) Comput Mater Sci 6:15–50
23. Kresse G, Hafner J (1993) Phys Rev B 47:558–561
24. Kresse G, Joubert D (1999) Phys Rev B 59:1758–1775
25. Blöchl PE (1994) Phys Rev B 50:17953–17979
26. Perdew JP, Chevary JA, Vosko SH, Jackson KA, Pederson MR, Singh DJ, Fiolhais C (1992) Phys Rev B 46:6671–6687
27. Harris J (1985) Phys Rev B 31:1770–1779
28. Henkelman G, Arnaldsson A, Jonsson H (2006) Comput Mater Sci 36:354–360
29. Sanville E, Kenny SD, Smith R, Henkelman G (2007) J Comput Chem 28:899–908
30. Bader RFW (1985) Acc Chem Res 18:9–15
31. Rasmussen MD, Molina ML, Hammer B (2004) J Chem Phys 120:14583
32. Ganduglia-Pirovano MV, Hofmann A, Sauer J (2007) Surf Sci Rep 62:219–270
33. Dudarev SL, Botton GA, Savrasov SY, Humphreys CJ, Sutton AP (1998) Phys Rev B 57:1505–1509
34. Deskins NA, Rousseau R, Dupuis M (2009) J Phys Chem C 113:14583–14586
35. Park JB, Graciani J, Evans J, Stacchiola D, Ma SG, Liu P, Nambu A, Sanz JF, Hrbek J, Rodriguez JA (2009) Proc Natl Acad Sci 106:4975–4980
36. Dion M, Rydberg H, Schrder E, Langreth DC, Lundqvist BI (2004) Phys Rev Lett 92:246401
37. Klimeš J, Michaelides A (2012) J Chem Phys 137:120901
38. Román-Pérez G, Soler JM (2009) Phys Rev Lett 103:096102
39. Klimeš J, Bowler DR, Michaelides A (2011) Phys Rev B 83:195131
40. Oviedo J, Miguel MAS, Sanz JF (2004) J Chem Phys 121:7427–7433
41. Márquez AM, Plata JJ, Sanz JF (2009) J Phys Chem C 113:19973–19980
42. Henderson MA (2004) J Phys Chem B 108:18932–18941
43. Plata JJ, Collico V, Márquez AM, Sanz JF (2011) J Phys Chem C 115:2819–2825
44. Koch W, Holthausen MC (2002) A chemist guide to density functional theory. Wiley, Mörlenbach
45. Tkatchenko A, Romaner L, Hofmann OT, Zojer E, Ambrosch-Draxl C, Scheffler M (2010) MRS Bull 35:435–442
46. Göltl F, Hafner J (2011) J Chem Phys 134:064102
47. Redhead PA (1962) Vacuum 12:203

REGULAR ARTICLE

Numerical investigation of the elastic scattering of hydrogen (isotopes) and helium at graphite (0001) surfaces at beam energies of 1 to 4 eV using a split-step Fourier method

Stefan E. Huber · Tobias Hell · Michael Probst ·
Alexander Ostermann

Abstract We report simulations of the elastic scattering of atomic hydrogen isotopes and helium beams from graphite (0001) surfaces in an energy range of 1–4 eV. To this aim, we numerically solve a time-dependent Schrödinger equation using a split-step Fourier method. The hydrogen- and helium-graphite potentials are derived from density functional theory calculations using a cluster model for the graphite surface. We observe that the elastic interaction of tritium and helium with graphite differs fundamentally. Whereas the wave packets in the helium beam are directed to the centers of the aromatic cycles constituting the hexagonal graphite lattice, they are directed toward the rings in case of the hydrogen beams. These observations emphasize the importance of swift chemical sputtering for the chemical erosion of graphite and provide a fundamental justification of the graphite peeling mechanism observed in molecular dynamics studies. Our investigations imply that wave packet studies, complementary to

S. E. Huber (✉) · M. Probst (✉)
Institute for Ion Physics and Applied Physics,
Innsbruck University, Technikerstrasse 25,
6020 Innsbruck, Austria
e-mail: s.huber@uibk.ac.at

M. Probst
e-mail: michael.probst@uibk.ac.at

T. Hell · A. Ostermann
Department of Mathematics, Innsbruck University,
Technikerstrasse 13, 6020 Innsbruck, Austria

classical atomistic molecular dynamics simulations open another angle to the microscopic view on the physics underlying the sputtering of graphite exposed to hot plasma.

1 Introduction

The divertor in the next step nuclear fusion device ITER is planned to be shielded against hot hydrogen plasma by plasma facing component (PFC) materials, in some scenarios carbon (target plates) and tungsten (upper divertor and dome) [1]. The choice of graphite (carbon) PFC materials has been made based on its high thermal conductivity, thermo-mechanical resistivity and the fact that it does not melt but only sublimates [2]. Major advantages of tungsten over carbon are its lower erosion rates and thus its longer lifetime. Its limitations are related to its thermal behavior under high off-normal heat loads like they might occur in large edge-localized modes [3, 4] and its stiffness. One of the major disadvantages of carbon is its high chemical reactivity with hydrogen (and/or isotopes) leading to erosion processes summarized under the expression chemical erosion [4–7]. This drastically limits the utilization of carbon-based materials due to safety concerns. In fact, it is planned to replace the carbon-based divertor plates with a full tungsten divertor at least in a later stage of the fusion experiment ITER [1]. The use of tungsten blocks below their ductile to brittle transition temperature could result in the total failure through the cracking of tungsten near cooling tubes [8]. Hence, carbon-based materials

should at least be kept in mind as an alternative for PFC materials in ITER and maybe even for reactors [8, 9]. Fusion experiments (for an overview, see for instance [4, 7, 10]) imply that carbon is significantly eroded by incoming hydrogen isotopes due to the sputtering of hydrocarbon molecules already at low plasma temperatures of about 1–10 eV. To complement these experiments as well as to provide an atomistic view on the chemical erosion of carbon-based materials, a variety of molecular dynamics (MD) simulations have been carried out [11–25]. These studies reveal several mechanisms that are important for the chemical erosion of carbon-based materials, of which three phenomena are of particular interest for the scope of this work. First, swift chemical sputtering has been proposed [11] as one of the main mechanisms leading to significant erosion at low impact energies. The sputtering proceeds as impinging hydrogen atoms penetrate the region between carbon–carbon bonds and, if their energies are in a certain range, subsequently break the bonds. Second, MD studies show that the addition of noble gases in small concentrations (less than 10 per cent) to the usual impact species hydrogen (isotopes) does not significantly alter the erosion yield [12]. Third, MD studies involving pure perfect hexagonal graphite reveal the interesting sputtering mechanism known as graphite peeling, that is as the graphite layers are bombarded they are peeled off one after the other [21].

In contrast to classical MD simulations, we treat elastic collisions of H/D/T (hydrogen/deuterium/tritium) and He with a graphite (0001) surface quantum-mechanically. For this purpose, we derive H/D/T-graphite and He-graphite potentials from quantum chemical cluster calculations and model the H/D/T and He nuclei as appropriate wave packets. The time-dependent Schrödinger equation is solved using a split-step Fourier method. This numerical approach has already proven to be beneficial in a variety of applications such as thermal energy atomic scattering or molecular beam scattering [26–29]. For example, in [30] elastic scattering of low energy He beams in the range of 10^{-2} eV with a rigid monolayer of xenon atoms was simulated in good agreement with experiment. More relevant in terms of nuclear fusion and the present work has been a wave packet study investigating the interaction of H and D with the basal plane of graphite in a range of impact energies of 0.1–0.9 eV [31]. In this work, we deal with beam energies in the range of 1–4 eV, which are substantially higher than the ones in these earlier applications, and therefore, they are believed to be of significant relevance for fusion research in the field of plasma-wall interaction. Though only elastic scattering is studied, our results are in agreement with the physics underlying the three previously described phenomena revealed by earlier MD studies and the mentioned wave packet study.

In Sect. 2, the numerical method is presented and the construction of the H/D/T-graphite and the He-graphite potentials is described. The results are discussed in Sect. 3, focusing especially on the significant difference between H/D/T and He scattering. Finally, in Sect. 4, a conclusion is given.

2 Methodology

We briefly describe the used numerical scheme, a Fourier split-step method, and the construction of the H/D/T- and He-graphite-potentials which is the input to the propagation method.

2.1 Numerical solution of the Schrödinger equation

In order to model the dynamics of elastic scattering from a quantum mechanical point of view, one has to solve the time-dependent Schrödinger equation, which—using mass-scaled coordinates—reads

$$i\partial_t \Psi(t,x) = -\Delta\Psi(t,x) + V(x)\Psi(t,x), \quad t > 0, x \in \mathbb{R}^3,$$
$$\Psi(0,x) = \Psi_0(x), \quad x \in \mathbb{R}^3, \tag{1}$$

where

$$\Psi : [0, \infty) \times \mathbb{R}^3 \to \mathbb{C}$$

denotes the wave function, V is the time-independent potential describing either the hydrogen- or the helium-graphite-potential and Ψ_0 is a given smooth initial function, in our case a Gaussian wave packet. Note that the original problem can be reformulated in such a way by choosing appropriate units. To numerically approach Eq. (1), we use a so-called splitting method. The basic idea of Lie splitting is to solve the potential part of Eq. (1)

$$i\partial_t \Psi_P(t,x) = V(x)\Psi_P(t,x) \tag{2}$$

and to use the respective solution as initial value for the kinetic part

$$i\partial_t \Psi_K(t,x) = -\Delta\Psi_K(t,x), \tag{3}$$

that is the free-particle Schrödinger equation. Given a temporal step size $h > 0$, set $t_n = nh$ and $\Psi_n^{(L)}$ denoting the numerical approximation obtained by the Lie splitting at time $t = nh$ and using the notion of groups, the resulting method reads as

$$\Psi_{n+1}^{(L)} = e^{-ihV}e^{ih\Delta}\Psi_n^{(L)},$$

where $\Psi_0^{(L)} = \Psi_0$. Under appropriate assumptions, the method is of order one, that is

$$\left\|\Psi_n^{(L)} - \Psi(nh, \cdot)\right\| \leq C_L h \quad \text{for } 0 \leq nh \leq T,$$

where $C_L > 0$ denotes a constant not depending on the step size h.

By symmetrizing the Lie splitting, one obtains the Strang splitting, cf. [32], given by

$$\Psi_{n+1}^{(S)} = e^{-i\frac{h}{2}V} e^{ih\Delta} e^{-i\frac{h}{2}V} \Psi_n^{(S)},$$

which is of order two, that is

$$\left\| \Psi_n^{(S)} - \Psi(nh, \cdot) \right\| \leq C_S h^2 \quad \text{for } 0 \leq nh \leq T$$

and some constant $C_S > 0$ independent of h.

By choosing an appropriately large domain, we can artificially impose periodic boundary conditions. As a result, Eq. (3) can be solved very efficiently by using a Fourier spectral method. Note that the other subproblem, Eq. (2), is simply solved by pointwise multiplication. The resulting numerical scheme is known as the split-step Fourier method, cf. [33].

Let $\Psi_n^{(K)}$ denote the numerical approximation at time $t = nh$ using $K \in \mathbb{N}$ grid points in each space dimension. As described in [33], a single time step is schematically computed as follows:

1. Pointwise multiplication: $\Psi_n^{(K)} := e^{-ihV/2} \Psi_n^{(K)}$
2. FFT: $\hat{\Psi}_n^{(K)} := \mathcal{F}_K \Psi_n^{(K)}$
3. Multiplication in Fourier space: $\hat{\Psi}_n^{(K)} := e^{ih\lambda} \hat{\Psi}_n^{(K)}$
4. IFFT: $\Psi_n^{(K)} := \mathcal{F}_K^{-1} \hat{\Psi}_n^{(K)}$
5. Pointwise multiplication: $\Psi_n^{(K)} := e^{-ihV/2} \Psi_n^{(K)}$

Here, λ refers to the respective eigenvalues associated with the Laplacian. Note that one can combine the computation in (1) and (5) if one does not need dense output. For further details, we refer to [34], where also structure-preserving properties of the split-step Fourier method as unitarity, symplecticity and time-reversibility are discussed.

We found this method to perform very well when applied to our problems and since our potential V is expressed in quintic smoothing splines, the applied split-step Fourier method is second-order convergent, that is if Ψ_n denotes the numerical approximation at $t = nh$, it holds that

$$\| \Psi_n - \Psi(nh) \| \leq C h^2 \quad \text{for } 0 \leq nh \leq T,$$

where the constant $C > 0$ does not depend on the step size h, see [33].

2.2 Construction of the potentials

The determination of energy barriers arising from the permeation of different atoms through models of a graphite (0001) surface is discussed in detail in Ref. [35], and we give only a short summary of the procedure to construct the H/D/T- and He-graphite potentials.

The permeation of an atom through a graphite (0001) surface is investigated by quantum chemical cluster calculations in which graphite is modeled by the PAH molecule coronene, see Fig. 1a. PAH molecules are adequate models for graphite surfaces [36] and graphite (0001) surfaces can be interpreted as infinitely sized PAH molecules. Recent studies [35] have shown that the size of coronene ($C_{24}H_{12}$) is already sufficient to estimate the potential energy within an accuracy of about 10 % at the barrier maximum. The three-dimensional H/D/T- and He-graphite potentials are constructed in the following way (Fig. 1). First, the distance between H or He and the plane of coronene is varied in 30 steps of 0.15 Å each. Thus, the atom is moved along the z-axis toward the molecular plane (the xy-plane) of the undisturbed PAH from 3 Å in front to 3 Å behind the PAH molecule (Fig. 1b). This process is performed for 21 distinct permeation sites in a quarter of the central aromatic cycle of coronene. At each position, the energy E(A-PAH, z) of the total system is calculated after relaxation of the PAH molecule, that is the adiabatic energy barrier $E(z)$ at the respective position z of the contaminant atom is then given as the difference $E(z) = E(\text{A-PAH}, z) - (E(A) + E(\text{PAH}))$, where $E(A)$ and E(PAH) denote the energies of the isolated atom and unperturbed coronene, respectively. Geometries and energies are obtained by density functional theory using the B3LYP functional [37] in conjunction with the small split-valence basis set 6-31G [38]. The appropriateness of this chemical model has been validated in [35]. By comparison with the results obtained with the more sophisticated ωB97XD functional [39] and larger basis sets, it has been found that shortcomings of the B3LYP functional, for example, the neglect of dispersion forces and the basis set truncation error cancel each other. Our model agrees with earlier theoretical [40–43] as well as experimental [44] studies on hydrogen adsorption on graphite: The potential minima are above the carbon atoms, that is at top sites about 1.4 Å from the undisturbed surface. Moreover, the change from sp^2 to sp^3 hybridization of the carbon atom where the hydrogen is adsorbed is well reproduced [35]. Quantitatively, however, the accordance is just reasonable. In [44] a desorption energy of 0.7 eV (the energy needed to release an adsorbed hydrogen from the surface) and an activation energy of 0.18 eV are given, whereas our constructed potential yields 0.43 and 0.4 eV, respectively. On average, the accuracy of our model is approximately 0.5 eV. All quantum chemical computations have been carried out with the Gaussian 09 software [45]. By using the symmetry and periodic continuation, the potentials obtained from coronene are extended to model the H/D/T- and He-graphite potential for an infinite surface, see Fig. 1c.
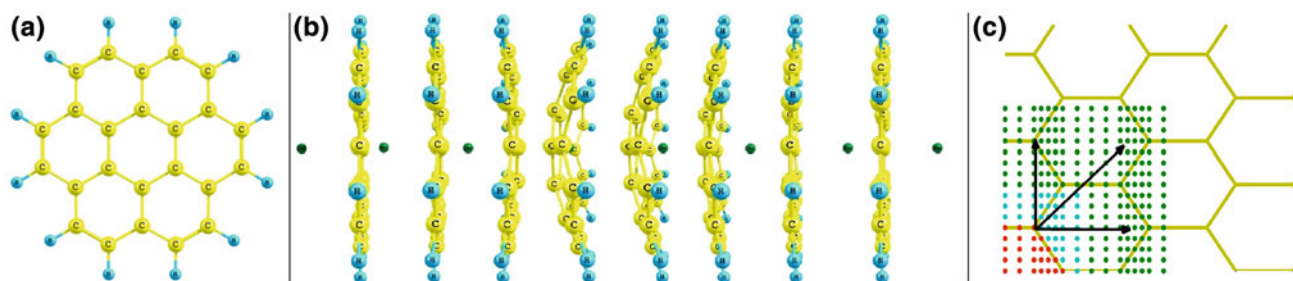
Fig. 1 a Coronene as a model for the graphite (0001) surface. **b** Permeation of He (*dark green sphere*) through coronene. The PAH is fully relaxed at each step. Geometries are calculated at the B3LYP/6-31G level of theory. **c** Schematics of the construction of the total surface potential by using symmetry. The energy barriers at the locations of the *blue dots* coincide with those calculated at 21 considered permeation sites marked by *red dots*) and periodic continuation (*arrows* and *green dots*) from energy barriers at the 21 permeation sites

The barrier heights, that is the energies that incoming classical particles need to have at least to penetrate the surface are about 4.5 eV in case of the H/D/T-graphite potential and about 10 eV in case of the He-graphite potential. Since we are only considering elastic reflection processes at energies below these limits, we believe that the approximations introduced by our adiabatic potentials are quite reasonable. Furthermore, it should be noted that the adiabatic approximation is well applicable for the range of projectile energies considered here since the highest energy of 4 eV is much below the HOMO–LUMO gap of coronene (about 9 eV [46]) used to model the graphite (0001) surface.

3 Results and discussion

We simulate the elastic scattering of H/D/T- and He beams, at a graphite (0001) surface with the wave packet method described in Sect. 2.1. The hydrogen isotopes are all chemically equivalent and therefore the potentials of H, D and T are the same. For convenience, we focus mainly on tritium for the following reasons. Since the mass of T is most similar to the one of helium, one can treat both elements using the same computational domain as well as spatial resolution. For lighter isotopes, the spatial width of the wave packet is larger and consequently the use of different computational domains becomes necessary to capture the essential features of the scattering process. In particular, the energy width of the beam $\Delta E = 0.2$. eV is inversely related to the spatial width via Heisenberg's uncertainty relation $\Delta k \Delta x \geq 1$ and $E = k^2/2\,m$, where k denotes the momentum of the beam and m the particle mass. However, the results for tritium scattering are qualitatively applicable for the lighter hydrogen isotopes too and treating H or D results in a scaling of the temporal and spatial dimensions proportional to a factor of $\sqrt{m}$ as will be shown below.

For T and He, except for the mass, the same input parameters are used in the simulations. The energies of the particles are 1, 2, 3 and 4 eV, the energy width has been chosen to be 0.2 eV and in all cases the angle of impact is 90°, corresponding to a perpendicular beam direction with respect to the surface. The size of the time steps are 200, 160, 140 and 120 a.u. for the four different beam energies, respectively, leading to 41 time steps in each simulation. The size of the computational domain is $12.5041 \times 14.5418 \times 18$ Å^3 in which $256 \times 256 \times 512$ grid points define the spatial resolution. The interaction potential is active only in the last third of the computational domain with respect to the z-direction, that is in the region 3 Å $< z <$ 9 Å, whereas the z-limits are between -9 Å and $+9$ Å. The interaction sites, that is the regions where the wave packet impinges at the surface, have been chosen to be the three high-symmetry sites of the graphite hexagonal lattice: The hollow site in the center of one aromatic cycle, the bridge site at the center of one C–C bond and the top site directly on the top of one of the carbon atoms. Thus, we have to simulate 24 scattering events in total (four energies and three sites for two nuclei).

One approach to assess the results is to inspect a cut through the computational domain parallel to the surface, the 'window' [29], and observe the wave packet as it moves through. The wave packet passes the window twice, that is when it approaches the interaction region and after it is reflected from the surface. The window is placed near the plane of the PAH at $z = 3$ Å.

We start with the discussion of the elastic reflective scattering of He. The He-graphite potential is purely repulsive and does not contain features like, for example, adsorption minima, whereas the H/D/T-graphite potentials do (see Sect. 2.2). Therefore, the physical interpretation is easier and can serve as a basis for the assessment of the more complex situation of H/D/T scattering.

3.1 Elastic scattering of He

In Fig. 2, the reflection patterns are depicted for the energy range of 1–4 eV at all three interaction sites (movies illuminating the dynamics of the reflection process are available as supplementary material in the online version of this article). He is deflected from the boundary of the aromatic cycles constituting the graphite lattice. For the impact at bridge site, this means that the initially spherical symmetric He wave packet (top row in Fig. 2) is symmetrically split into two parts by the carbon–carbon bond. At the hollow site, the interaction with the potential valley causes a focusing effect similar to light being reflected by a parabolic mirror. At the top site, the wave packet is symmetrically split into three parts pointing toward the centers of the adjacent hexagonal rings. In addition, for higher energies, the quantum nature of the scattering process becomes visible. Several maxima and minima form in the reflection pattern due to quantum interferences. At the highest energies of 3 and 4 eV, the interaction is strong enough that even far outlying parts of the initial wave packet are focused considerably in hexagonal rings adjacent to the central one, that is six, two and three side maxima are observable besides the strong central feature of the reflection pattern in the case of the hollow, the bridge and the top site, respectively.

The He-graphite potential is purely repulsive and most repulsive at the boundary of the aromatic cycles, that is at the position of the carbon atoms and at carbon–carbon bonds. This can nicely be seen in plots of the dwell time, that is the probability distribution inside a certain interaction region versus time. The interaction region has been chosen as the interval from 4.5 to 9 Å in the computational domain. The part of the system beyond the potential maximum is never reached at the considered energies, and thus, the actual interaction region ranges from 1.5 Å in front of the undisturbed surface to the turning point of the wave packet. For purely repulsive interaction, the dwell time exhibits a Gaussian shape since no partial reflection occurs on the way back of the wave packet due to the monotone shape of the potential. This is the case at all energies as can be seen in Fig. 3. In addition, no difference between the distinct impact sites is observed. The different maxima of the dwell time curves for the individual energies simply refer to the fact that at low energies not the whole wave packet is able to penetrate the interaction region, whereas at high energies parts of the wave packet have already left the interaction region as recently as other parts are not there yet. The shift of the dwell time curves to the left with increasing energy results simply from the fact that all wave packets have the same starting location at 6 Å in front of the undisturbed surface.

Additional information can be gathered from projecting the probability distribution in the momentum space into p_x and p_y. In Fig. 4, this is done for the initial wave packet

(away from the interaction zone) as well as for the final reflected beam (again away from the interaction zone) for the three different interaction sites for 1 eV. Initially, the wave packet corresponds to a peak at the origin of the xy-plane with a width corresponding to the energy width of 0.2 eV as discussed at the beginning of Sect. 3. After reflection at the bridge site, four main scattering channels symmetrically shifted to two opposite sites from the origin are observed, in accordance with the direct inspection of the reflection patterns. Analogously, for an impact at the top site, one observes a partial shift of the central peak to three main peaks separated from each other by an angle of 120° in accordance with the broad diffuse zone between the three main peaks in Fig. 2. For impact at the hollow site, the central peak becomes tighter and is symmetrically encircled by a ring-like distribution, again corresponding well to the same feature in Fig. 2. The projection of the probability distribution in momentum space has the advantage that, due to energy conservation, the angular dependence of the reflection pattern can be inspected simply by plotting Ewald's circle referring to the initial momentum of the wave packet. The part of the probability distribution located at the origin then refers to a reflection perpendicular to the surface, that is a reflection angle of 0°, whereas toward Ewald's circle the reflection angle increases and on it corresponds to a reflection angle of 90°, that is, parallel to the surface.

3.2 Elastic scattering of H/D/T

In Fig. 5, the reflection patterns are depicted for the energy range of 1–4 eV at all three interaction sites (movies illuminating the dynamics of the reflection process are available as supplementary material in the online version of this article). The results exhibit more complexity than in the case of He scattering and a stronger energy dependence. For 1 eV one recognizes similar features as in the case of He scattering. However, already at this low impact energy, some differences are apparent. At the bridge site, the wave packet appears to be deflected by the carbon–carbon bond but exhibits much more structure due to interference. The initially spherical wave packet is distributed into an ellipsoid-like shape around the region of the bond and is not separated into two distinct wave packets at the two sides of the bond. At the hollow site, the wave packet is not focused but rather smeared out, pointing toward the edges of the aromatic ring. Inspection of the reflection patterns at different locations of the window reveals that in case of the hollow site, the wave packet is mainly split into two parts which are deflected toward the carbon–carbon bonds on the left and right of the hexagon, see Fig. 5. For the top site, one observes a similar triangular shape of the reflection pattern as for He, but again, it is smeared out and does not

Fig. 2 Reflection patterns for He-graphite interaction in the range of 1–4 eV (from *top* to *bottom*) at the three interaction sites bridge (*left*), hollow (*center*) and top (*right*). As reference, the initial wave packets are shown in the *top row*
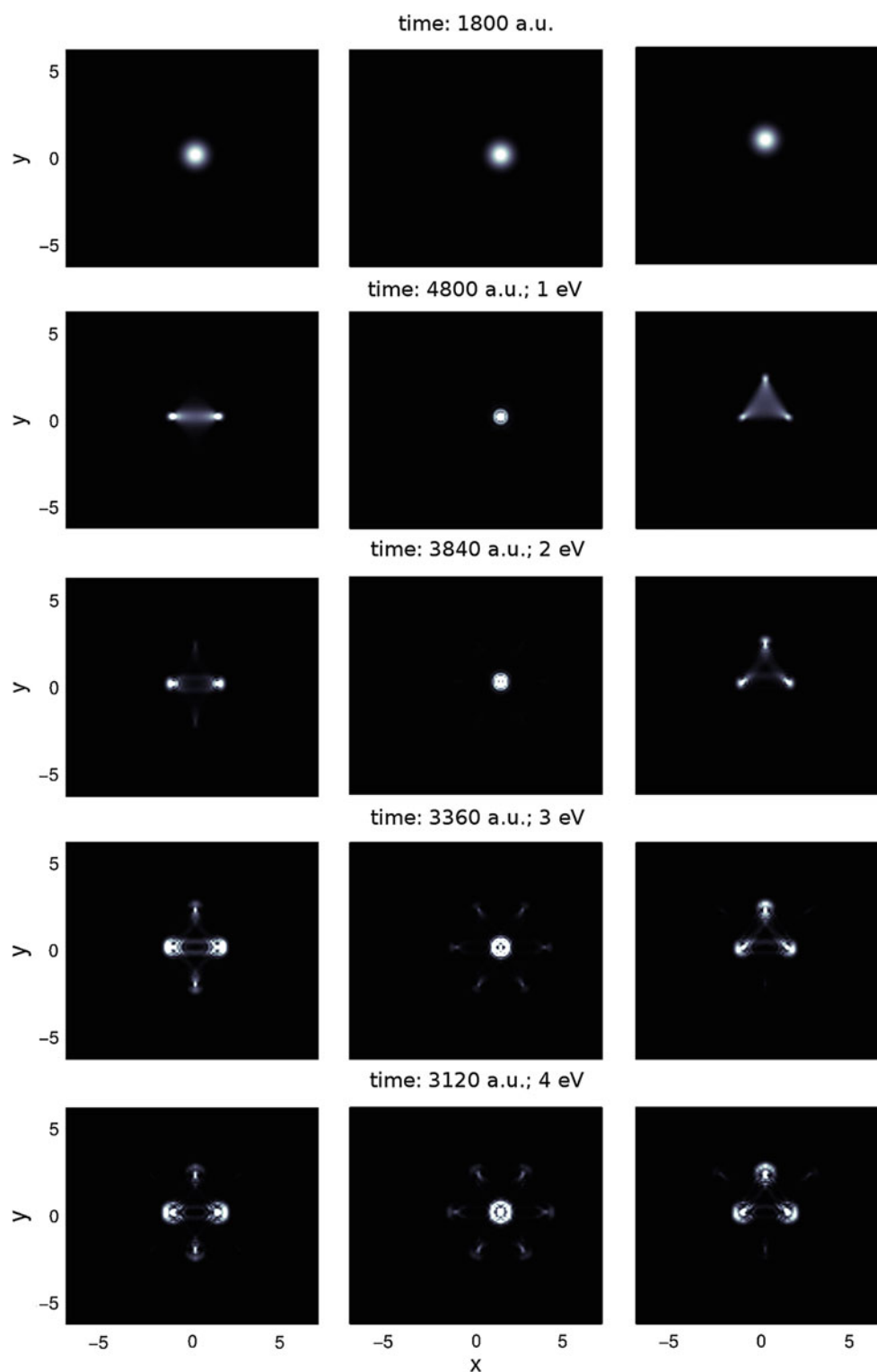
yield three well-separated reflection peaks but is distributed rather around the carbon atom. At 2–4 eV, the situation becomes clearer. It turns out that the T wave packet is attracted by the boundary of the hexagonal aromatic cycles that constitute the graphite surface. This can be seen best in the reflection patterns for the hollow site, where at the

starting energy of 2 eV a hexagonal structure in the reflection pattern becomes observable. The higher the energy, the more concise and compact the structure of the reflection pattern becomes. In addition, parts of the adjacent aromatic cycles can be observed in it. The same is true for the bridge and the top sites. Altogether, the T wave

190

packet is rather deflected away from the centers of the aromatic cycles and attracted to their boundary in the energy range of 2–4 eV.

The reason for the difference between He and H/D/T are the various local minima in the H/D/T-graphite potential. This is clearly seen in the plots of the dwell time for hydrogen, deuterium and tritium in Fig. 6. For an energy of 1 eV, the dwell time exhibits a long tail which is a consequence of partial reflection when the wave packet passes through the minima in front of the steep repulsive increase of the potential. Thus, the wave packet is first reflected by the potential wall and then by the edges of the minima, parts of it several times. Hence, the dwell time decreases significantly slower than in the case of He in agreement with an earlier wave packet study focusing on impact energies in a range of
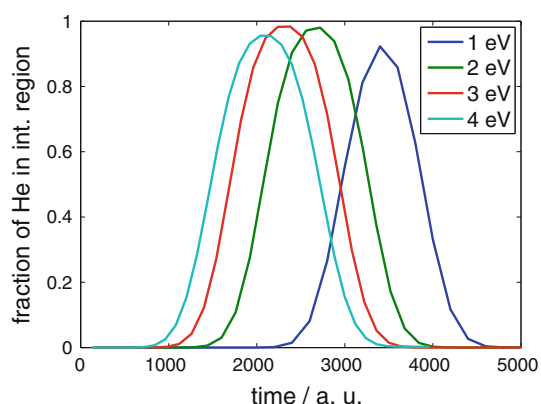
Fig. 3 Dwell time in the case of He reflection at graphite (0001) surfaces for different impact energies
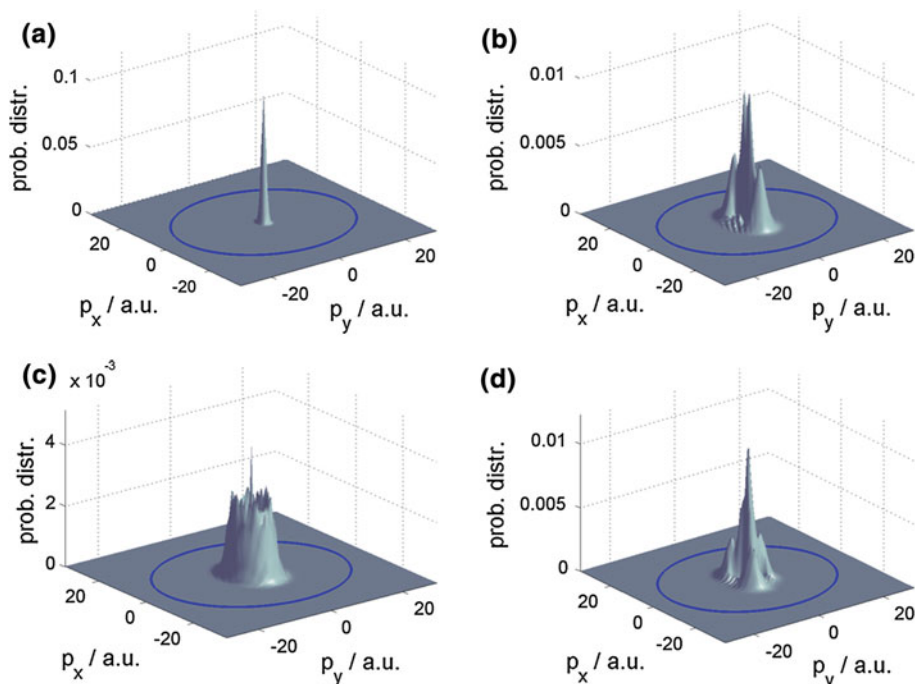
0.1–0.9 eV [31]. At higher energies, the dwell time becomes more Gaussian-shaped as the influence of these features of the H/D/T-graphite potential become less significant in agreement with physical intuition, because as energies become higher the importance of quantum dynamic effects diminishes. Nevertheless, the results at higher energies are in agreement also with a recent investigation of the sticking coefficients of H/D/T on graphite which have been shown to decrease with increasing impact energy [47]. These features as well as the scaling with the square root of the impinging particle's mass are well reproduced in Fig. 6. The dwell time at 2 eV approaches zero still slower as in the case of He reflection, but the difference is much smaller than at 1 eV. The dwell time curve for 3 eV is already Gaussian-shaped. The curve for the impact energy of 4 eV again exhibits a small tail. This could be due to the fact that the classical permeation energy at bridge site is 4.5 eV and thus is only slightly higher than the impact energy. This subsequently increases the possibility of encountering small potential walls when the wave packet is reflected into regions of higher energy.

Inspection of the projected momentum space probability distributions (Fig. 7) illuminates this. For an impact energy of 1 eV, the different scattering channels extend far to the boundary of Ewald's circle corresponding to the complex situation of multiple re-reflection at the several potential walls arising from the small-scale structures in the potential. When the impact energy is increased, the scattering channels are well located in the central region of Ewald's circle, that is the influence of the small scale structures vanishes. For an impact energy of 4 eV, one observes that the probability distribution becomes fuzzy for the reasons

Fig. 4 Projection of the probability distribution in momentum space for an impact energy of 1 eV. **a** Initial wave packet. **b** Final probability distribution after reflection at the bridge site. **c** Final distribution after reflection at the hollow site. **d** Final distribution after reflection at the *top* site. Ewald's circle is indicated as a *blue line*

Fig. 5 Reflection patterns for the T-graphite interaction in the range of 1–4 eV (from *top* to *bottom*) at the three interaction sites bridge (*left*), hollow (*center*) and top (*right*). As reference, the initial wave packets are shown in the *top row*
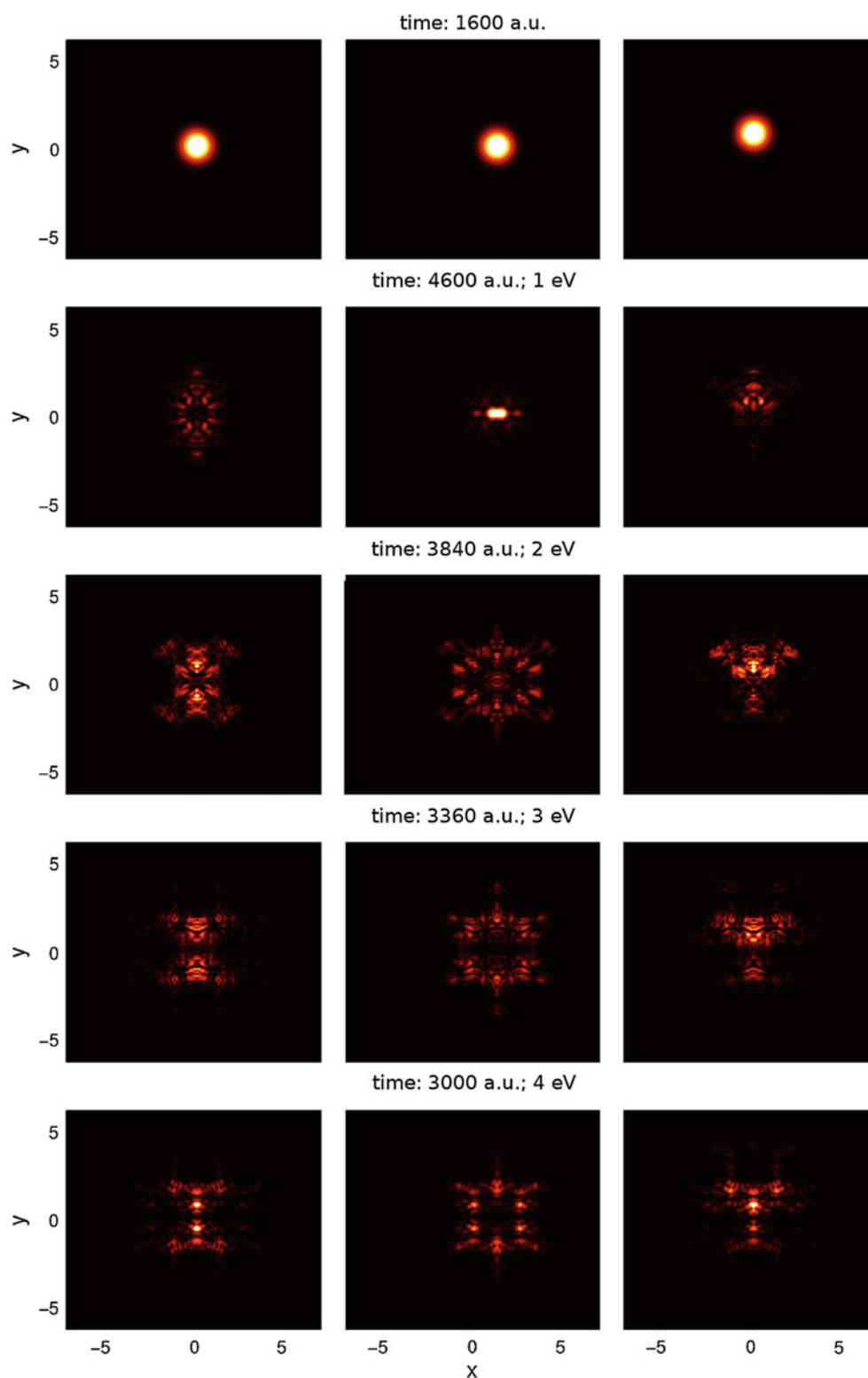
discussed above. Since the probability distribution is more centered and Ewald's circle is larger as the impact energy increases, the overall reflection angle becomes smaller.

This explains why the reflection patterns become more concise and compact as the energy is increased, see the discussion of Fig. 5.
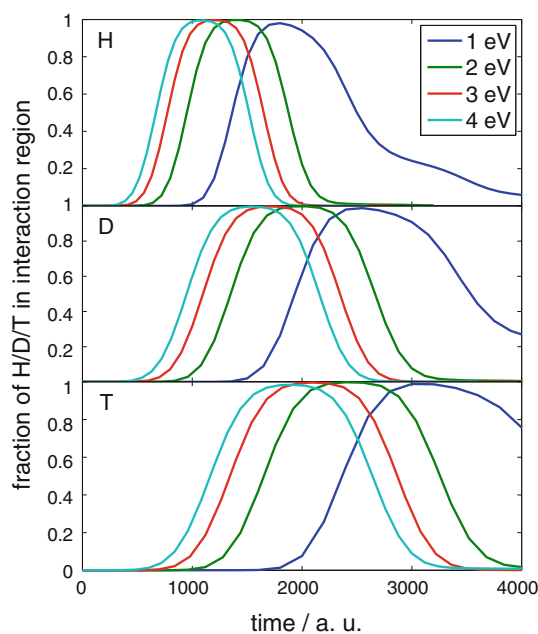
Fig. 6 Dwell times in the case of H, D or T scattering at graphite (0001) surfaces for different impact energies

3.3 Discussion

We observe a substantial difference between the reflective scattering of He and H/D/T at graphite (0001) surfaces. Whereas He is deflected from the boundary of the hexagonal aromatic cycles constituting the graphite surface toward their centers, the opposite is the case for H/D/T. Though this can be expected as hydrogen more likely reacts with carbon than He, our method provides a microscopic view on the atom-surface scattering process which can be related to several observations from earlier MD simulations [11, 12, 21]. First the role of swift chemical sputtering [11] of carbon-based materials by hydrogen isotopes is significantly emphasized by our findings. They show that hydrogen is very likely to interact with the surface predominantly in this relevant region of the aromatic cycle, especially in those low energy regimes in which chemical sputtering takes place. Second, the non-significance on the sputtering yield at low energies (<10 eV) of small amounts of noble gases mixed into the bombarding species [12] is easily explained. It turns out that He is likely to be deflected toward the centers of the

Fig. 7 Projection onto the *xy*-plane of the probability distribution in ▶ momentum space for an impact of *T* at the bridge site. **a** Initial wave packet. **b** Final probability distribution after reflection for an impact energy of 1 eV. **c** Final distribution after reflection for 2 eV. **d** Final distribution after reflection for 3 eV. **e** Final distribution after reflection for 4 eV. Ewald's circle is indicated as a *blue line*
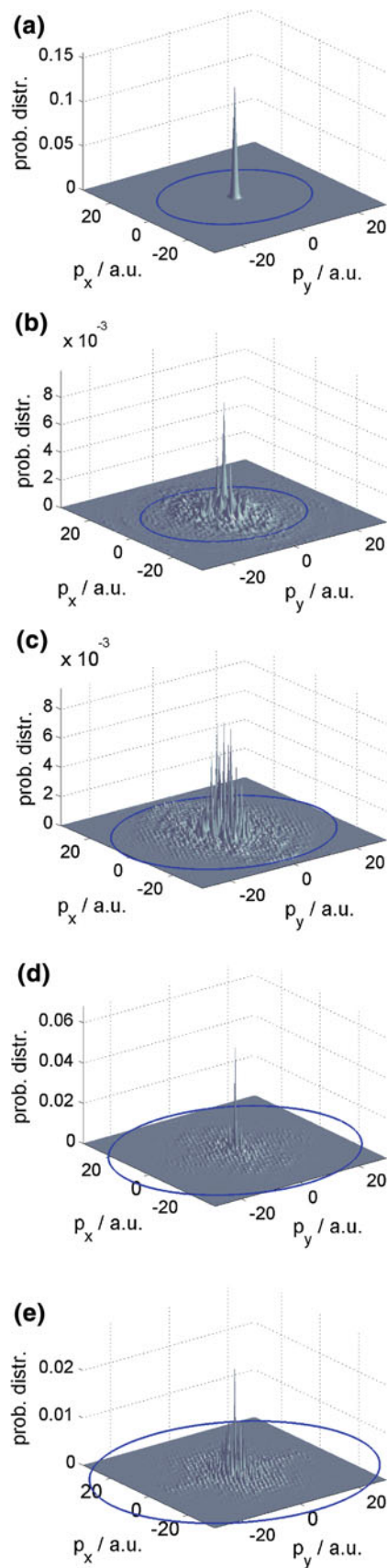
aromatic rings, which means that swift chemical sputtering is not applicable to this situation. In addition, the energy that a He atom might transfer is spread over a larger amount of carbon atoms and bonds such that at low energies He does not contribute much to the sputtering. Since He is prototypical for the chemical behavior of other noble gases, one can directly apply these results to these cases. Third, our results serve as a simple explanatory tool for the graphite peeling process [21], that is the observation that graphite layers are sputtered off one by one. Since hydrogen is deflected toward the boundary of the aromatic cycles, it is unlikely to go through the first layer at low energies without disturbing at least the surface layer via adsorption or bond breaking mechanisms. Thus, the second layer cannot effectively be sputtered until the degradation of the first layer has sufficiently advanced such that hydrogen can go through without being deflected toward the boundary.

4 Conclusion

We present a 3D time-dependent wave packet simulation to investigate elastic scattering of tritium and helium on graphite (0001) surfaces. We gain some insight into the physics underlying important mechanisms for chemical erosion of carbon-based materials exposed to hot plasma or gas. In particular, the importance of swift chemical sputtering [11] for the chemical erosion of carbon-based materials is underlined, and a first-principle explanation for observed graphite peeling during hydrogen bombardment [21] is given.

References

1. The ITER organization. see http://www.iter.org/
2. Miyahara A, Tanabe T (1988) Graphite as plasma facing material. J Nucl Mater 155–157:49–57
3. Roth J, Tsitrone E, Loarte A, Loarer T, Counsell G, Neu R, Philipps V, Brezinsek S, Lehnen M, Coad P, Grisolia C, Schmid K, Krieger K, Kallenbach A, Lipschultz B, Doerner R, Causey R, Alimov V, Shu W, Ogorodnikova O, Kirschner A, Federici G, Kukushkin A (2009) Recent analysis of key plasma wall interactions issues for ITER. J Nucl Mater 390–391:1–9
4. Federici G, Skinner CH, Brooks JN, Coad JP, Grisolia C, Haasz AA, Hassanein A, Philipps V, Pitcher CS, Roth J, Wampler WR, Whyte DG (2001) Plasma-material interactions in current tokamaks and their implications for next step fusion reactors. Nucl Fusion 41:1967–2137
5. Samm U (2008) Plasma-wall interaction in magnetically confined fusion plasmas. Trans Fusion Sci Technol 53:223–228
6. Philipps V (2006) Plasma-wall interaction, a key issue on the way to a steady state burning fusion device. Phys Scr T123:24–32
7. Lipschultz B, Bonnin X, Counsell G, Kallenbach A, Kukushkin A, Krieger K, Leonard A, Loarte A, Neu R, Pitts RA, Rognlien T, Roth J, Skinner CH, Terry JL, Tsitrone E, Whyte D, Zweben S, Asakura N, Coster D, Doerner R, Dux R, Federici G, Fenstermacher M, Fundamenski W, Ghendrih P, Herrmann A, Hu J, Krasheninnikov S, Kirnev G, Kreter A, Kurnaev V, LaBombard B, Lisgo S, Nakano T, Ohno N, Pacher HD, Paley J, Pan Y, Pautasso G, Philipps V, Rohde V, Rudakov D, Stangeby P, Takamura S, Tanabe T, Yang Y, Zhu S (2007) Plasma-surface interaction, scrape-off layer and divertor physics: implications for ITER. Nucl Fusion 47:1189–1205
8. Tanabe T (2010) Tritium issues in plasma wall interactions. AIP Conf Proc 1237:106–121
9. Tanabe T (2006) On the possibility of ITER starting with full carbon. Fusion Eng Des 81:139–147
10. Roth J, Tsitrone E, Loarte A (2008) Plasma-wall interaction: a complex combination of surface processes critical for thermonuclear fusion. J Phys Conf Ser 100:062003. doi:10.1088/1742-6596/100/6/062003
11. Nordlund K, Salonen E, Krasheninnikov S, Keinonen J (2006) Swift chemical sputtering of covalently bonded materials. Pure Appl Chem 78:1203–1211
12. Traeskelin P, Nordlund K, Keinonen J (2006) H, He, Ne, Ar-bombardment of amorphous hydrocarbon structures. J Nucl Mater 357:1–8
13. Traeskelin P, Saresoja O, Nordlund K (2008) Molecular dynamics simulations of C2, C2H, C2H2, C2H3, C2H4, C2H5, and C2H6 bombardment of diamond (111) surfaces. J Nucl Mater 375:270–274
14. Krstic PS, Reinhold CO, Stuart SJ (2007) Chemical sputtering from amorphous carbon under bombardment by deuterium atoms and molecules. New J Phys 9:209/201–225
15. Stuart SJ, Krstic PS, Embry TA, Reinhold CO (2007) Methane production by deuterium impact at carbon surfaces. Nucl Instrum Methods Phys Res B 255:202–207
16. Marian J, Zepeda-Rutz LA, Gilmer GH, Bringa EM, Rognlien T (2006) Simulations of carbon sputtering in amorphous hydrogenated samples. Phys Scr T124:65–69
17. Alman DA, Ruzic DN (2003) Molecular dynamics calculation of carbon/hydrocarbon reflection coefficients on a hydrogenated graphite surface. J Nucl Mater 313–316:182–186
18. Ito A, Nakamura H (2006) Molecular dynamics simulation of collisions between hydrogen and graphite. J Plasma Phys 72:805–808
19. Ito A, Nakamura H (2007) Molecular dynamics simulation of sputtering process of hydrogen and graphene sheets. Mol Simul 33:121–126
20. Ito A, Nakamura H (2008) Hydrogen isotope sputtering of graphite by molecular dynamics simulation. Thin Solid Films 516:6553–6559
21. Ito A, Nakamura H (2008) Molecular dynamics simulation of bombardment of hydrogen atoms on graphite surface. Commun Comput Phys 4:592–610

22. Ito A, Wang Y, Irle S, Morokuma K, Nakamura H (2009) Molecular dynamics simulation of hydrogen atom sputtering on the surface of graphite with defect and edge. J Nucl Mater 390–391:183–187

23. Ito A, Ohya K, Inai K, Nakamura H (2010) Dependency of tritium retention in graphite on temperature control of molecular dynamics. Contrib Plasma Phys 50:464–469

24. Ohya K, Mohara N, Inai K, Ito A, Nakamura H, Kirschner A, Borodin D (2010) Molecular dynamics and dynamic Monte Carlo studies of mixed materials and their impact on plasma wall interactions. Fusion Eng Des 85:1167–1172

25. Saito S, Ito AM, Takayama A, Kenmotsu T, Nakamura H (2011) Hybrid simulation between molecular dynamics and binary collision approximation codes for hydrogen injection into carbon materials. J Nucl Mater 415:S208–S211

26. Varga G (1999) Investigation of thermal energy atomic scattering from solid surfaces using the 3D time-dependent Schroedinger equation. Surf Sci 441:472–478

27. Balazs E, Varga G, Fuestoess L (2001) Comparison of 3D classical and quantum mechanical He scattering on Rh(311). Surf Sci 482–485:1145–1151

28. Varga G (2001) Computer simulated thermal energy atomic scattering on solid surfaces. Surf Sci 482–485:1152–1158

29. Varga G (2002) Computer simulation by the quantum mechanical time-dependent wavepacket method, especially for atom/molecule-solid-surface interaction. J Phys Condens Matter 14:6081–6107

30. Tribe L (2006) Wave packet calculations for helium scattering by a xenon monolayer. Chem Phys 327:468–473

31. Sha X, Jackson B, Lemoine D, Lepetit B (2005) Quantum studies of H atom trapping on a graphite surface. J Chem Phys 122:014709/014701–014708

32. Strang G (1968) On the construction and comparison of difference schemes. SIAM J Numer Anal 5:506–517

33. Jahnke T, Lubich C (2000) Error bounds for exponential operator splittings. BIT 40:735–744

34. Lubich C (2008) From quantum to classical molecular dynamics: reduced models and numerical analysis. European Mathematical Society (EMS), Zuerich

35. Huber SE, Mauracher A, Probst M. Permeation of low-Z atoms through carbon sheets: density functional theory study on energy barriers and deformation effects. arXiv:1201.4014v2 [physics.chem-ph]

36. Jug K, Bredow T (2004) Models for the treatment of crystalline solids and surfaces. J Comput Chem 25:1551–1567

37. Becke A (1993) Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys 98:5648–5652

38. Ditchfield R, Hehre WJ, Pople JA (1971) Self-consistent molecular orbital methods. 9. Extended Gaussian-type basis for molecular-orbital studies of organic molecules. J Chem Phys 54:724–728

39. Chai J, Head-Gordon M (2008) Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. Phys Chem Chem Phys 10:6615–6620

40. Sha X, Jackson B (2002) First-principles study of the structural and energetic properties of H atoms on a graphite (0001) surface. Surf Sci 496:318–330

41. Ferro Y, Brosser C, Allouche A (2004) Quantum study of hydrogen interaction with plasma-facing graphite and boron doped graphite surfaces. Phys Scr T108:76–79

42. Ferro Y, Jelea A, Marinelli F, Brosset C, Allouche A (2005) Density functional theory and molecular dynamics studies of hydrogen interaction with plasma-facing graphite surfaces and the impact of boron doping. J Nucl Mater 337–339:897–901

43. Jeloaica L, Sidis V (1999) DFT investigation of the adsorption of atomic hydrogen on a cluster-model graphite surface. Chem Phys Lett 300:157–162

44. Zecho T, Guettler A, Sha X, Jackson B, Kueppers J (2002) Adsorption of hydrogen and deuterium atoms on the (0001) graphite surface. J Chem Phys 117:8486–8492

45. 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.1. Gaussian, Inc., Wallingford, CT

46. Kayanuma M, Ikeshoji T, Ogawa H (2011) Theoretical study of hydrogen chemisorption to nitrogen-substituted graphene-like compounds. Bull Chem Soc Jpn 84:52–57

47. Ehemann RC, Krstic PS, Dadras J, Kent PRC, Jakowski J (2012) Detection of hydrogen using graphene. Nanoscale Res Lett 7:198/191–114

REGULAR ARTICLE

First-principles study of structure and stability in Si–C–O-based materials

A. Morales-García · M. Marqués · J. M. Menéndez · D. Santamaría-Pérez · V. G. Baonza · J. M. Recio

Abstract The decomposition of silicon–carbon–oxygen (SiCO)-based materials into their binary oxides is studied at different pressure conditions by first-principles total energy calculations. We evaluate how the influence of pressure modifies the enthalpy of this reactive process as the carbon concentration increases for a number of compounds with variable stoichiometries within the general chemical formula $Si_xC_yO_{24}$ ($x = 4$–11, $y = 1$–8, $x + y = 12$). At low carbon content, pressure destabilizes all SiCO structures examined, whereas at high carbon content, pressure tends to reduce the exothermic character of the decomposition reaction. After the evaluation of the equation of state parameters for these crystals, we found a similar trend for the bulk modulus with the highest values found for the richest carbon compounds. According to our results, we propose SiC_2O_6 as the most plausible stoichiometry.

Keywords First-principles · SiCO-based materials · High pressure · Decomposition enthalpy

A. Morales-García (✉) · D. Santamaría-Pérez · V. G. Baonza
MALTA Consolider Team, Departamento de Química Física I, Universidad Complutense de Madrid, 28040 Madrid, Spain
e-mail: amoralesg@quim.ucm.es

M. Marqués · J. M. Menéndez · J. M. Recio
MALTA Consolider Team, Departamento de Química Física y Analítica, Universidad de Oviedo, 33006 Oviedo, Spain

1 Introduction

The analogies and differences between the stability and reactivity of CO_2 and SiO_2 polymorphs have been the subject of several investigations [1–3]. Although both compounds belong to the group IV oxides, they are remarkably different under ambient conditions: CO_2 is a molecular gas, and SiO_2 is a crystalline solid. However, such differences progressively vanish at extreme conditions. Thus, although isolated CO_2 molecules are characterized by double bonds, the bonding pattern changes dramatically in the solid state at extreme conditions of pressure (p) and temperature (T). Non-molecular CO_2 crystalline phases in close resemblance to high-pressure SiO_2 polymorphs have been discovered above 30 GPa [4–7].

Such structural similarities suggest the possible existence of exotic mixed oxides based on silicon and carbon, which will be referred to as SiCO compounds. Only few theoretical efforts have been invested to face this attractive, yet difficult challenge. Considering α-quartz (SiO_2) as the reference structure, da Silva et al. [8] carried out a molecular dynamics study with the aim of proposing crystalline structures for SiCO-based compounds. A hypothetical α-quartz silicon oxycarbide ($Si_{1-x}C_xO_2$) was then generated by replacing Si atoms with C ones. This strategy provided two potential and competing crystalline phases, though an ab initio study of the (meta)stability of these compounds was not undertaken. Other density functional theory calculations were performed by Aravindh et al. [9] by assuming that SiO_2, CO_2, and their alloys took a β-cristobalite-like structure. In that work, the authors found that, despite mixed SiO_2–CO_2 alloys do not appear to be thermodynamically stable at ambient pressure, it might be possible to create them in metastable forms. They observed some tendency for C atoms in polymeric phases

to prefer threefold coordination rather than fourfold coordination.

Recently, Santoro et al. [7] succeeded in the synthesis of a silicon carbonate phase, which was obtained by reacting silicalite and fluid CO_2 in a diamond anvil cell at 18–26 GPa and 600–980 K. Interestingly, their spectroscopic results were in agreement with the observations of Aravindh et al. predictions. These experiments definitively reveal a unique oxide chemistry at extreme conditions, which can be summarized in the chemical equation $xSiO_2 + yCO_2 \rightarrow Si_xC_yO_{2(x+y)}$, thus opening routes for the synthesis of a novel class of chemical compounds of interest in different scientific areas, from geochemistry to materials science.

In this article, we study the stability of SiCO-based materials under pressure, with specific emphasis on the decomposition reaction of these hypothetical compounds into their simple oxides: SiO_2 and CO_2. Since no clues about the possible structure of potential SiCO compounds are available, we choose a monoclinic (space group $C2/c$, $Z = 4$) reference structure for our first-principles electronic structure calculations. This choice is justified by the expected analogy with the structure of UB_2O_6 [10] using well-established crystal-chemistry concepts based on the Zintl-Klemm concept applied to oxides [11]. This reference structure has been selected after a careful and systematic search among simple lattices of ternary compounds compatible with the carbonate-like environment proposed for C in the new synthesized SiCO compound [7]. Several U/Si and B/(C,Si) substitutions have been considered, as detailed in the following section.

This crystal modeling allows us to understand the different role played by Si and C in the stabilization of SiCO-based materials. The carbon content, along with the influence of hydrostatic pressure in the decomposition enthalpy, are the two variables considered in our study. After presenting the crystallographic models and the computational details, we evaluate and discuss the calculated decomposition enthalpies for several compounds within the stoichiometries involved in the general chemical formula $Si_xC_yO_{24}$ ($x = 4$–11, $y = 1$–8, $x + y = 12$). Our analysis is completed with the calculation of the static equations of state (EOS) for all the optimized structures. The conclusions are summarized at the end of the paper.

1.1 Modeling stoichiometries

The cell parameters and atomic positions of the reference UB_2O_6 structure are listed in Table 1 [10]. Our calculations are performed making use of the conventional monoclinic cell containing 4 formula units. With the aim of extending the number of potential stoichiometries and atomic environments for C and Si in this monoclinic lattice, it is

pertinent to substitute the eight B sites for C and Si atoms. Boron atoms are in threefold coordination with oxygen (BO_3), and the uranium atoms are in sixfold coordination (UO_6) (see Fig. 1).

Since the experiments of Santoro et al. [7] suggested the existence of carbonate units (CO_3), the reference structure should be rewritten as $U_4[B_nB'_{8-n}]O_{24}$ ($n = 1$–8), where we have distinguished two types of boron atoms: B and B'. The SiCO configurations are built just substituting U and B' for Si and B for C. Notice that n stands for the number of carbon atoms in the chemical formula above. Therefore, up to eight different stoichiometries involving different content of threefold coordinated carbon are to be considered (see Table 2). These structures can be generalized in the following carbon silicate or silicon carbonate chemical formula: $Si_xC_yO_{24}$, $x = 4$–11, $y = 1$–8, with $x + y = 12$. In addition, for some of these stoichiometries, more than one isomeric structure is compatible depending on the specific B and B' atoms chosen to be replaced by C and Si, respectively. This results in C–D, E–F and G–H isomers for $Si_4[C_3Si_5]O_{24}$, $Si_4[C_4Si_4]O_{24}$ and $Si_4[C_5Si_3]O_{24}$, respectively. Although the stoichiometries are the same for each pair of isomers, the environments are different, leading to energetically non-equivalent structures with the same carbon content. As a result, the final number of different SiCO-based compounds examined is eleven, covering a carbon content (defined as the ratio $n/8$) ranging between 0.125 and 1.

1.2 Calculation details

We perform first-principles enthalpy calculations within the framework of the density functional theory (DFT) with a plane-wave pseudopotential approach, as implemented in the Vienna ab initio simulation package (VASP) [12]. We use the projector augmented wave (PAW) all-electron description of the electron–ion–core interaction [13] and the Perdew-Burke-Ernzerhof generalized gradient exchange–correlation functional (PBE) [14]. Brillouin zone integrals were approximated using the method of Monkhorst and

Table 1 Lattice parameters and internal coordinates of the UB_2O_6 structure

Atom type	Wyckoff positions	x	y	z
U	4e	0	0.2407	0.25
B	8f	0.3013	0.1484	0.2509
O	8f	0.1026	0.2433	0.7562
O	8f	0.2983	0.1830	0.7614
O	8f	0.0893	0.2439	0.4487

Space group $C2/c$, $Z = 4$. $a = 12.504$ Å, $b = 4.183$ Å, $c = 10.453$ Å, and $\beta = 122.18°$
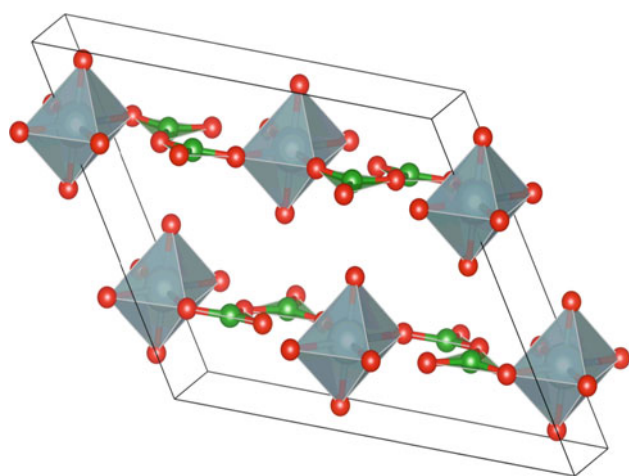
Fig. 1 Reference structure, UB_2O_6. *Gray*, *green* and *red* spheres represent uranium, boron and oxygen atoms, respectively. Boron (Uranium) atoms are tri(hexa)-coordinated to oxygen atoms

Table 2 Chemical formula and number and type of symmetry-inequivalent SiCO configurations for each carbon content

Carbon atoms	Structure	Carbon content	Structures (Fig. 3 nomenclature)
1	$Si_4[CSi_7]O_{24}$	0.125	1 (A)
2	$Si_4[C_2Si_6]O_{24}$	0.250	1 (B)
3	$Si_4[C_3Si_5]O_{24}$	0.375	2 (C–D)
4	$Si_4[C_4Si_4]O_{24}$	0.500	2 (E–F)
5	$Si_4[C_5Si_3]O_{24}$	0.625	2 (G–H)
6	$Si_4[C_6Si_2]O_{24}$	0.750	1 (I)
7	$Si_4[C_7Si]O_{24}$	0.875	1 (J)
8	$Si_4[C_8]O_{24}$	1.000	1 (K)

Carbon content = Carbon atoms/8

Pack [15], and the energies were converged with respect to the k-points density and the plane-wave cutoff (600 eV).

Guided by the experimental conditions under which the silicon carbonate phase was stabilized [7] and the known overestimation of pressure by the PBE functional, we restricted our calculations to a hydrostatic pressure window between 22.5 and 30 GPa. It is expected that temperature plays a crucial role in the kinetics (energy barrier) but not in the thermodynamics of the decomposition reaction. Therefore, only static (zero temperature and zero point vibrational contributions neglected) results will be presented here.

Due to the consideration of static conditions, enthalpy, $H = E + pV$, is the most suitable thermodynamic potential in our study, where E is the energy and V the volume, and all the quantities are considered per formula unit. We perform enthalpy calculations for the different SiCO configurations as well as for SiO_2 stishovite and the so-called phase III of solid CO_2, since they are the thermodynamically stable phases in the proposed pressure range. In
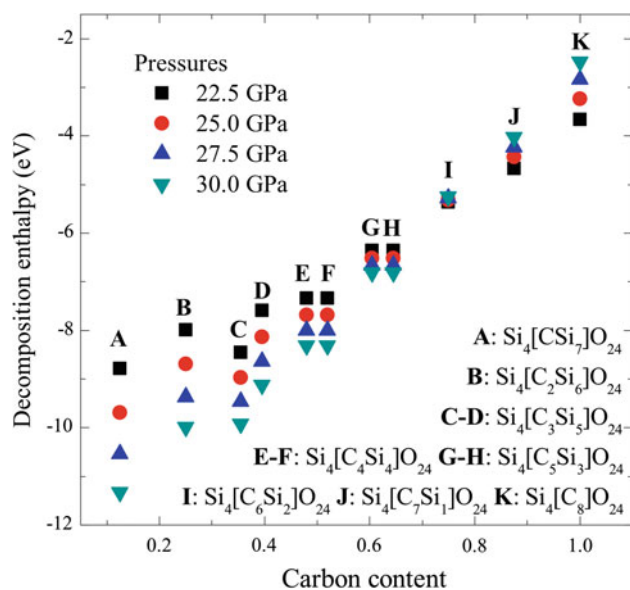
Fig. 2 Decomposition enthalpy versus carbon content for all the SiCO configurations at selected pressures. For sake of clarity, although the C and D structures have the same carbon content, the corresponding enthalpy values are slightly shifted to the *left/right*. The same applies to the E–F and G–H structures

particular, $4\times8\times4$, $6\times6\times8$ and $6\times6\times6$ k-meshes were employed for the SiCO structures, SiO_2-stishovite and CO_2-III, respectively. Full structural relaxations of both lattice parameters and atomic coordinates at each pressure were performed via a conjugate-gradient minimization of the enthalpy using the Hellmann–Feynman forces on the atoms and stresses on the unit cell. The geometry relaxation was considered to be complete when the total force on atoms was less than 1 meV/Å.

Additionally, the (p,V) calculated points have been described using the Murnaghan equation of state [16]:

$$V(p) = V_0\left(1 + B_0'\frac{p}{B_0}\right)^{-\frac{1}{B_0'}} \tag{1}$$

where V_0 is the zero-pressure volume, and B_0 and B_0' are the bulk modulus and its pressure derivative, respectively, both evaluated at zero pressure. The results are summarized in Table 4. The pressure derivative of the bulk modulus has been fixed to 3.5 and 4, which are typical values observed in oxides [17]. This option minimizes the relative uncertainty in the numerical determination of B_0 since large covariance terms between B_0 and B_0' are usually present in the fitting procedure.

2 Results and discussion

In order to study the stability of the optimized structures summarized in Table 2, the decomposition enthalpy into

Table 3 Unit formula volumes of stishovite (SiO_2), phase III (CO_2), $Si_4[C_4 Si_4]O_{24} \equiv Si_2CO_6$, $Si_4[C_6 Si_2]O_{24} \equiv SiCO_4$ and $Si_4[C_8]O_{24} \equiv SiC_2O_6$, at selected pressures

p (GPa)	SiO_2 ($Z = 2$)	CO_2 ($Z = 4$)	$Si_4[C_4 Si_4]O_{24}$ (E–F)		$Si_4[C_6 Si_2]O_{24}$ (I)		$Si_4[C_8]O24$ (K)	
	V (Å^3/Z)	V (Å^3/Z)	V (Å^3)	ΔV (%)	V (Å^3)	ΔV (%)	V (Å^3)	ΔV (%)
22.5	22.450	26.875	309.57	7.8	293.20	-0.93	277.35	−9.0
25.0	22.305	26.265	304.26	7.3	288.65	-0.95	273.51	−8.6
27.5	22.160	25.720	299.82	7.0	284.86	-0.84	270.36	−8.2
30.0	22.025	25.227	295.66	6.7	281.29	-0.78	267.37	−7.8

The volume changes (ΔV (%)) involved in the formation of the SiCO compounds (inverse reaction to the decomposition process) are also indicated
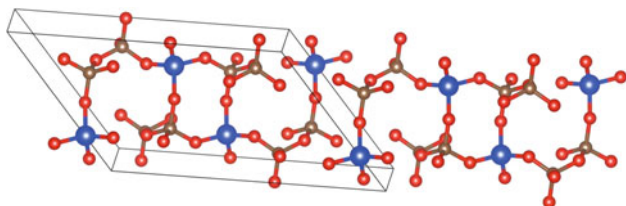
Fig. 3 SiC_2O_6 monoclinic structure. Lattice parameter: $a = 12.651$ Å, $b = 4.133$ Å, $c = 8.011$ Å, $\beta = 129.9°$. Wyckoff positions: Si(4e) (0, 0.9808, 0.75), C(8f) (0.3270, 0.2274, 0.3145) and O(8f) × 3 (0.1219, 0.1970, 0.7886) (0.2971, 0.1177, 0.7741) (0.0738, 0.2465, 0.4634). *Brown*, *red* and *blue* spheres represent carbon, oxygen and silicon atoms, respectively

Table 4 Zero-pressure bulk modulus (B_0) and first pressure derivative of the bulk modulus (B_0') for all the structures

Carbon atoms	Structures (Fig. 3 nomenclature)	B_0 (GPa)	B_0'
1	A	44.9–31.9	3.5–4
2	B	52.7–39.7	
3	C–D	65.3–52.3	
4	E–F	71.9–58.9	
5	G–H	80.1–67.1	
6	I	89.7–76.7	
7	J	101.4–88.4	
8	K	113.7–100.6	

B_0' is fixed at 3.5 and 4. The range of B_0 comes from different parameters and uncertainties in the fitting procedure

their binary oxides has been calculated. We consider the general chemical reaction $Si_xC_yO_{2(x+y)} \rightarrow xSiO_2 + yCO_2$ and evaluate $\Delta H_r = xH_{SiO_2} + yH_{CO_2} - H_{Si_xC_yO_{2(x+y)}}$ at four different pressures: 22.5, 25, 27.5 and 30 GPa. The decomposition enthalpy of the different SiCO structures as a function of the carbon content is plotted in Fig. 2. Specific pressure–volume data for SiO_2 stishovite, CO_2-III and SiCO structures with Si_2CO_6, $SiCO_4$ and SiC_2O_6 stoichiometries, along with the associated volume changes involved in their formation are also collected in Table 3. Analysis of these results allows us to draw some

conclusions regarding the role played by pressure and carbon content in the chemistry of SiCO-based materials.

The first general remark to emphasize is the exothermic character ($\Delta H_r < 0$) of the decomposition process, regardless the stoichiometry of the SiCO compound and the pressure involved in the reaction. As a result, these structures are not thermodynamically stable, releasing between 100 and 250 kcal/mol after decomposition into SiO_2 stishovite and CO_2–III. This fact agrees with the difficulty in finding crystalline forms containing simultaneously C, Si and O, and reinforces the well-known fact that CO_2 presents difficulties to become part of a solid solution [18].

It is now interesting to explore the combined influence of pressure and carbon content on ΔH_r. Bringing together these two factors in our simulations allows us to explore the existence of potential SiCO-based materials, since at the Earth depths corresponding to the pressure window of our simulations, traces of carbon have been found, and Si and O are common elements of many minerals [19, 20]. The results are summarized in Fig. 2.

It is observed that in those structures with carbon content below 0.75, a pressure increase favors the decomposition process, but the effect is reversed as the carbon content increases. Thus, it is predicted that pressure has negligible influence on ΔH_r in the structure with a carbon content of 0.75 (I configuration: $Si_4[C_6Si_2]O_{24} \equiv SiCO_4$), whereas for the J and K stoichiometries ($Si_4[C_7Si]O_{24}$ and $Si_4[C_8]O_{24} \equiv SiC_2O_6$), where the carbon content is higher, the role of pressure makes ΔH_r to increase around 35 kcal/mol from 22.5 to 30 GPa. Globally, it is interesting to condense the combined action of the two factors in these general rules: (1) the effect of increasing pressure changes from favoring to disfavoring the decomposition reaction as the carbon content increases, and (2) the decomposition reaction is less exothermic ($|\Delta H_r|$ decreases) as the carbon content increases for a given pressure. As a result, we found that SiC_2O_6 is the most plausible stoichiometry with the lowest value for $|\Delta H_r|$. Figure 3 shows this structure with the unit cell description given in the caption.

These results can be explained taking into account that pressure favors processes encompassing volume reductions [21], and these can be achieved by increasing the carbon content, as illustrated in Table 3. For each of the pressures examined, there is a continuous tendency to reduce the total volume involved in the reaction of formation of the SiCO compound (ΔV) as the carbon content increases. Thus, the highest volume reduction is found for the formation of SiC_2O_6.

The reason why is not difficult to understand has to do with the compressibilities of the three compounds involved in the reaction. In Table 4, we collect B_0 values corresponding to all the SiCO stoichiometries examined in this work. For CO_2-III and SiO_2 stishovite, our computed values are 7.6 GPa and 260 GPa, respectively. Since CO_2-III is the most compressible reactant, a high y stoichiometric coefficient for CO_2-III favors a great reduction of volume in the formation of a given SiCO compound: $\Delta V = \frac{V_{Si_xC_yO_{(2x+2y)}} - (xV_{SiO_2} + yV_{CO_2})}{xV_{SiO_2} + yV_{CO_2}} \times 100$. Besides, those compounds with higher carbon content are the ones with higher B_0 values (see Table 4), thus also favoring the reduction of volume in the reaction. Both factors make SiC_2O_6 to be the compound with highest volume reduction.

For a given stoichiometry, pressure variations from 22.5 to 30 GPa play a lower role on ΔV values. If we look, for example, at the sequence of ΔV in SiC_2O_6, it is found a slightly lower reduction of volume at the highest pressure compared with the value at 22.5 GPa. This fact does not contradict the result shown in Fig. 2. In energetic terms, we should consider the product $p\Delta V$, and this is greater in absolute value as the pressure increases making the decomposition reaction to be less favorable, in agreement with our discussion above.

3 Conclusions

In the search for relative stability of in SiCO-based compounds at high pressure, we have carried out a computational study of the chemical decomposition of eleven structures with varying stoichiometries into the binary oxides SiO_2 stishovite and CO_2-III. We have evaluated how the influence of pressure and carbon content modify the enthalpy of the reactive process: $Si_xC_yO_{(2x+2y)} \rightarrow xSiO_2 + yCO_2$. Although we found $\Delta H_r < 0$ for all SiCO structures analyzed, at high carbon content pressure tends to reduce their exothermic character. According to the results of our investigation, SiC_2O_6 is the most plausible stoichiometry. This agrees with the fact that this SiCO structure with the highest carbon content is less compressible than the corresponding decomposition products, due to the great concentration of the molecular CO_2-III phase. This behavior suggests new synthetic routes of this family of materials.

References

1. Iota V, Yoo CS, Klepeis JH, Jenei Z, Evans W, Cynn H (2006) Nat Mater 6:34–38
2. Serra S, Cavazzoni C, Chiarotti GL, Scandolo S, Tosatti E (1999) Science 284:788–790
3. Santoro M, Gorelli F, Bini R, Haines J, Cambon O, Levelut C, Montoya MA, Scandolo S (2012) Proc Natl Acad Sci 109:5176–5179
4. Iota V, Yoo CS, Cynn H (1999) Science 283:1510–1513
5. Yoo CS, Cynn H, Gygi F, Galli G, Iota V, Nicol M, Carlson S, Hausermann D, Mailhiot C (1999) Phys Rev Lett 83:5527–5530
6. Dong J, Tomfohr JK, Sankey OF (2000) Phys Rev B 61:5967–5971
7. Santoro M, Gorelli F, Haines J, Cambon O, Levelut C, Garbarino G (2011) Proc Natl Acad Sci 108:7689–7692
8. da Silva CRS, Justo JF, Pereyra I, Assali LVC (2005) Diam Relat Mater 14:1142–1145
9. Aravindh A, Arkundato A, Barman S, Baroni S, Bhargava BL, Chandrakumar KRS, Chen W, Cherian R, Dal Corso A, Datta S, de Gironcoli S, Dhayal SS, Dixit AK, Dutta S, D'yachkov S, Floare CG, Ganguli N, Ganguly S, Gebauer R, Ghosh S, Giannozzi P, Govind, Hatt AJ, Hembram KPSS, Imam M, Jayalakshmi V, Jayanthi CS, Kelkar T, Kumar A, Lee JH, Lee MS, Lonappan D, Mahadevan P, Mallajosyula SS, Marathe M, Marzini N, Melot B, Miller N, Morrone J, Nanavati S, Nanayakkara A, Nandi PK, Narasimhan S, Natarajan B, Parvin F, Paul S, Pradhan K, Praveena G, Prasad DLKV, Poswal HK, Pujari B, Pushpa R, Reddy KHH, Saha SK, Sbraccia C, Scandolo S, Seal P, Shafai GS, Shavanas KV, Simrall JOH, Srirangarajan A, Srivastava V, Talati GS, Tantirungrotechai Y, Tarafder K, Thomas T, Uthayathasan T (2007) Solid State Commun 144:273–276
10. Gasperin PM (1987) Acta Cryst C43:2031–2033
11. Santamaría-Pérez D, Vegas A (2003) Acta Cryst B 59:305–323
12. Kresse G, Furthmuller J (1996) Phys Rev B 54:11169
13. Kresse G, Joubert D (1999) Phys Rev B 59:1758–1775
14. Perdew JP, Burke K, Ernzerhof M (1996) Phys Rev Lett 77:3865–3868
15. Monkhorst HJ, Pack JD (1976) Phys Rev B 13:5188–5192
16. Murnaghan FD (1944) Proc Natl Acad Sci 30:244–247
17. Schulte O, Holzapfel WB (1993) Phys Rev B 48:767–773
18. Sleep NH, Zahnie K (2001) J Geophys Res-Planet 106:1373–1399
19. Panero WR, Kabbes JE (2008) Geophys Res Lett 35:L14307
20. Boulard E, Gloter A, Corgne A, Antonangeli D, Auzende AL, Perrillat JP, Guyoy F, Fiquet G (2011) Proc Natl Acad Sci 108:5184–5187
21. Mujica A, Rubio A, Muñoz A, Needs RJ (2003) Rev Mod Phys 75:863–912

REGULAR ARTICLE

Simulating the optical properties of CdSe clusters using the RT-TDDFT approach

Roger Nadler · Javier Fdez Sanz

Abstract The structural and electronic properties of the CdSe nanoclusters, which have been intended to model quantum dots, have been examined by means of time-dependent density functional (TDDFT) calculations. The optical spectra were first simulated using the standard linear response implementation of the TDDFT (LR-TDDFT) in a series of calculations performed using different basis sets and exchange–correlation functionals. In a second step, the real-time TDDFT implementation (RT-TDDFT) was used to simulate the optical absorption spectra of the CdSe nanoclusters, both naked and capped with ligands. In general, we found that the RT-TDDFT approach successfully reproduced the optical spectrum of CdSe clusters offering a good compromise to render both the optical and the geometrical properties of the CdSe clusters at lower computational costs. While for small systems, the standard TDDFT is better suited, for medium- to large-sized systems, the real-time TDDFT becomes competitive and more efficient.

Keywords TDDFT · CdSe · Absorption spectrum · UV–Vis · Quantum dot · Real-time TDDFT

R. Nadler · J. F. Sanz (✉)
Department of Physical Chemistry, University of Seville,
41012 Seville, Spain
e-mail: sanz@us.es

1 Introduction

The flexibility and variety of structures of quantum-confined, semiconducting clusters nowadays opens a vast field of applications. These so-called quantum dots (QDs) can be tuned in size and shape by controlling temperature, solvents, and ligands. For instance, different organic ligand molecules prefer to bind to different facets of a QD, and therefore, the nanocluster will grow along specific directions. Such properties make available the growth of complex nanostructures like tetrapods [1] or hourglass-like nanoclusters [2]. Most certainly, advances in experimental and theoretical methodologies will lead to applications and structures that today just exist in our imagination.

Quantum dots are thought of being artificial atoms since their electronic wave functions are confined in 3 dimensions. This confinement leads to a quantization of the electronic states, resulting in localized orbitals rather than in a band structure description of a bulk semiconductor. The size of a QD is typically within the range of a few nm, which is an order of magnitude smaller than the Bohr exciton radius in the corresponding bulk material. Another consequence of the small size of the QDs is a high surface-to-volume relation, which leads to a particular restructuring of the QD surface. For a CdSe QD, it was shown that the tendency is such that the Cd surface atoms move into the cluster while the Se atoms are puckered out [3, 4]. It is clear that the electronic environment of these surface atoms is different from bulk atoms. Unsaturated bonds occur on the surface, they being a source of surface trap states where electronic charges are trapped. Trap states are responsible for bleaching, that is, when no more excitons can be generated due to the large amounts of energy needed to add another exciton. This process is highly unwanted since excitons formed upon irradiation should be harvested to,

for example, produce electricity in solar cells sensitized with QDs or emit light in a diode. To avoid trap states, these unsaturated bonds, or dangling bonds, need to be passivated.

The QD's popularity comes also from the fact that they possess a high photostability and, most importantly, their absorption spectra can be tuned by simply changing their size. Moreover, the organic capping ligands can easily be replaced by different ligands that could render the QDs water soluble, for instance [5]. This, in turn, offers applications such as bio-labeling material.

Experimentally, there were several attempts to identify the size of CdSe nanoclusters [6–9]. For instance, the cage/core clusters containing 13, 33, and 34 CdSe units were found to be exceptionally stable CdSe clusters by time-of-flight mass spectroscopy [7]. In the smallest of these so-called magic-sized clusters (MSC), the core is formed by one Se atom. A 6-unit cluster just fits nicely inside the $(CdSe)_{28}$ cage which results in a higher binding energy per CdSe unit than for the $(CdSe)_{33}$ cage/core cluster [4]. The first absorption peak of 415 nm (2.98 eV) was attributed to these larger clusters. However, Nguyen et al. [10] found in their calculations that the experimentally observed $(CdSe)_{34}$ cluster is actually the least stable one. Other work used the cluster model consisting of 33 CdSe units [11–16]. Nguyen et al. correlated their result to the fact that the clusters observed by Kasuya et al. by laser ablation experiments might not follow the thermodynamics of $(CdSe)_n$ as they investigated it. Then, Del Ben et al. were able to assign the excitonic transition between 350 and 360 nm (3.44–3.54 eV) to the $(CdSe)_{13}$ cluster by comparing the experimental spectrum obtained by Kudera et al. [8] with that simulated for this specific cluster. Just recently, an experimental paper was published where the authors selectively synthesized $[(CdSe)_{13}(n\text{-octylamine})_{13}]$ and $[(CdSe)_{13}(oleylamine)_{13}]$ nanoclusters [9]. In this paper, the correctness of the assignation was justified by presenting elemental analysis and mass spectroscopy. Also, they used the corresponding spectrum that Del Ben et al. [11] did present for the $(CdSe)_{13}$ QD capped with formate/hydrogen ligand pairs. Both spectra coincide nearly perfectly, even though the two groups used different ligands to obtain their respective spectrum. This is somewhat surprising since Del Ben et al. presented a spectrum for the $(CdSe)_{33}$ cluster that has the dangling bonds saturated with methylamine which is 0.5 eV lower in energy than the QD saturated with the formate/hydrogen ligand pair.

Typically, organic ligand molecules like phosphines, phosphine oxides, thiols, and amines are used to cap the QD, but there are reports that describe the ligation of the QDs with metal-free inorganic ligands like HS^-, S^{2-}, OH^-, and TeS_3^{2-} [17]. Usage of such ligands could increase the efficiency of QDSSC due to a stronger electronic QD–QD coupling [18]. However, it provides the possibility to tune the properties as such as to use QDs as electronic devices, too. Unfortunately, these ligands are not easily incorporated in a model system that could be used for such calculations due to their charged nature.

Albert et al. [15] investigated the dependence of the electronic structure of the $(CdSe)_{33}$ cluster on the computational method and stated that at least a LANL2DZ basis set is necessary to obtain accurate geometries of the clusters. For organic ligands, they propose to also include polarization functions in order to calculate correct binding energies between clusters and ligand molecules. The effect of the exchange–correlation (xc) functional on the lowest-energy optical transitions is more pronounced. Generalized gradient approximation (GGA) functionals underestimate band gaps and miss relevant excitonic effects in the optical transitions, whereas hybrid functionals give results that are more satisfying. On the other hand, they found that solvent effects have only little influence on the absorption spectrum, blue-shifting it by 0.2–0.3 eV. A different study investigated the influence of amines and phosphine oxides on the optical spectrum [12]. They also varied the number of capping ligands in order to observe a change in the spectra. These hybridized surface states near the band gap increase the density of electronic states, the effect being stronger for phosphine oxides than for amines. This behavior suggests a more likely electron–phonon coupling with high-frequency vibrations in the ligand, thus providing a possible explanation of why electron relaxation is more effective when CdSe QDs are capped with phosphine oxides than with amines.

Most commonly, to calculate the electronic spectra to study optical properties of QDs, the linear response time-dependent density functional theory (LR-TDDFT) approach [19] is used [6, 15, 20]. Another methodology is the so-called configuration interaction (CI) [21], applied, for example, in the work of Isborn et al. [22]. The frequency domain or standard implementation of TDDFT and CI is, however, demanding for calculations of medium- and large-sized systems. Recently, it was shown that real-time TDDFT (RT-TDDFT) is a convenient alternative in the computation of the absorption spectra of organic dyes used in solar cells (DSSC) [23–27]. The technique allows for the simulation of not only the dye, but also a relatively well-sized [23–25] cluster representing the semiconducting support. The methodology allows even more elaborate approaches used to calculate the rate of the photoinduced electron transfer [28].

In this paper, we report a computational analysis of the electronic absorption spectra of CdSe QDs by means of RT-TDDFT using a variety of xc-functionals and basis sets available in an all-purpose solid-state chemistry code. In order to facilitate the analysis of the spectra thus obtained,

several spectra have also been computed applying the LR-TDDFT approach. Our aim is to benchmark several xc-functionals and basis sets and to check if and to what extent such technical details influence geometry and optical spectrum of the CdSe quantum dots. After that, we calculated and compared the absorption spectrum of a given nanocluster where different Cd surface atoms were saturated.

2 Models and computational details

Three differently sized CdSe clusters and its ligated variants were employed in this work. The smallest cluster consists of 6 units of CdSe (Fig. 1a, b). It is the smallest possible cluster that still reproduces the wurtzite structure of bulk CdSe. The bond distances d_1 are determined as the Cd–Se bond within the 6-membered ring, whereas d_2 is the bond between a Cd and Se atom located in two adjacent rings. The bond angle α_1 is calculated as the bond angle including Se–Cd–Se, all of them form part of the hexameric ring. Finally, α_2 is the bond angle that includes again Se–Cd–Se but one Se atom is in the other ring. To saturate the Cd atoms, we used 6 molecules of methylamine (MA) (Fig. 1c, d). Yang et al. [29] reported that even for the small $(CdSe)_6$ cluster, the 2-coordinated Cd atoms are not saturated with two ligand molecules; therefore, we will not employ such model systems for ligated clusters. We employed these clusters to simulate the absorption spectra based upon both LR-TDDFT frequency domain calculations and real-time propagation of the wave function in the time domain (RT-TDDFT). Following this procedure, we are able to benchmark the spectra obtained with RT-TDDFT against the LR-TDDFT methodology. Additionally, modeled after the cage/core clusters reported by Kasuya et al. [7], we employed here the $(CdSe)_{13}$ and the $(CdSe)_{34}$ clusters. The medium-sized $(CdSe)_{13}$ cluster consists of a 1 Se atom core as represented in Fig. 2a, b. The saturated cluster, $(CdSe)_{13}(MA)_9$, is sketched in Fig. 2c, d. This is somewhat different to the model employed by Yang et al. [29] where they use 10 ligand molecules. Here, we find that one Cd atom indeed forms bonds to 4 adjacent Se atoms; therefore, we see no need for it to be saturated with a ligand molecule. Finally, the largest cluster, which we have considered for this work, consists of a cage of 28 CdSe units into which a $(CdSe)_6$ core is introduced, cf. Fig. 3a, b. From the 28 surface Cd atoms, six are 4-coordinated. The remaining 22 surface Cd atoms are all 3-coordinated and, consequently, are saturated with one MA molecule (Fig. 3c, d).

From a computational point of view, this work consists of two parts. The first part includes LR-TDDFT calculations, which were performed using the GAUSSIAN 03 program suite [30]. In order to simulate the absorption spectra, over 400 singlet transitions were employed in selected calculations. Four different density functionals and a large number of different basis sets were employed as to investigate the influence of them on the spectra. The basis sets used were LANL2DZ, cc-pVDZ, def2-SVP, and def2-TZVP for both Cd and Se. Effective core potentials (ECP) were employed to describe the core region. For the LANL2DZ basis set, the valence electron density is defined by the $4d^{10}5s^2$ electrons of each Cd and the $4s^2p^4$ electrons for the Se atoms. The valence region described with the cc-pVDZ basis set contains the $4s^2p^6d^{10}5s^2$ electrons for Cd and for Se the $3s^2p^64s^2p^4$ electrons. The def2-SVP and def2-TZVP basis sets include the same valence for Cd as for the cc-pVDZ basis set, while for Se it is an all-electron basis set. When the $(CdSe)_6$ cluster is ligated with methylamine, then the 6-31G(d) basis set was applied to C, N, and H. The xc-functionals were represented by the GGA density functionals PBE [31] and BLYP [32, 33], and two hybrid functionals, PBE0 [34–37] and B3LYP [33, 38]. Since the B3LYP functional was parameterized against the G1 database of Pople et al. [39], for which it gives excellent results, its performance might differ when a system is investigated that contains heavier atoms, as it is the case for semiconductors. It is worth noting that the combination LANL2DZ/B3LYP was reported to reproduce adequately the electronic properties of CdSe systems [40]. The PBE0 hybrid functional, on the other hand, is known to overestimate the band gaps of semiconducting material. The $(CdSe)_6(MA)_6$ system was only calculated with PBE. We did not introduce a solvent model in the calculations as it was shown that the effect on the spectrum is not significant [15].

The second part consists of the calculation of absorption spectra in the time domain using the CP2K/QUICKSTEP program [41, 42], which uses a hybrid Gaussian and plane-wave basis set based on DFT. Here, the PBE and the hybrid PBE0 functionals were considered together with norm-conserving Goedecker–Teter–Hutter (GTH) pseudopotentials [43–45]. For both Cd and Se, the short range, molecularly optimized double-ζ single-polarized basis set (m-SR-DZVP) was applied [46]. For Cd, the $4d^{10}5s^2$ valence electrons are included and for Se the $4s^2p^4$ electrons. It was reported that increasing the basis set for Cd and Se to triple-ζ quality does not result in relevant differences [11]. The m-DZVP basis set is applied on the ligand atoms only. It contains more diffuse functions [46], which is in spirit with Albert et al. [15] who proposed to use polarizable functions at least on the ligand atoms. The geometries were optimized until the gradients were smaller than 0.01 eV/Å and the cutoff for the SCF procedure was set to 10^{-7}. The box length of the simulation cell was set as such that in each direction at least 15 Å of vacuum is

Fig. 1 Sketch of the $(CdSe)_6$ cluster. The *top* view and *side* view of the bare cluster are represented in **a** and **b**, and for the ligated cluster, they are sketched in **c** and **d**. In **a**, the bond distance d_1 and the bond angle α_1 are indicated, whereas in **b**, d_2 and α_2 are shown. See the main text for details. Cd atoms are represented in *green*, Se in *yellow*, N in *blue*, C in *black*, and H in *white*
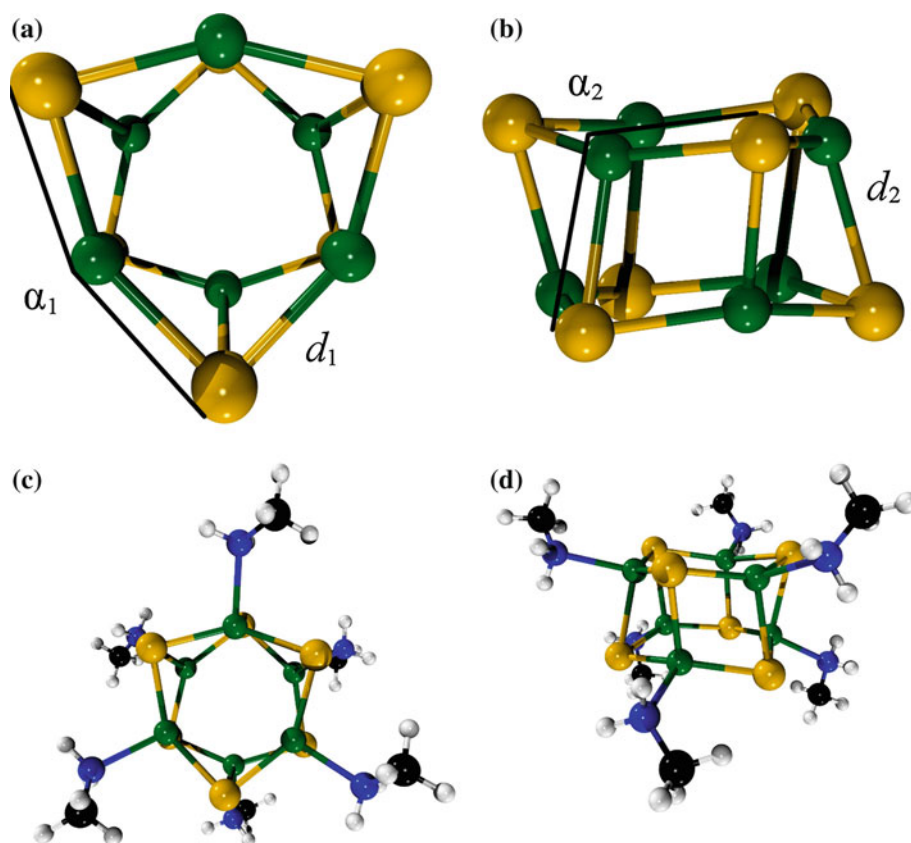
introduced. The absorption spectra were calculated using the RT-TDDFT methodology introduced by Chen et al. [27]. Shortly, this method consists in applying an ultra-short electromagnetic pulse on the ground-state electronic wave function. It follows the propagation of the perturbed wave function during which the dipole moment for each time step was calculated. Through a Fourier transformation of the induced dipole moment, the absorption cross section will be obtained. Similar to their work, we set a stepwise electric field pulse with duration of 0.0121 fs and a magnitude of 0.5 V/Å. The excited wave function was propagated during 4000 steps, using the ETRS propagator [47], with a time step of 0.0121 fs.

3 Results

First, we look at the geometrical properties of the bare cluster. We observe that the performance of the different basis sets changes. Larger basis sets give better geometries than the smaller ones [15]. In Table 1 are the geometrical parameters tabulated that define the bare cluster. At the B3LYP/LANL2DZ level of theory, the same geometry was reported earlier in literature [6, 10]. The bond lengths obtained with the PBE functional and the m-SR-DZVP basis set almost exactly reproduced the bond lengths

obtained when the def2-TZVP basis set is used, indicating a very nice performance of these basis sets in terms of geometrical properties. Again, going along with the findings of Albert et al. [15], we observe that the geometry is quite basis set dependent.

Now, we turn our attention to the optical absorption spectra. In a first step, we quantify the blue-shift induced by saturating the dangling bonds on the Cd atoms by comparing the absorption spectra of $(CdSe)_6$ and the ligated cluster. The spectra computed with the LANL2DZ basis set and the PBE functional, within the LR-TDDFT approach, are represented in Fig. 4. In the bare cluster, the transition occurs at a first absorption peak maximum of 2.68 eV. The excitation takes place between the HOMO-2, which is composed by a large part of $4p$ Se orbitals, and the LUMO, which mainly consists of the $5s$ orbitals on the Cd atoms. This can be rationalized by taking into account the symmetry of the involved MOs. The $(CdSe)_6$ cluster belongs to the D_{3d} point group, and its HOMO-2 has A_{2u} symmetry and its LUMO A_{1g}. The HOMO-1 and the HOMO have E_g symmetry; therefore, the transition between these two MOs and the LUMO is symmetry-forbidden. The first absorption peak maximum for $(CdSe)_6(MA)_6$ is found at 3.08 eV, in very good agreement with the experimental value of 3.14 eV obtained in a TOP/TOPO solvent mixture [6]. The excitonic transition at

Fig. 2 Sketch of the $(CdSe)_{13}$ cluster. The *top* view and *side* view of the bare cluster are represented in **a** and **b**, and for the ligated cluster, they are sketched in **c** and **d**. There are four 4-coordinated Cd atoms in this cluster. Therefore, only 9 molecules of MA are necessary to saturate the under-coordinated Cd atoms. Cd atoms are represented in *green*, Se in *yellow*, N in *blue*, C in *black*, and H in *white*
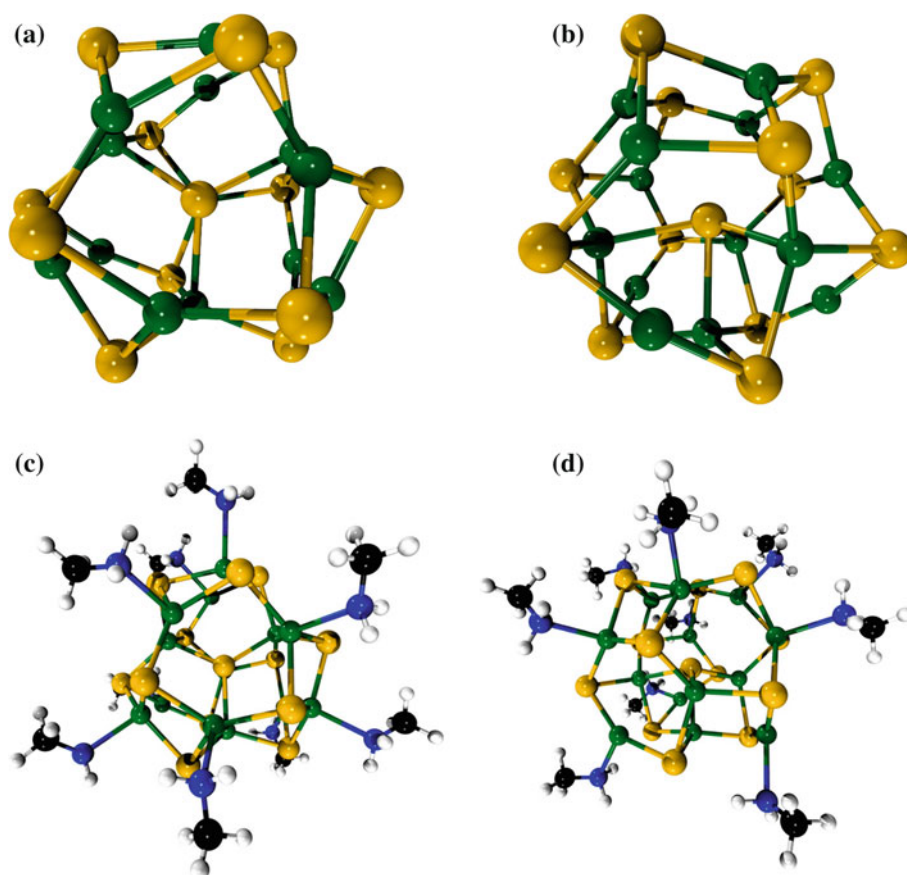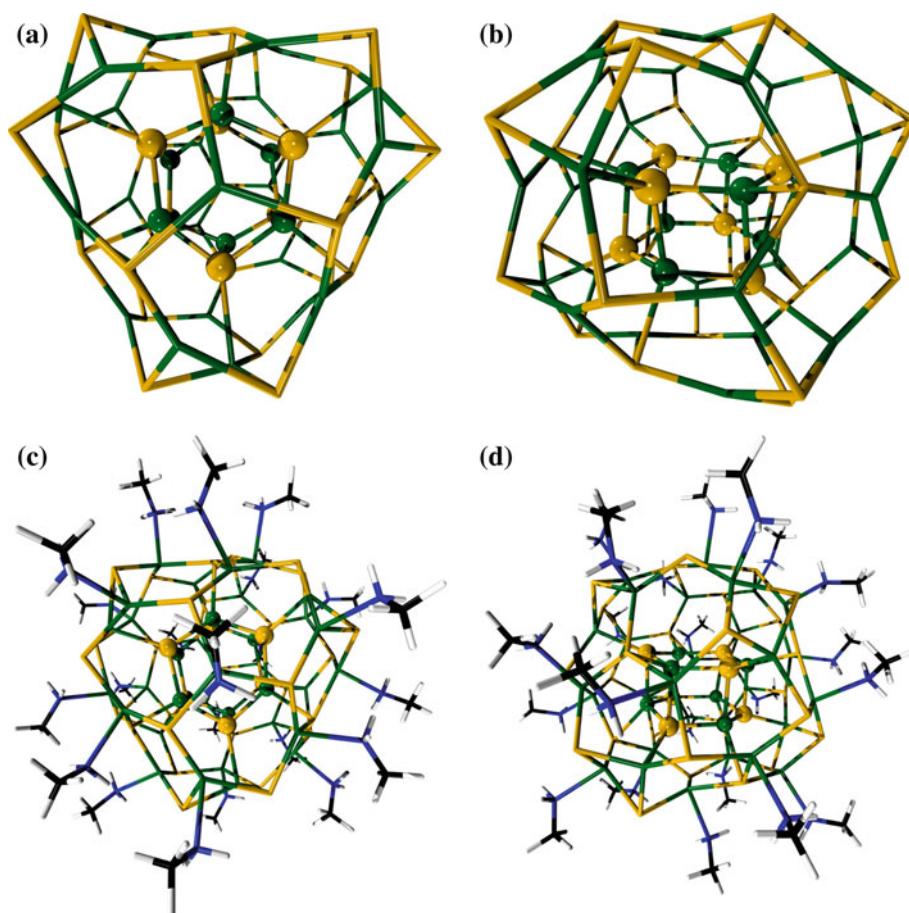
3.08 eV includes the HOMO-2, HOMO-1, and HOMO orbitals and excites an electron into the LUMO. The two-fold degeneracy of the HOMO-1 and HOMO must have been broken due to the presence of ligand molecules. We also calculated the spectra for bare cluster where the cc-pVDZ, def2-SVP, and the def2-TZVP basis sets were applied in combination with the PBE functional. We find that these larger basis sets shift the spectrum to the red, which is not desirable in this context.

Let us now compare the spectra obtained with the different xc-functionals. The spectra computed using the LANL2DZ basis set for both Cd and Se are reported in Fig. 5. Clearly, the BLYP functional does not account satisfactorily for the position of the first absorption maximum; it underestimates it by roughly 0.6 eV. On the other extreme, the PBE0 hybrid functional does suffer from its known limitations: it overestimates the first absorption peak by almost 0.4 eV. The B3LYP functional performs much better as it gives a value that is only slightly below the PBE/LANL2DZ result, the difference being as small as 0.06 eV.

We now compare the spectra obtained with RT-TDDFT to the ones simulated with LR-TDDFT methodology. The spectra for the bare and ligated $(CdSe)_6$ clusters are plotted in the top panel of Fig. 6. Compared to the PBE/

LANL2DZ spectra of the bare cluster, we observe a red shift of 0.3 eV. This energy is equivalent to a difference of 0.6 eV compared to literature values [4, 22]. When the dangling bonds on Cd are saturated, the performance is slightly better but the energy of the first peak is still underestimated by ~ 0.14 eV compared to the experimental value [6], and 0.08 eV below the PBE/LANL2DZ value. We think that this is an acceptable deviation since this method allows for the simulation of much larger systems for which LR-TDDFT would be too cumbersome. The blue-shift when the $(CdSe)_6$ is completely ligated, compared to the bare cluster, is found to be 0.6 eV. Surprisingly, the band edge absorption for the bare $(CdSe)_6$ and $(CdSe)_{13}$ clusters nearly coincides. In principle, the band edge for the bare $(CdSe)_{13}$ cluster should be red-shifted due to its larger diameter. For the fully saturated $(CdSe)_{13}(MA)_9$ cluster, the expected red shift is observed again, it being 0.5 eV. For the largest bare cluster $(CdSe)_{34}$, represented in Fig. 3, the first absorption peak is located at 1.86 eV which agrees quite well with results reported earlier in literature [11, 15], the differences being of only 0.1 eV. It should be noted that the herein presented values are not shifted by the difference between the band gap of the bulk material and the theoretical band gap as it was done in Ref. [11]. Our spectrum of the $(CdSe)_{34}$ cluster

Fig. 3 Illustration of the $(CdSe)_{34}$ cluster. The *top* view and *side* view of the bare cluster are represented in **a** and **b**, and for the ligated cluster, they are sketched in **c** and **d**. Each Se atom of the $(CdSe)_6$ core forms one bond with a Cd atom in the cage. This leads to six 4-coordinated Cd surface atoms, and therefore, 28 MA molecules bound to the remaining under-coordinated Cd atoms. Cd atoms are represented in *green*, Se in *yellow*, N in *blue*, C in *black*, and H in *white*

which shows up at slightly higher energies: 1.86, 2.02, and 2.16 eV, respectively. However, the second peak here is the most intense out of the three peaks while in Ref. [11] it is the least intense one. This could be explained by the orientation of the ligands. Depending on their orientation, the symmetry of the cluster's MOs would be broken in a way that favors a given transition while a different orientation of ligands will favor a different one.

Table 1 Bond lengths, bond angles, and optical band gaps for the bare $(CdSe)_6$ cluster

	d_1	d_2	α_1	α_2	E_{abs}
PBE, LANL2DZ	2.71	2.87	141.0	100.8	2.68
PBE0, LANL2DZ	2.66	2.82	140.7	100.7	3.03
BLYP, LANL2DZ	2.75	2.91	139.9	100.5	2.08
B3LYP, LANL2DZ	2.70	2.86	139.9	100.4	2.63
PBE, cc-pVDZ	2.60	2.83	145.7	100.7	2.03
PBE0, cc-pVDZ	2.58	2.79	144.6	100.6	–
PBE, def2-SVP	2.61	2.83	145.3	100.4	2.12
PBE, def2-TZVP	2.60	2.83	145.4	100.5	2.11
PBE, m-SR-DZVP	2.61	2.83	145.3	100.3	2.38
PBE0, m-SR-DZVP	2.56	2.80	145.9	100.2	3.41

The distances d_1 and d_2 and the angles α_1 and α_2 are defined in Fig. 1. Distances are in (Å), angles in (°), and the first absorption maximum, E_{abs}, in (eV)

nicely reproduces the one obtained by Del Ben et al. in that their spectrum of the bare $(CdSe)_{33}$ cluster also shows two peaks next to the first absorption peak, located at 1.76 eV, at 1.96 and 2.09 eV (these values are obtained by subtracting 0.43 eV from their results). This is nearly identical to the triple peak in the lowest panel shown in Fig. 6,

Fig. 4 Optical absorption spectra obtained for the $(CdSe)_6$ and $(CdSe)_6(MA)_6$ clusters. *Yellow lines* correspond to the bare cluster's spectrum, and *black lines* to the caped cluster
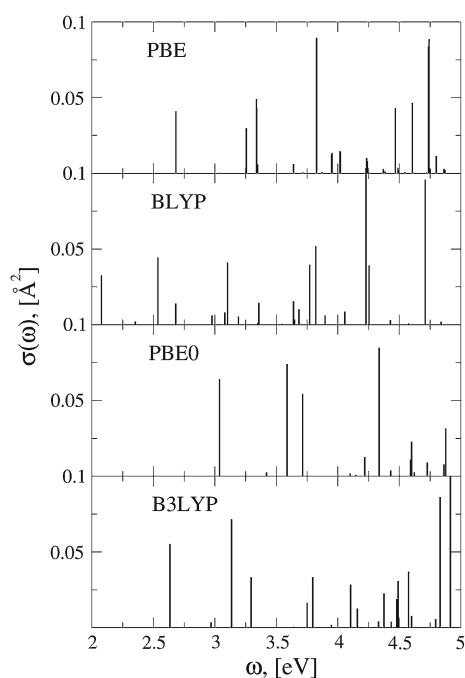
Fig. 5 Absorption spectra of the bare $(CdSe)_6$ cluster obtained with different xc-functional as indicated within each *panel*
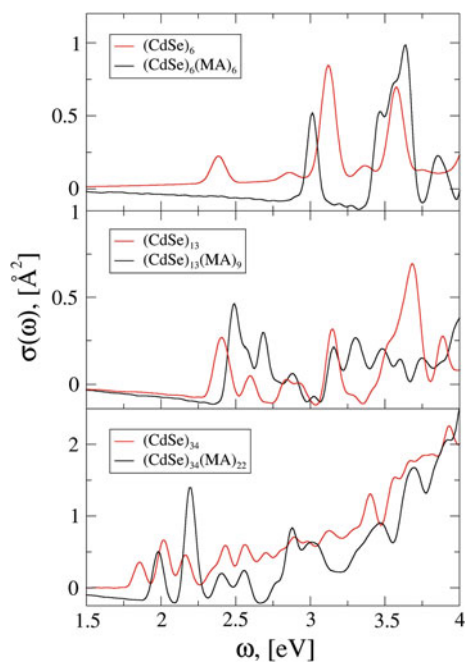
Fig. 6 Absorption spectra calculated with the RT-TDDFT methodology. All spectra were obtained using the PBE functional. In the *top panel* are the spectra of the bare and ligated $(CdSe)_6$ cluster plotted. In the *middle panel*, the corresponding spectra that represent the $(CdSe)_{13}$ and $(CdSe)_{13}(MA)_9$ clusters. In the *lowest panel* are the spectra of $(CdSe)_{34}$ and $(CdSe)_{34}MA_{22}$ plotted. *Red lines* correspond to the bare clusters, and *black lines* to the saturated ones
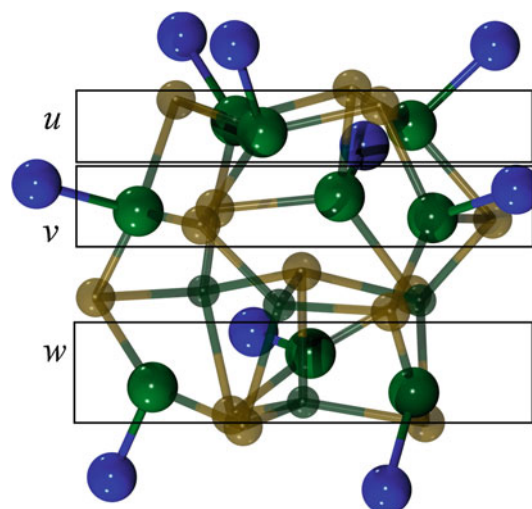
Fig. 7 Represented are the three planes that each consists of 3 Cd atoms that have a similar bonding environment. In plane *u*, the geometry resembles to the 6-membered ring in $(CdSe)_6$. The Cd atoms in plane *v* have a similar bond order than those in plane *u*. Finally, in plane *w*, the Cd atoms are only 2-coordinated
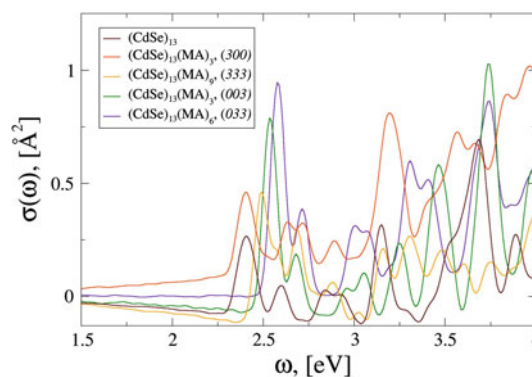
Fig. 8 Illustrated is a series of spectra obtained from the $(CdSe)_{13}$ and $(CdSe)_{13}(MA)_n$ clusters, each having different Cd surface atoms saturated. See the main text for details. The spectra are ordered in ascending order with respect to the first absorption peak

For the saturated $(CdSe)_{34}(MA)_{22}$ cluster, we find $E_{abs} = 1.98$ eV, an underestimation compared to theoretical literature values by approximately 0.16 eV [11] and 1.0 eV when compared to experiment where the CdSe clusters are dissolved in toluene [7]. From this discussion, the disadvantage of the RT-TDDFT methodology emerges. It is very difficult to assign the absorption features in the spectrum and compare them to other experimental and theoretical work because no information about the MOs involved in the excitation is available.

We now analyze the effect of a different number of ligands on the spectrum. The side view in Fig. 7 shows

Table 2 First absorption peak maximum, binding energies of the ligands and bond Cd–N bond lengths for the different ligand/cluster models

	E_{abs}	E_{bind}	$d_{Cd–N}$, u	$d_{Cd–N}$, v	$d_{Cd–N}$, w
(000)	2.39	–	–	–	–
(300)	2.39	0.50	2.46	–	–
(333)	2.49	0.45	2.51	2.50	2.47
(003)	2.54	0.60	–	–	2.45
(033)	2.58	0.52	–	2.47	2.48
(030)	–	0.57	–	2.44	–

The bond lengths are given for each plane as they are defined in Fig. 7. The (030) cluster is tabulated for the sake of completeness; no spectrum was calculated for it. E_{abs} and E_{bind} are given both in (eV); all bond lengths are given in (Å)

how the $(CdSe)_{13}$ cluster is divided into three planes u, v, and w, respectively. Each plane consists of 3 Cd atoms that mainly differ in their coordination number. In the planes u and v, each Cd atom is 3-coordinated while in plane w they are only 2-coordinated. The difference between the Cd atoms in plane u and v is the charge that they carry in the bare cluster which we estimate to be 0.340 e in plane u and 0.346 e in planes v and w [48]. From these differences, it might be possible that the optical absorption spectrum will be different, too. To investigate the influence of each ligand plane on the absorption spectrum, we modeled in total 4 cluster–ligand systems: $(CdSe)_{13}(MA)_n$ with $n = 3, 6, 9$. We label them in the following manner. The Cd atoms of each plane are saturated with at most three ligands, or they are not saturated at all. We did not consider adding only one or two MAs. Therefore, the fully saturated cluster will be labeled as (333), while the bare cluster is labeled (000), the labels referring to the planes u, v, and w, respectively. The following cluster–ligand systems were employed: (003), (300), (033), and (333), plus the bare cluster: (000). The corresponding spectra are plotted in Fig. 8. The plots are ordered with respect to their first absorption maximum feature. Their values are tabulated in Table 2. We observe that the largest blue-shift compared to the bare cluster is induced by the (033) model ($E_{abs} = 2.57$ eV), followed by the (003) cluster. Between the two extremes, we find the fully saturated cluster, (333), with an $E_{abs} = 2.49$ eV. The second $(CdSe)_{13}(MA)_3$ cluster/ligand system, where the u plane Cd atoms are saturated, hardly shows a shift in the first absorption peak maximum. Obviously, it is not the (333) system that experiences the largest blue-shift but the one where the Cd atoms in the two lowest planes v and w are saturated, (033). Clearly, the impact on the spectrum is largest if the 2-coordinated Cd atoms are saturated, while the influence of saturating 3-coordinated Cd atoms in plane v is smaller and it is nearly inexistent when ligands are bound to the

u plane Cd atoms. The small impact of the u plane ligands on the spectrum can be rationalized by analyzing the binding energies of the 3 ligands to the cluster, $E_{bind} = 0.5$ eV, the smallest energy among the systems where only 3 ligands are bound. Yet, the Cd–N bond distance is similar in all three cases. In addition, when no ligands are added, the average charge is not as positive on the w plane Cd atoms than on the u and v plane cadmiums (0.340 e vs. 0.346 e). For comparison, the average charge is most positive on the 4-coordinated Cd atoms (0.362 e). Obviously, the need for saturation of the Cd atoms is smaller in plane u than on other surface Cd atoms in the planes v and w.

When it comes to the PBE0 functional, we find that it overestimates the first absorption peak significantly, both for the $(CdSe)_6$ and the $(CdSe)_{13}$ bare clusters. In the latter, the blue-shift is even more pronounced. While for the smaller QD the difference between the first absorption peaks is 1.0 eV, it increases to 1.3 eV when 13 CdSe units form the QD. Albert et al. [15] reported that one should employ an asymptotically corrected functional like the LC-ωPBE to calculate the energy of the lowest transition closer to experiments.

4 Conclusion

Concerning the geometrical properties of the CdSe clusters, we find that the m-SR-DZVP basis set and the PBE xc-functional give results that are close to results obtained with larger basis sets like def2-TZVP or cc-pVDZ. When the PBE functional is used together with the LANL2DZ basis set, the bond lengths are overestimated compared to the large basis sets.

We also find that the optical properties of CdSe QDs can be successfully reproduced with the RT-TDDFT method using the m-SR-DZVP basis sets and the PBE GGA xc-functional. Therefore, this level of theory offers a good compromise to calculate both geometries and spectra. The PBE0 functional is not recommended for the calculation of optical properties of CdSe clusters since the optical band gap is severely overestimated while B3LYP would be a better choice.

For the $(CdSe)_{13}$ cluster, we find that the absorption spectrum depends on which Cd surface atoms the methylamine ligands are adsorbed to. In particular, when planes v and w are saturated, the cluster/ligand system shows the strongest blue-shift in the optical spectrum. Moreover, compared to clusters with fewer capping ligands, the fully saturated cluster (9 ligand molecules) shows lower excitation energies.

In summary, the comparison of the conventional LR-TDDFT approach with the less conventional but still already mature RT-TDDFT shows the latter competitive

for large systems, but in contrast, it is not able to give the assignment of the spectral feature. This is an important limitation of RT-TDDFT, which is very stimulating for new theoretical research to overcome or, on the other hand, to improve the LR-TDDFT algorithm in order to become more efficient in large systems but keeping the possibility to give an assignment to the spectral feature, which is a crucial point for the merit of a computational method.

References

1. Manna L, Milliron DJ, Meisel A, Scher EC, Alivisatos AP (2003) Nat Mater 2:382–385
2. Watt J, Yu C, Chang SLY, Cheong S, Tilley RD (2012) J Am Chem Soc. doi:10.1021/ja311366k
3. Puzder A, Williamson AJ, Gygi F, Galli G (2004) Phys Rev Lett 92:217401–217405
4. Botti S, Marques MAL (2007) Phys Rev B 75:035311–035317
5. Wang J, Xu J, Goodman MD, Chen Y, Cai M, Shinar J, Lin Z (2008) J Mater Chem 18:3270–3274
6. Jose R, Zhanpeisov NU, Fukumura H, Baba Y, Ishikawa M (2006) J Am Chem Soc 128:629–636
7. Kasuya A, Sivamohan R, Barnakov YA, Dmitruk IM, Nirasawa T, Romanyuk VR, Kumar V, Mamykin SV, Tohji K, Jeyadevan B, Shinoda K, Kudo T, Terasaki O, Liu Z, Belosludov RV, Sundararajan V, Kawazoe Y (2004) Nat Mater 3:99–102
8. Kudera S, Zanella M, Giannini C, Rizzo A, Li Y, Gigli G, Cingolani R, Ciccarella G, Spahl W, Parak WJ, Manna L (2007) Adv Mater 19:548–552
9. Wang Y, Liu YH, Zhang Y, Wang F, Kowalski PJ, Rohrs HW, Loomis RA, Gross ML, Buhro WE (2012) Angew Chem Int Ed 51:6154–6157
10. Nguyen KA, Day PN, Pachter R (2010) J Phys Chem C 114:16197–16209
11. Del Ben M, Havenith RWA, Broer R, Stener M (2011) J Phys Chem C 115:16782–16796
12. Kilina S, Ivanov S, Tretiak S (2009) J Am Chem Soc 131(7717):7726
13. Kilina S, Kilin DS, Prezhdo OV (2009) ACS Nano 3:93–99
14. Kilina S, Velizhanin KA, Ivanov S, Prezhdo OV, Tretiak S (2012) ACS Nano 6:6515–6524
15. Albert VV, Ivanov SA, Tretiak S, Kilina SV (2011) J Phys Chem C 115:15793–15800
16. Abuelela AM, Mohamed TA, Prezhdo OV (2012) J Phys Chem C 116:14674–14681
17. Nag A, Kovalenko MV, Lee JS, Liu W, Spokoyny B, Talapin DV (2011) J Am Chem Soc 133:10612–10620
18. Hillhouse HW, Beard MC (2009) Curr Opin Colloid Interface Sci 14:245–259
19. Casida ME (1996) Recent developments and applications of modern density functional theory. In: Seminario M (ed) Theoretical and computational chemistry, vol 4. Elsevier, Amsterdam, p 391
20. Inerbaev TM, Masunov AE, Khondaker SI, Dobrinescu A, Plamadă AV, Kawazoe Y (2009) J Chem Phys 131:044106–044112
21. Nakatsuji H, Hirao K (1978) J Chem Phys 68:2053–2066
22. Isborn CM, Kilina SV, Li X, Prezhdo OV (2008) J Phys Chem C 112:18291–18294
23. Sánchez-de-Armas R, Oviedo J, San Miguel MA, Sanz JF, Ordejón P, Pruneda M (2010) J Chem Theory Comput 6:2856–2865
24. Sánchez-de-Armas R, Oviedo J, San Miguel MA, Sanz JF (2011) J Phys Chem C 115:11293–11301
25. Sánchez-de-Armas R, San Miguel MA, Oviedo J, Márquez A, Sanz JF (2011) Phys Chem Chem Phys 13:1506–1514
26. Chen H, Blaber MG, Standridge SD, DeMarco EJ, Hupp JT, Ratner MA, Schatz GC (2012) J Phys Chem C 116:10215–10221
27. Chen H, McMahon JM, Ratner MA, Schatz GC (2010) J Phys Chem C 114:14384–14392
28. Chen H, Ratner MA, Schatz GC (2011) J Phys Chem C 115:18810–18821
29. Yang P, Tretiak S, Ivanov S (2011) J Clust Sci 22:405–431
30. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, 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, Bakken V, 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 (2004) Gaussian 03, revision C.02. Gaussian, Inc., Wallingford
31. Perdew JP, Burke K, Ernzerhof M (1996) Phys Rev Lett 77:3865–3868
32. Becke AD (1988) Phys Rev A 38:3098–3100
33. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785–789
34. Perdew JP, Ernzerhof M, Burke K (1996) J Chem Phys 105:9982–9985
35. Ernzerhof M, Perdew JP, Burke K (1997) Int J Quantum Chem 64:285–295
36. Ernzerhof M, Scuseria GE (1999) J Chem Phys 110:5029–5036
37. Adamo C, Barone V (1999) J Chem Phys 110:6158–6170
38. Becke AD (1993) J Chem Phys 98:5648–5653
39. Pople JA, Head-Gordon M, Fox DJ, Raghavachari K, Curtiss LA (1989) J Chem Phys 90:5622–5630
40. Yang P, Tretiak S, Masunov AE, Ivanov S (2008) J Chem Phys 129:074709–074721
41. The CP2K Developers Group (2000–2013) http://www.cp2k.org
42. VandeVondele J, Krack M, Mohammed F, Parrinello M, Chassaing T, Hutter J (2005) Comput Phys Commun 167:103–128
43. Goedecker S, Teter M, Hutter J (1996) Phys Rev B 54:1703–1710
44. Hartwigsen C, Goedecker S, Hutter J (1998) Phys Rev B 58:3641–3662
45. Krack M (2005) Theor Chem Acc 114:145–152
46. VandeVondele J, Hutter J (2007) J Chem Phys 127:114105–114114
47. Castro A, Marques MAL, Rubio A (2004) J Chem Phys 121:3425–3433
48. Blöchl PE (1995) J Chem Phys 103:7422–7429

Low-energy nanoscale clusters of (TiC)$_n$ n = 6, 12: a structural and energetic comparison with MgO

**Oriol Lamiel-Garcia · Stefan T. Bromley ·
Francesc Illas**

Abstract Global optimization and density functional theory-based calculations are used to investigate the structure of (TiC)$_6$ and (TiC)$_{12}$ clusters. In both cases, the isomers exhibiting the bulk-like rock salt structure appear to be the most energetically stable. This is in contrast to a number of other materials that have the rock salt structure as their bulk ground state, such as MgO, for which more open tubular motifs are observed. A major factor contributing to this difference in nanoscale structure is likely to be the larger covalent character of TiC as compared to MgO. This is supported by our finding that some of the low-energy isomers of both (TiC)$_6$ and (TiC)$_{12}$ exhibit C–C bonding, whereas the bonding in all analogously sized MgO isomers is solely based on Mg–O ionic bonding. The simulated IR spectra indicate that it should be experimentally possible to differentiate (TiC)$_n$ clusters with or without C–C bonds.

O. Lamiel-Garcia · S. T. Bromley · F. Illas (✉)
Departament de Química Física and Institut de Química Teòrica
i Computacional (IQTCUB), Universitat de Barcelona,
08028 Barcelona, Spain
e-mail: francesc.illas@ub.edu

S. T. Bromley (✉)
Institució Catalana de Recerca i Estudis Avançats (ICREA),
08010 Barcelona, Spain
e-mail: s.bromley@ub.edu

Keywords TiC · Clusters · Nanoclusters ·
Density functional theory · Global optimization

1 Introduction

The technological need to improve our understanding of important inorganic solids (e.g. ZnO, SiC, TiO$_2$) at the nanoscale has renewed the interest in the detailed modelling of the atomic and electronic structure of nanosized clusters of such materials. Nanoclusters of only a few tens of atoms are placed somewhere between molecules and bulk-like nanoparticles of hundreds of atoms. Not only are nanoclusters thus fundamentally interesting for studying how a material's properties scale with increasing size, but also with respect to their physical or chemical properties which may significantly differ from those exhibited by larger nanoparticles and extended solids. Probing and exploiting the specific nanoscale properties of inorganic solids have become major objectives in numerous different fields such as nanotechnology, catalysis, solar cells, electronic components, and material science [1–5]. In addition to their singular properties, the study of inorganic nanoclusters can give insights into growth and nucleation processes which lead to different observed crystal structures [6, 7] in a given material and/or even suggesting the possible existence of new cluster-based materials [8, 9]. In this sense, the ability to predict the most stable ground state structures of nanoclusters, assumed to be those most likely to be found in the experiment, becomes an urgent and important subject. Exhaustive exploration of the potential energy surface of even small-sized clusters is a formidable task due to the large number of possible stable structures which grows exponentially with the number of atoms in the cluster. For this reason, different specific global

optimization approaches have been developed to efficiently search the low-energy isomer spectrum of nanoclusters [10–12]. In this study, we employ global optimization at a classical level of theory followed by refinement using quantum mechanical-based methods to investigate the most stable isomers of nanoclusters of titanium carbide (TiC) with six and twelve formula units.

TiC is a material in the fringe zone between ionic and covalent, with bonding properties that go from metallic, covalent, to ionic [6, 7]. Bulk TiC exhibits the rock salt structure which is commonly found in strongly ionic materials such NaCl or MgO, the latter being a prototypical non-reducible ionic oxide. In spite of its rather simple crystal structure, TiC has shown interesting properties, for instance in heterogeneous catalysis. In this context, TiC, and other simple transition metal carbides, has been suggested as the potential substitutes for expensive noble metal-based catalysts of the platinum group and, furthermore, can display advantages over these bulk transition metals in terms of catalytic selectivity or resistance to poisoning [13, 14]. Transition metal carbides have already proven to be efficient in some reactions of great technological importance, as the water–gas shift reaction [15], and received recent attention in a variety of processes [16–19]. In addition, transition metal carbides have proven to be an excellent support for metallic nanoparticles, especially after the pioneering work of Ono et al. [20] on the catalytic properties of Au nanoparticles supported on TiC. Subsequent work has shown that Au/TiC systems exhibit extraordinary reactivity in DeSOx reactions [21–23], being active to dissociate molecular oxygen [24] and molecular hydrogen [25, 26] and, more recently, even to hydrogenate CO_2 to methanol [27]. In the specific case of TiC, it has been shown that, for the water–gas shift reaction, small clusters adsorb reactants and products too strongly, thus resulting in excessively large energy barriers [15]. A deeper knowledge of TiC nanoparticles seems necessary to be able to properly tune their properties for such applications.

Non-stoichiometric clusters containing Ti and C are known since the early cluster beam work of Castleman and co-workers [28] who termed these clusters as metallocarbohedrenes, or just metcars, where Ti_8C_{12} was found to be particularly stable. These metcars have been studied in detail for a long time from both theoretical and experimental point of view [29–35]. Metcars were found to be especially stable and abundant when obtained by laser vaporization under certain experimental conditions and were clearly identified by mass spectroscopic techniques. Subsequent theoretical investigations were carried out to elucidate the geometry and structure of metcars, showing that they have a cage shape similar to fullerenes formed by pentagonal rings [28, 30]. These results increased the interest of the scientific community in the possible properties and utilities of these

particular clusters [28–30]. Nevertheless, succeeding works found that, depending on the experimental conditions, metcars are not necessarily the most abundant and stable structures, with $Ti_{14}C_{13}$ and also larger 1:1 stoichiometric clusters, being found to be as abundant as the metcars or even more so [29, 30]. Such stoichiometric TiC clusters may also have relevance to astrophysical environments, with experimental measurements and observations both pointing to the presence of such clusters around stars as possible precursors in the nucleation of carbonaceous stardust [36].

In this work, we focus on the study of the ground state atomic structures of the stoichiometric $(TiC)_n$ clusters with $n = 6, 12$ and compare the lowest energy structures to the corresponding clusters exhibited by MgO. For the TiC clusters, we also calculate the vibrational IR spectra in order to find fingerprints to assist in the identification of such clusters in experiment.

2 Methodology and computational details

Since TiC and MgO exhibit the same bulk rock salt structure, it is interesting to see whether this correspondence also holds for the extreme case of small nanoscale clusters. For this comparison, we take the sizes $(TiC)_6$ and $(TiC)_{12}$ and use the structures of the low-energy isomers of the correspondingly sized MgO clusters obtained from global optimization calculations based on the basin-hopping algorithm [37] using simple ionic Born interatomic potentials. As nanocluster structures are highly dependent on the atomic charges employed in such potentials, in order to capture a wide structural variety, the effective ionic charges, $Mg^{q+}O^{q-}$, were varied (between 1e and 2e) and global optimizations were performed for each value of q [38, 39].

More than 15 different isomer structures resulting from the basin-hopping global optimizations were considered for $(TiC)_6$ and $(TiC)_{12}$ in a first round of calculations. The structures and energies of these isomers were first optimized employing a Lennard–Jones two body potential with an Axilrod–Teller potential for the three body interactions following the parameterization by Erkoc [40], using the General Utility Lattice Program (GULP) code [41]. The lowest energy geometries thus obtained were then refined using density functional theory (DFT)-based calculations performed with the VASP code [42–45]. In the DFT calculations, the valence electron density was expanded in a plane wave basis set with a cut-off of 415 eV for the kinetic energy and the effect of the core electrons on the valence density was taken into account using the projector augmented wave method [46, 47]. The exchange–correlation effects were described by the Perdew–Wang (PW91) [48, 49] implementation of the generalized gradient approach (GGA) functional. The calculations were

performed using a cubic cell of $20 \times 20 \times 20$ Å, to guarantee there was no significant interaction between the images of the replicated cells. All calculations were carried out considering the Γ point of the reciprocal space only. Structures were considered converged when forces on the atoms were smaller than 0.03 eV/Å.

In order to further confirm that the lowest energy optimized structures found in the PW91/DFT calculations corresponded to genuine energy minima on the potential energy surface, frequency calculations including dipole moment calculations were carried out. This type of calculation also allows one to compute the intensities corresponding to the lowest transition for each normal mode and thus simulate the infrared (IR) vibrational spectra. The dipole intensity of a given normal mode k corresponding to the z component (μ_z) of the dipole is computed according to the following expression [50–52]

$$I^k = \left(\frac{\partial \mu_z}{\partial Q_k} \right)^2 = \sum_{i=1}^{3n} \left(\frac{P_{ik}}{\sqrt{m_i}} \frac{\partial \mu_z}{\partial \Delta r_i} \right)$$

where μ_z is the z component of the dipole moment, Δr_i are the Cartesian displacements, and $\frac{P_{ik}}{\sqrt{m_i}}$ is the mass-weighted coordinate matrix of this normal mode. Summing the contributions for the x, y and z components permits one to obtain the full theoretical IR spectra for each of the considered structures of the $(TiC)_6$ and $(TiC)_{12}$ clusters.

3 Results and discussion

The ground state structures for both $(TiC)_6$ and $(TiC)_{12}$ are found to exhibit the bulk rock salt structure, see Fig. 1. In both cases, the $(TiC)_6$ and $(TiC)_{12}$ resemble small elongated rock salt slabs. The TiC distances in each of these clusters were found to be similar to the corresponding values in the bulk, but with a small contraction observed in the $(TiC)_6$ structure. This is most clearly noticeable for the TiC bonds at the ends of the cuboid $(TiC)_6$ cluster and can probably be attributed to the low coordination of these surface atoms together with the relatively small cluster size. In the case of $(TiC)_{12}$, the bonds involving the atoms on the "rim" of the slab also show a contraction, although here a small expansion relative to the bulk, with noticeable larger bond distances, is observed in the innermost part of the structure. A selection of bond distances for the ground state $(TiC)_6$ and $(TiC)_{12}$ clusters and the bulk rock salt structure are compared in Table 1.

It is interesting to compare the differences between these structure and those corresponding to $(MgO)_n$ clusters with the same number of atoms. The ground state structures for small $(MgO)_n$, $n = 6, 12$, clusters tend to exhibit an hexagonal tubular shape [53]. Although the lowest energy

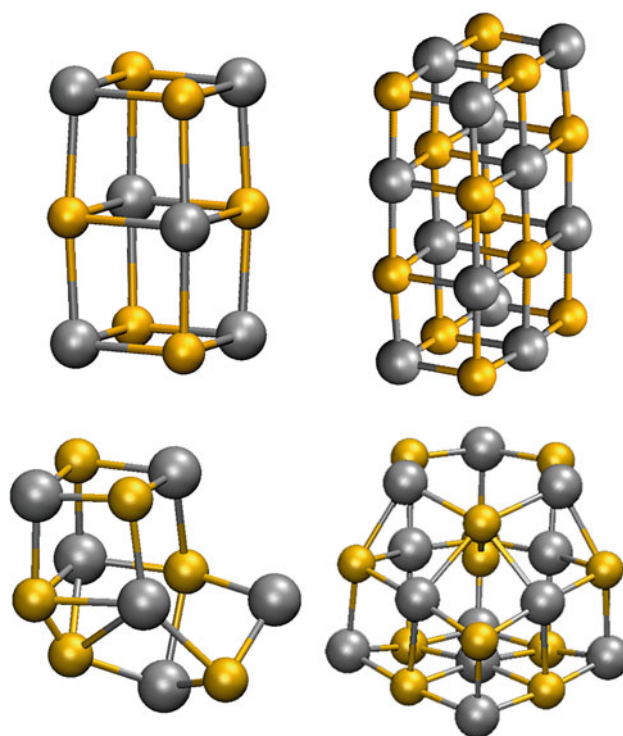**Fig. 1** Lowest energy structures for the $(TiC)_6$ and $(TiC)_{12}$ nanoclusters (*top*), and lowest energy C–C-containing $(TiC)_6$ and $(TiC)_{12}$ nanoclusters (*bottom*) as predicted from global optimization and refined by DFT calculations

Table 1 Smallest and largest TiC distances (Å) found in $(TiC)_6$ and $(TiC)_{12}$ clusters

TiC bonds	$(TiC)_6$	$(TiC)_{12}$
Smallest	1.937	1.899
Largest	2.072	2.195
Bulk calculated	2.161	2.161
Bulk experimental	2.164	2.164

Calculated and experimental TiC distances in the bulk rock salt structure are reported for comparison

structures for $(TiC)_6$ and $(TiC)_{12}$ are bulk-like and thus structurally unlike the corresponding ground states for MgO, most low-energy TiC nanocluster isomers are also found to be low in the energy for MgO. In other words, it is the order of nanocluster isomer stability that is different in each material.

Taking first the comparison between $(TiC)_6$ and $(MgO)_6$ clusters, in addition to the differences in the order of energetic stability, the different $(TiC)_6$ structures are more widely separated in energy (see Fig. 2). Whereas the $(TiC)_6$ isomers are fairly evenly separated throughout a 0.8 eV per TiC unit (eV/TiC) range, the $(MgO)_6$ isomers span a smaller energy interval (<0.6 eV/MgO) and are tightly grouped into two groups: (1) a low-lying pair of isomers

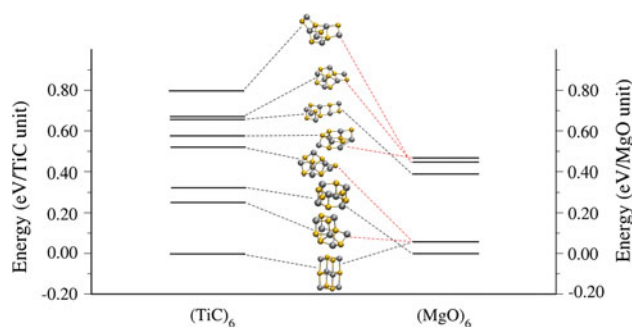**Fig. 2** Relative energies (in eV per TiC or MgO unit) corresponding to the low-energy isomer structures of $(TiC)_6$ and $(MgO)_6$ clusters. The *central part* of the figure shows the calculated structures which are linked to their respective relative energies via *dashed lines*. *Highlighted* by *red dashed lines* are those where structural differences between the two types of clusters are the largest (see Supplementary Information for more detail) (color figure online)

(the hexagonal drum ground state and a bulk cut corresponding to the $(TiC)_6$ ground state) separated by <0.1 eV/MgO and (2) the remaining isomers in a higher lying narrow (<0.2 eV/MgO) energy interval. Of particular note is that some structures obtained for $(TiC)_6$ after full DFT optimization were found to be not stable energy minima for the MgO system. This is the case, for instance, for isomers 2, 4, 5 and 8, where the numbering refers to the order of decreasing stability, in $(TiC)_6$. For $(TiC)_6$, isomers 2 and 5 are especially interesting because both exhibit a C–C bond which is not found to have a correspondence (i.e. a O–O bond) in any low-energy MgO clusters. Attempting to optimize $(TiC)_6$ isomer 2 as a $(MgO)_6$ isomer, for example, converts it to a relatively high-energy isomer with a bulk-like structure. This striking difference between the two materials at the nanoscale is particularly noteworthy as $(TiC)_6$ isomer 2 is only 0.252 eV/TiC higher in energy than
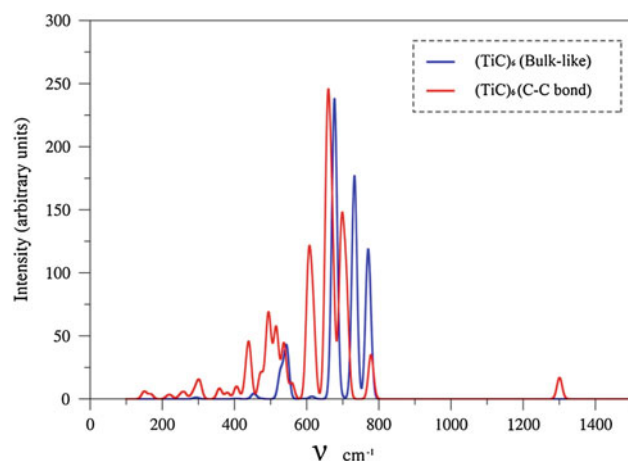
the ground state, and it is thus the second most stable cluster found for $(TiC)_6$ (see structure in Fig. 1). Although C–C bonds have not, to our knowledge, been reported previously in stoichiometric TiC clusters, their presence in C-rich non-stoichiometric TiC clusters (e.g. metcars) is well known and is thought to be energetically stabilising [54]. The appearance of the structures with C–C bonds prompted us to study the IR spectra that are reported in Fig. 3 for the bulk ground state cut isomer 1 and for isomer 2. Both calculated IR spectra are quite similar with respect to the position and intensity of their main peaks, except for a low intensity peak observed in the spectra of the C–C bonding isomer that appears at approximately 1,300 cm^{-1} which is due to to the C–C vibrations present in this isomer. Experimentally C-C vibrations in this frequency region have been used to indentify metcars while ruling our bulk-like stoichiometric TiC nanocrystals [55]. The calculated spectra should permit one to differentiate and assign our new stoichiometric, yet C-C-containing, structures from IR measurements. With respect to electronic structure, we also note that the gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) found for the $(TiC)_6$ ground state is 1.67 eV, while isomer 2 with a C–C bond shows a considerably smaller HOMO–LUMO gap of 0.56 eV.

In the case of $(TiC)_{12}$, one finds trends similar to those already described for the smaller $(TiC)_6$ cluster. Comparing the lowest energy structures obtained for the $(TiC)_{12}$ system with those of $(MgO)_{12}$, we also observe differences in the order of isomer stabilities (see Fig. 4). It is also interesting to remark that the structures for $(TiC)_{12}$ span a significantly larger energy range (0.85 eV/TiC) than in the

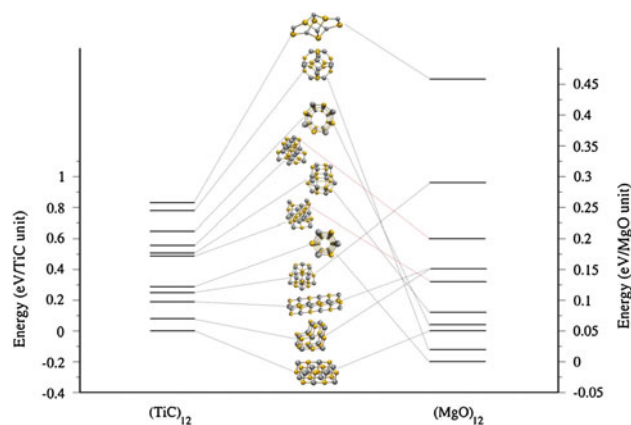**Fig. 3** Calculated IR spectra for two $(TiC)_6$ isomers: the bulk-like ground state isomer (*blue*) and the lowest energy isomer exhibiting C–C bonding (*red*) (color figure online)

Fig. 4 Relative energies (in eV per TiC or MgO unit) corresponding to the low-energy isomer structures of $(TiC)_{12}$ and $(MgO)_{12}$ clusters. The *central part* of the figure shows the calculated structures which are linked to their respective relative energies via *dashed lines*. *Highlighted* by *red dashed lines* are those where structural differences between the two types of clusters are the largest (see Supplementary Information for more detail) (color figure online)
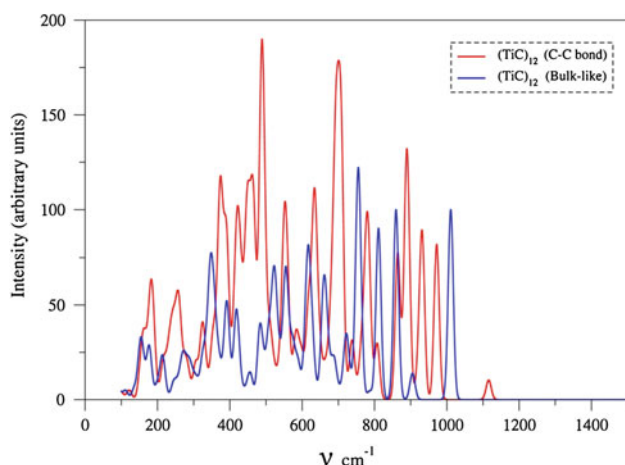
Fig. 5 Calculated IR spectra for two $(TiC)_{12}$ isomers: the bulk-like ground state isomer (*blue*) and the lowest energy isomer exhibiting C–C bonding (*red*) (color figure online)

case of the corresponding $(MgO)_{12}$ cluster (0.47 eV/MgO). Moreover, as for $(TiC)_6$, reasonably stable structures with C–C bonding also appear in the set of low-energy $(TiC)_{12}$ isomers. In particular, $(TiC)_{12}$ isomer 4 contains a C–C bond and is 0.25 eV/TiC higher in energy than the bulk-like-structured $(TiC)_{12}$ ground state (see structure in Fig. 1). Figure 5 presents the calculated IR spectra for the bulk-like ground state $(TiC)_{12}$ isomer and for $(TiC)_{12}$ isomer 4. Here, the spectra are much more complex than for the case of the smaller $(TiC)_6$ clusters. A direct comparison of both spectra is not straightforward, but, as in the case of $(TiC)_6$ cluster, there is a new peak in the spectrum of the isomer with a C–C bond at approximately 1,100 cm^{-1}. As in the previous case, this can be assigned to the C–C vibration. Finally, the $(TiC)_{12}$ HOMO–LUMO gap calculated for the ground state is 1.04 eV, which is smaller than in the case of the $(TiC)_6$ ground state and closer to the bulk, which shows metallic behaviour. The isomer with a C–C bond shows a significantly smaller gap of 0.11 eV, which is in line with the results discussed for $(TiC)_6$.

4 Conclusions

Based on global optimization with interatomic potentials and DFT-based calculations, we confirm that small stoichiometric TiC clusters tend to energetically prefer to exhibit a bulk-like structure. For the $(TiC)_6$ and $(TiC)_{12}$ clusters investigated here, this rock salt-like structure is between 0.1 and 0.3 eV/unit more stable than the next most stable isomers found. This is in clear contrast with some other materials which exhibit the rock salt structure in the bulk such as MgO, for which the lowest energy $(MgO)_6$ and $(MgO)_{12}$ clusters correspond to hexagonal tubular

structures. This difference in atomic structure at the nanoscale is likely to be, at least partially, related to the much more covalent character of TiC as compared to MgO, which is a prototypal ionic oxide material. This interpretation is supported by the observation that some of the low-energy isomers of both $(TiC)_6$ and $(TiC)_{12}$ exhibit clear C–C bonding which has no natural low-energy correspondence on MgO clusters. From the calculated IR spectra for the lowest energy isomers of both $(TiC)_6$ and $(TiC)_{12}$, and of the two lowest energy isomers exhibiting C–C bonding, we find distinguishing features which is due to the presence of C–C vibrations. These data may be used to differentiate between these two types of $(TiC)_n$ cluster in experiment.

References

1. Alivisatos AP (1996) Science 271:933–937
2. Daniel MC, Astruc D (2004) Chem Rev 104:293–346
3. Haruta M (1997) Catal Today 36:153–166
4. Kamat PV (2007) J Phys Chem C 111:2834–2860
5. Erwin SC, Zu LJ, Haftel MI, Efros AL, Kennedy TA, Norris DJ (2005) Nature 436:91–94
6. Calais JL (1977) Adv Phys 26:847–885
7. Schwartz K (1987) CRC Crit Rev Solid State Mater Sci 13:211–257
8. Sangthong W, Limtrakul J, Illas F, Bromley ST (2008) J Mater Chem 18:5871–5879
9. Berkdemir C, Castleman AW, Sofo JO (2012) Phys Chem Chem Phys 14:9642–9653
10. Catlow CRA, Bromley ST, Hamad S, Mora-Fonz M, Sokol AA, Woodley SM (2010) Phys Chem Chem Phys 12:786–811
11. Hartke B (2002) Angew Chem Int Ed 41:1468–1487
12. Bromley ST, Moreira I, Neyman KM, Illas F (2009) Chem Soc Rev 38:2657–2670
13. Levy RB, Boudart M (1973) Science 181:547–549
14. Hwu HH, Chen JGG (2005) Chem Rev 105:185–199
15. Viñes F, Rodriguez JA, Liu P, Illas F (2008) J Catal 260:103–112
16. Esposito DV, Chen JG (2011) Energy Environ Sci 4:3900–3912
17. Ren H, Hansgen DA, Kelly TG, Stottlemyer AL, Chen JG (2011) ACS Catal 1:390–398
18. Esposito DV, Hunt ST, Kimmel YC, Chen JG (2012) J Am Chem Soc 134:3025–3033
19. Hsu IJ, Kimmel YC, Willis BG, Chen JG (2012) Chem Comm 48:1063–1065
20. Ono LK, Sudfeld D, Cuenya BR (2006) Surf Sci 600:5041–5050
21. Rodriguez JA, Liu P, Viñes F, Illas F, Takahashi Y, Nakamura K (2008) Angew Chem Int Ed 47:6685–6689
22. Rodriguez JA, Liu P, Takahashi Y, Nakamura K, Viñes F, Illas F (2009) J Am Chem Soc 131:8595–8602
23. Rodriguez JA, Liu P, Takahashi Y, Nakamura K, Viñes F, Illas F (2010) Top Catal 52:393–402

24. Rodriguez JA, Feria L, Jirsak T, Takahashi Y, Nakamura K, Illas F (2010) J Am Chem Soc 132:3177–3186
25. Florez E, Gomez T, Liu P, Rodriguez JA, Illas F (2010) ChemCatChem 2:1219–1222
26. Florez E, Gomez T, Rodriguez JA, Illas F (2011) Phys Chem Chem Phys 13:6865–6871
27. Vidal AB, Feria L, Evans J, Takahashi Y, Liu P, Nakamura K, Illas F, Rodriguez JA (2012) J Phys Chem Lett 3:2275–2280
28. Guo BC, Wei S, Purnell J, Buzza SA, Castleman AW (1992) Science 256:515–516
29. Pillgrim JS, Duncan MA (1993) J Am Chem Soc 115:6858–6961
30. Wei S, Guo BC, Deng HT, Kerns K, Purnell J, Buzza SA, Castleman AW (1994) J Am Chem Soc 116:4475–4476
31. Duncan MA (1997) J Clust Sci 8:239–266
32. Rohmer MM, Bérnard M, Poblet JM (2000) Chem Rev 100:495–542
33. Patzschke M, Sundholm D (2005) J Phys Chem B 109:12503–12508
34. Li S, Wu H, Wang LS (1997) J Am Chem Soc 119:7417
35. Reddy BV, Khanna SN (1994) J Phys Chem 98:9946–9949
36. von Helden G, Tielens AGGM, van Heijnsbergen D, Duncan MA, Hony S, Waters LBFM, Meijer G (2000) Science 288:313–316
37. Wales DJ, Doye JPK (1997) J Phys Chem A 101:5111–5116
38. Carrasco J, Illas F, Bromley ST (2007) Phys Rev Lett 99:235502
39. Roberts C, Johnston RL (2001) Phys Chem Chem Phys 3:5024–5034
40. Erkoc S (1989) Phys Stat Sol (b) 152:447–454
41. Gale JD, Rohl AL (2003) Mol Simul 29:291–341
42. Kresse G, Hafner J (1993) Phys Rev B 47:558–561
43. Kresse G, Hafner J (1994) Phys Rev B 49:14251–14269
44. Kresse G, Furthmüller J (1996) Phys Rev B 54:11169–11186
45. Kresse G, Furthmüller J (1996) J Comp Mat Sci 6:15–50
46. Blöchl PE (1994) Phys Rev B 50:17953–17979
47. Kresse G, Joubert D (1999) Phys Rev B 59:1758–1775
48. Perdew JP, Wang Y (1992) Phys Rev B 45:13244–13249
49. Perdew JP, Wang Y (1993) Phys Rev B 48:4978
50. Valcárcel A, Ricart JM, Illas F, Clotet A (2004) J Phys Chem B 108:18297–18305
51. Happel M, Luckas N, Viñes F, Sobota M, Laurin M, Görling A, Libuda J (2011) J Phys Chem C 115:479–491
52. Lukas N, Viñes F, Happel M, Desikusumastuti A, Libuda J, Görling A (2010) J Phys Chem C 114:13813–13824
53. Bawa F, Panas I (2002) Phys Chem Chem Phys 4:103–108
54. Joswig J-O, Springborg M (2008) J Chem Phys 129:134311
55. van Heijnsbergen D, von Helden G, Duncan MA, van Roij AJA, Meijer G (1999) Phys Rev Lett 83:4983

A theoretical analysis of the magnetic properties of the low-dimensional copper(II)X_2(2-X-3-methylpyridine)$_2$ (X = Cl and Br) complexes

Sergi Vela · Mercè Deumal · Mark M. Turnbull · Juan J. Novoa

Abstract The First-principles Bottom-up (FPBU) procedure is applied to rationalize the different macroscopic magnetic properties of two compounds that were expected to be isostructural: bis(2-bromo-3-methylpyridine)dibromocopper(II), **1**, whose crystals present dominant ferromagnetic interactions, and bis(2-chloro-3-methylpyridine)dichlorocopper(II), **2**, that shows dominant antiferromagnetic behavior. Our FPBU analysis concludes that **1** presents a dominant ferromagnetic interaction of 1.16 cm^{-1} and other two non-negligible smaller interactions of opposite sign (−0.11 and 0.13 cm^{-1}). Contrarily, the dominant radical-pair interaction in **2** is antiferromagnetic (−2.37 cm^{-1}), in addition to three other non-negligible smaller magnetic couplings (0.48, −0.29, and −0.20 cm^{-1}). In **1**, these magnetic interactions generate a 2D magnetic topology of isolated planes, each made of weakly interacting parallel ferromagnetic chains, while in **2** they generate a 2D magnetic topology that can be described as isolated parallel double-decker planes, each of them made by weakly connected antiferromagnetic dimers.

The computed magnetic susceptibility curve that results after applying the FPBU procedure fully matches the experimental one in both systems. Furthermore, since in both systems, the weaker magnetic interactions are one order of magnitude smaller than the dominant coupling, the magnetic susceptibility curve does not vary significantly whether including all interactions or only the dominant ones. Thus, the FPBU analysis quantitatively traces down the origin of the different magnetic behavior of **1** and **2** as due to the change in sign of their dominant magnetic interactions. We have been able to connect such a change in nature of the dominant magnetic interaction with a change in the conformation of the ligands, which converts from *anti* in bis(2-bromo-3-methylpyridine) (**1**) to *syn* in bis(2-chloro-3-methylpyridine) (**2**), confirming the previous hypothesis.

S. Vela · M. Deumal · J. J. Novoa (✉)
Departament de Química Física and IQTCUB,
Facultat de Química, Universitat de Barcelona,
Martí i Franquès 1, 08028 Barcelona, Spain
e-mail: juan.novoa@ub.edu

M. M. Turnbull
Carlson School of Chemistry and Biochemistry,
Clark University, 950 Main St., Worcester,
MA 01610, USA

1 Introduction

Low-dimensional molecule-based magnets, materials where the observed magnetic exchange is limited to less than three dimensions over a broad range of temperatures, have become a major source of interest for the study of viable super exchange pathways and magneto-structural correlations. The advantage to the use of low-dimensional materials for magnetic study is clear; fewer interactions makes them easier to define and provides fewer parameters that affect the sign and magnitude of the magnetic interactions. However, for such studies to be effective, the true nature of the magnetic lattice must be known so that all

non-negligible interactions are taken into account. Low-dimensional materials are also interesting for the possible presence of new physical phenomena associated with their reduced dimensionality.

The problem with many studies of low-dimensional magnets is that the dimensionality of the magnetic lattice is generally assumed from the physical lattice (obtained from X-ray or neutron diffraction studies) and these may, or may not, correlate with the true dimensionality determined from more accurate and expensive physical measurements (e.g. muon spin rotation determinations). One prime example is that of $(VO)_2(P_2O_7)$ which was studied in detail as a magnetic ladder for nearly a decade [1–5] before neutron scattering experiments showed the material to be better described as an alternating magnetic chain [6, 7]. Magnetic dimensionality can also be accurately determined from accurate theoretical studies of the magnetic interactions within a crystal, as we show hereafter in the case of bis(2-bromo-3-methylpyridine)dibromocopper(II) (1) and bis(2-chloro-3-methylpyridine)dichlorocopper(II) (2) (see Fig. 1).

Despite the similar crystal structure of 1 and 2, they exhibit very different macroscopic magnetic properties. As observed in experimental studies, [8] 1 is a molecule-based magnet whose magnetic susceptibility is fitted using a ferromagnetic chain model with weak antiferromagnetic interchain corrections, while the magnetic data of 2 is fitted to an isolated antiferromagnetic dimer model. It has been postulated that the reason for such a different magnetic behavior is the change in conformation that the $Cu(II)X_2(2-X-3-Mepy)_2$ (X = Cl, Br; Mepy = methylpyridine) radical centers show in 1 and 2. Specifically, while in 1 all centers are present in the *anti*-conformation, in 2 all are found in the *syn*-conformation (see Fig. 1). However, this proposal has not yet been supported by theoretical studies on the magnetism of these two compounds. Our main aim in this work is to confirm the validity of such hypothesis by applying the First-principles Bottom-up (FPBU) procedure [9] to 1 and 2 in order to gain an in-depth insight into the nature of their magnetic interactions and, by comparing them, rationalize the origin of their difference.

The FPBU procedure computes the macroscopic magnetic properties of any molecule-based compound based only upon knowledge of the crystal structure. It begins by evaluating, using First-Principles methods (i.e. using high-level ab initio [10] or DFT [11] methods), the microscopic J_{AB} magnetic exchange parameters for all symmetry-unique radical-pairs in the crystal (note that these parameters uniquely define the Heisenberg Hamiltonian that describes the magnetic interactions within the crystal, in a pair approximation). The matrix representation of the Heisenberg Hamiltonian is then computed on a properly selected finite subset of the crystal (called the magnetic
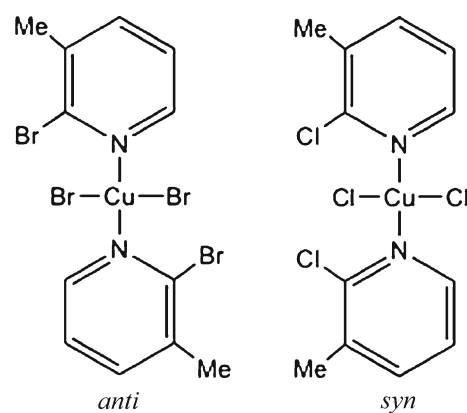

Fig. 1 Chemical structure of the *anti*-bis(2-bromo-3-methylpyridine)dibromocopper(II) (1) and *syn*-bis(2-chloro-3-methylpyridine) dichlorocopper(II) (2) conformers

model). Within a regionally reduced density matrix approach, the corresponding eigenvalues are obtained after diagonalization and used in the appropriate Statistical Mechanics expression of the property of interest, in this case the magnetic susceptibility, in order to evaluate such property in an unbiased and accurate form. The FPBU methodology has been previously shown to reproduce well the magnetic properties of molecule-based magnets presenting a wide variety of magnetic behaviors [12–16]. It also allows one to connect the macroscopic magnetic property of interest with its microscopic origin (i.e. the values of the J_{AB} magnetic couplings and the network of connections that the non-negligible J_{AB} interactions create among the radicals of the crystal, known as the *magnetic topology* of the crystal). The procedure is called *bottom-up* because the macroscopic magnetic properties are determined from the computed values of the microscopic radical···radical magnetic interactions, which are calculated using *first-principles* methods (no "a priory" assumptions are made about the size or topology of the magnetic interactions present in the crystal). In this paper, the FPBU procedure is used to identify in a numerical and unbiased way the origin of the different magnetic behavior of compounds 1 and 2.

2 Methodological details

The *First-principles Bottom-up* FPBU procedure [9] is a four-step methodology that computes the macroscopic magnetic properties of a molecule-based crystalline material using its crystal structure as the only input. The four steps of the FPBU procedure can be summarized as follows (they are valid for any crystal, although in the description below they have been adapted to the case of compounds 1 and 2):

1. *Identification of all unique radical···radical pairs within the crystal likely to be magnetically active.* All symmetry-unique radical-pairs (d_i), whose radical···radical distance is smaller than a given threshold, are identified by doing an in-depth analysis of the crystal. The threshold is selected in such a way that all relevant first and second nearest neighbor radical···radical pairs are included. No differences arise if the selected radical···radical magnetic interactions are through-bond or through-space [16].

2. *Calculation of the radical···radical magnetic interactions* (J_{AB}) *for all unique pairs found in Step 1.* The $Cu(II)X_2(2\text{-}X\text{-}3\text{-Mepy})_2$ radicals present a doublet ground state. Thus, the strength of the magnetic interaction J_{AB} between any potentially relevant radical-pair d_i in a crystal can be calculated as $J_{AB} = (E_{BS}^S - E^T)$,[1] where E^T is the energy of the triplet, and E_{BS}^S, the energy of the open-shell singlet computed using the broken-symmetry approximation [17, 18]. The general Heisenberg Hamiltonian for a pair of A and B radicals

$$\hat{H} = -2 \sum_{A,B} J_{AB} \hat{S}_A \cdot \hat{S}_B \qquad (1)$$

has been used, where $\hat{S}_A$ and $\hat{S}_B$ are the total spin operators acting on radicals A and B of each radical-pair. The radical-pair energies are computed at their crystal geometry. When available, it is recommended to use a crystal structure determined at a low temperature, where possible anisotropic thermal effects are minimized [19].

Note that the previous J_{AB} expression assumes that the overlap S_{ab} between the singly occupied molecular orbitals SOMOs of radicals A and B is small.[2] This hypothesis is valid in most through-space interactions, and also in some through-bond magnetic interactions (for instance, when the radical containing atoms are not directly bonded) [16]. The E_{BS}^S and E^T energies were evaluated using the UB3LYP [20–23] DFT functional, the Ahlrich's DZP basis set [24] for Cu, and the $6-31+G(d)$ basis set [25] on the remaining atoms. The choice of the B3LYP functional is based on previous studies that showed its performance to reproduce the experimental J_{AB} values and those from high-level ab initio methods on properly characterized

systems [12–16]. The convergence criterion of the energy values was forced to have an accuracy of 10^{-7} au in order to guarantee an accuracy of 0.04 cm^{-1} in the J_{AB}'s values. All DFT calculations were carried out using GAUSSIAN09 [26].

Let us further comment on the information held by the SOMOs. In crystals of **1** and **2**, the magnetic interactions formally originate in the overlap between the SOMOs hosting the unpaired d electron in the tetra-coordinated Cu(II) ions of the $Cu(II)X_2(2\text{-}X\text{-}3\text{-Mepy})_2$ (X = Cl in **1** and X = Br in **2**) radicals (Fig. 2a). However, UB3LYP calculations indicate that in both radicals their SOMO, and therefore spin density, spreads over the Cu-coordinated halide atoms and also onto the Mepy nitrogen atoms bonded to the Cu(II) atom (see Fig. 2b, c). As magnetic interactions originate in the overlap between the occupied orbitals, and in particular the SOMO orbitals, the shape of the SOMO in isolated radicals of **1** and **2** suggests that the strongest magnetic pathways between their radical-pairs are expected to be those involving short-distance Cu···halide interactions, halide···halide interactions, and also (Mepy)N···N(Mepy) interactions.

3. *Determination of the magnetic topology of the crystal and selection of the appropriate model space.* The magnetic topology is straightforwardly defined by the network of connectivities among the spin centers linked by non-negligible $J_{AB}(d_i)$ interactions (previous tests have shown that when $|J_{AB}(d_i)| < 0.05$ cm^{-1}, the magnetic interaction can be considered as negligible). Once the magnetic topology of the full crystal is known, magnetic models can be selected, namely subsets of radicals whose propagation along the crystallographic axes reproduces the magnetic topology of the infinite crystal and includes all non-negligible J_{AB} magnetic interactions in a ratio as close as possible to that found in the infinite crystal. The smallest of those finite magnetic models is the minimal magnetic model. When the minimal magnetic model is properly defined, the macroscopic properties computed by enlarging it converge smoothly toward the computed data obtained using the minimal model and, in turn, toward the experimental results.

4. *Calculation of the macroscopic magnetic properties of the crystal.* In Step 2 of the FPBU procedure, from each two $S = ½$ radicals, the energies and therefore the magnetic interaction parameter (the J_{AB}) of the effective Hamiltonian are calculated. Then, all the values (i.e. energy eigenvalues and S quantum numbers) are mapped on the corresponding matrix elements of the model Hamiltonian, whose model space is selected in Step 3. In the present simple case, within the framework of a regionally reduced density matrix

---

[1] The criterion chosen to compute the energy difference is $E^S - E^T = 2(E_{BS}^S - E^T)/(1 + S_{ab})$. Open-shell singlet systems separate alpha spin density and beta spin density on different radicals. In our case, once the broken-symmetry approximation is applied, the resulting overlap S_{ab} between the alpha SOMO and beta SOMO is zero. Thus, those orbitals are localized on each of the two radicals. This leads to $S_{ab} = 0$. As a conclusion, $J_{AB} = E_{BS}^S - E^T$.

[2] See footnote one.

Fig. 2 a Molecular units of **1** (*left*) and **2** (*right*). **b** SOMO orbitals of isolated radicals of **1** and **2** (the isosurface of 0.03 au has been plotted; *blue* and *white* correspond to positive and negative regions, respectively). **c** Spin densities of monomers of **1** and **2** (the isosurface of 0.008 au has been plotted). *Color code*: Cu (*cyan*), N (*blue*), C (*black*), H (*pink*), Cl (*green*), Br (*brown*)
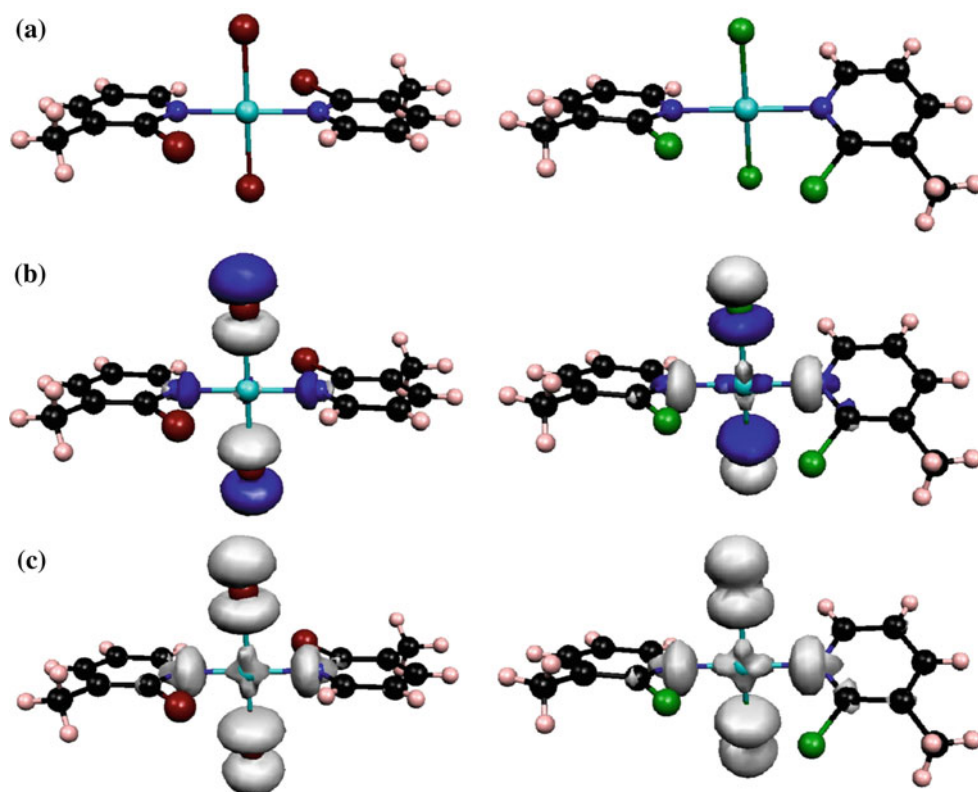
approach, the correspondence between model Hamiltonian and effective Hamiltonian is done by mapping the energies of the determinants on the diagonal elements of the Heisenberg Hamiltonian.[3] The energy eigenvalues and S quantum numbers obtained after full diagonalization of the matrix representation of (1) are then used to evaluate the magnetic properties of interest using the required statistical mechanics expressions (magnetization, heat capacity, magnetic susceptibility, etc.).

Note that the spectrum of the full crystal would require including all radical-pairs in the summation of Eq. 1, which is computationally impractical. Instead, it has been demonstrated [9] that a proper reproduction of the energy spectra is obtained by using a properly defined minimal magnetic model. If the radical is a doublet, as in compounds **1** and **2**, the size of the matrix representation increases with the number n of radicals in the model space as $n!/[(n/2)!(n/2)!]$. In practice, this means that we are limited to model spaces of 18 spin centers or fewer.

---

[3] Note that the model Hamiltonian in the context of magnetic interactions is the HDVV spin-only Hamiltonian, while the effective Hamiltonian is a projection onto an appropriate model space of calculations from the exact Hamiltonian.
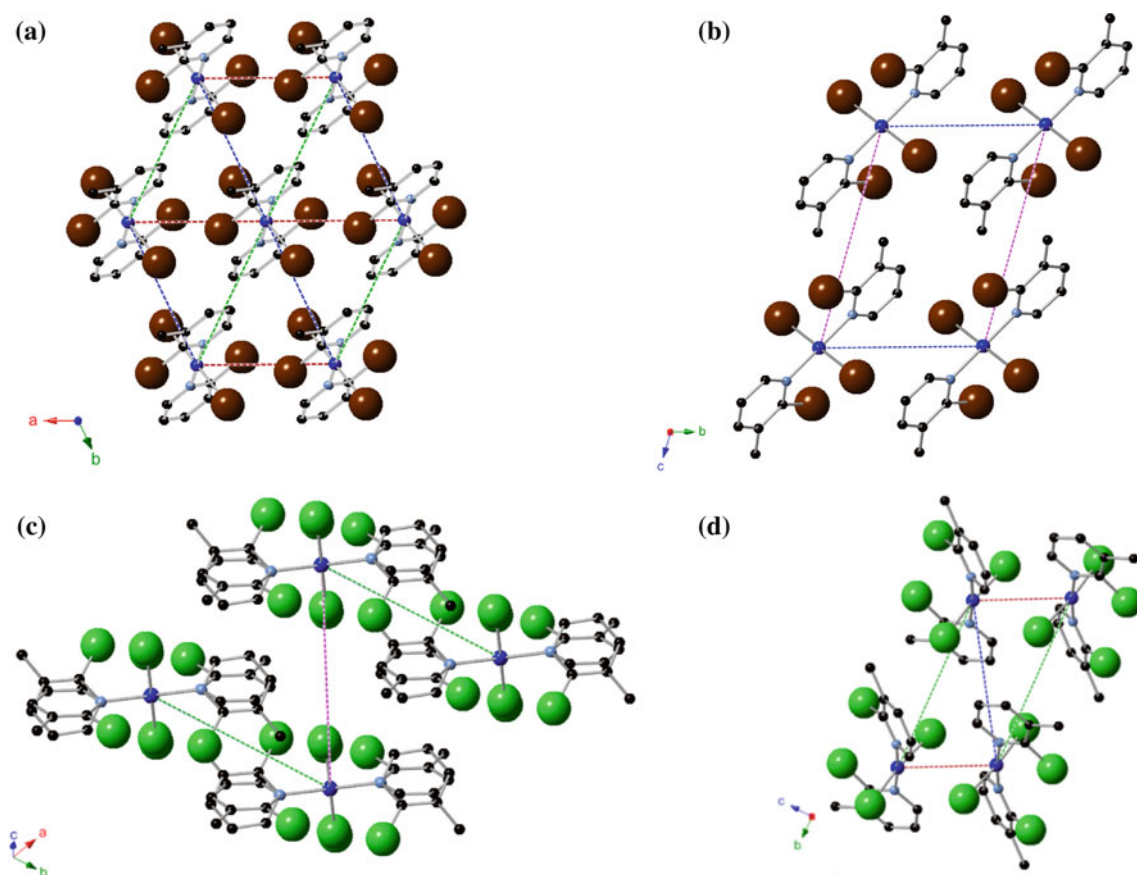
3 Results and discussion

3.1 Crystal packing analysis of the crystals of **1** and **2**

The FPBU studies on crystals of **1** and **2** were done using the X-ray structures reported in the literature [8]. Both crystal structures have been obtained at low enough temperatures (120 K for **1** and 165 K for **2**) to expect the anisotropic contraction of the crystal to be minimized. This fact guarantees that the geometry of any radical-pair extracted from these crystal structures will be similar to the geometry for that pair at the low temperatures where magnetic collective phenomena are most evident (local maxima/minima) in most molecule-based magnets.

Both crystals belong to the P$\bar{1}$ space group and present similar cell parameters. The cell parameters for crystals of **1** are as follows: $a = 6.2440(2)$ Å, $b = 7.4588(3)$ Å, $c = 9.5897(4)$ Å, $\alpha = 104.777(2)°$, $\beta = 90.043(2)°$, $\gamma = 114.151(2)°$, $V = 391.24(3)$ Å^3. For crystals of **2** they are as follows: $a = 6.0949(4)$ Å, $b = 7.4718(4)$ Å, $c = 9.5654(5)$ Å, $\alpha = 104.579(2)°$, $\beta = 91.809(2)°$, $\gamma = 112.825(2)°$, $V = 384.47(6)$Å^3. However, as shown in Figs. 3 and 4, the relative arrangement of the radical molecules in each crystal is quite different. These differences are better appreciated in Fig. 4, when looking at their superstructure, that is, at the general arrangement of the radicals within the crystal, where each radical is only represented by its Cu atom. As observed in Fig. 4, the radicals

Fig. 3 For **1**, **a** top *ab*- and **b** lateral *bc*-views of the geometry of d_1–d_4 radical-pairs selected in Step 1 (see *dashed lines*). For **2**, **c** view along the $\alpha = 45°$, $\beta = 45°$, $\gamma = 45°$ direction and **d** lateral *bc*-view

of the geometry of d_1–d_4 radical-pairs (see *dashed lines*). Hydrogen atoms have been removed for clarity. *Color code*: Cu (*deep-blue*), N (*blue*), C (*black*), Cl (*green*), Br (*brown*)

in **1** pack as stacks of flat planes, while in **2** they pack as stacks of double-decker corrugated planes, where each double-decker plane results from the aggregation of dimers.

3.2 First-principles bottom-up analysis of **1** and **2**

The results of applying the FPBU methodology to the study of the magnetic interactions in **1** and **2** are hereafter presented, grouped according to the four steps of the procedure.

Step 1 Identification of all unique radical⋯radical pairs from the crystal that are likely to be magnetically active.

All radical-pairs having a Cu⋯Cu distance shorter than 10 Å in both crystals were considered in **1** and **2** (the selected radical-pairs are numbered according to their Cu⋯Cu distance, d_1 being the dimer with the shortest distance). For **1**, four unique radical-pairs were found with Cu⋯Cu distances that range from 6.244 Å to 9.590 Å (see Fig. 5). For **2**, eleven unique radical-pairs were selected (see Supporting Information Figure S1 and Fig. 6 for magnetically relevant pairs). Their shortest Cu⋯Cu distance is 4.329 Å (d_1), while d_2–d_{11} present Cu⋯Cu distances ranging from 8 to 10 Å.

Step 2 Computation of the microscopic magnetic interactions, $J_{AB}(d_i)$, for all symmetry-unique radical-pairs selected in Step 1.

Calculation of the $J_{AB}(d_i)$ values done at the UB3LYP level (Table 1) indicates that **1** has only three magnetically non-negligible radical-pairs (i.e. $|J_{AB}(d_i)| > 0.05$ cm^{-1}): d_1, d_2, and d_3. The dominant radical-pair magnetic interaction is $J(d_2)$ (1.16 cm^{-1}), which is about one order of magnitude larger than $J(d_1)$ and $J(d_3)$ (−0.11 and 0.13 cm^{-1}, respectively). Note that our computed J_{AB} exchange coupling values are close to the experimental one, $J_{chain} = 0.89$ cm^{-1}, obtained by fitting the magnetic susceptibility curve with a ferromagnetic chain model with a correction term that was added to account for the weak interchain interactions (best fit value J'(interchain) = −0.25 cm^{-1})[4] [8]. As observed in Fig. 7a, the dominant magnetic interaction (d_2) is a through-space Cu–Br⋯Br–Cu interaction.

---

[4] Note that $J_{chain} = 0.89$ cm^{-1} corresponds to an experimentally fitted 2J parameter of 2.58 K. The J'(interchain) value has been translated from the fitted Curie–Weiss mean-field parameter theta (−0.74 cm^{-1}) assuming two neighbor radicals.

Fig. 4 Superstructure of **a 1** and **b 2**. Each radical is represented by its central Cu atom. The perfectly collinear chain in **1** (*red-dashed lines* in **a**) becomes corrugated in **2** and also non-regular since there is now a short Cu···Cu and a long Cu···Cu contact (*red-* and *black-dashed lines* in **b**, respectively)
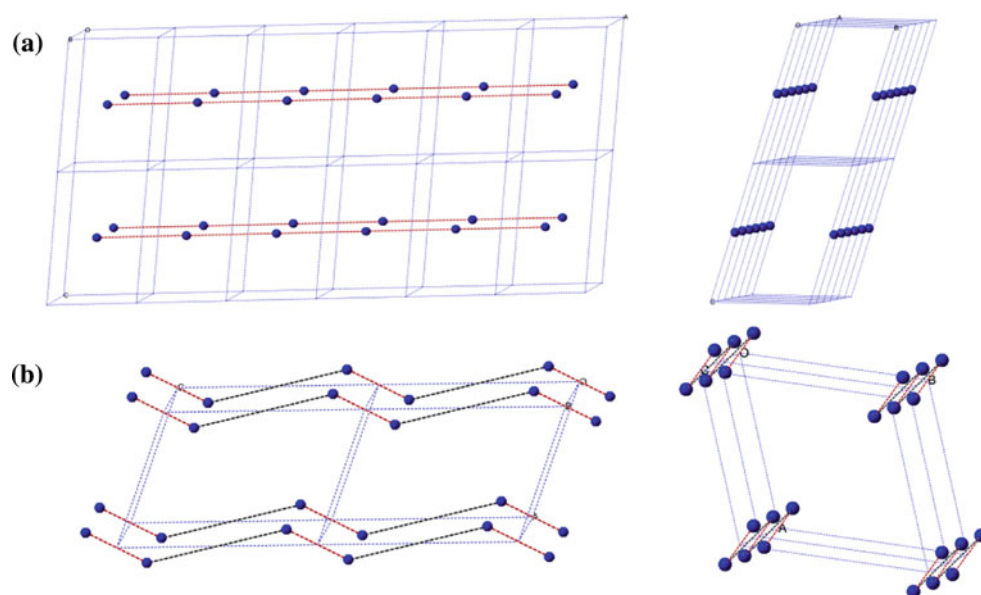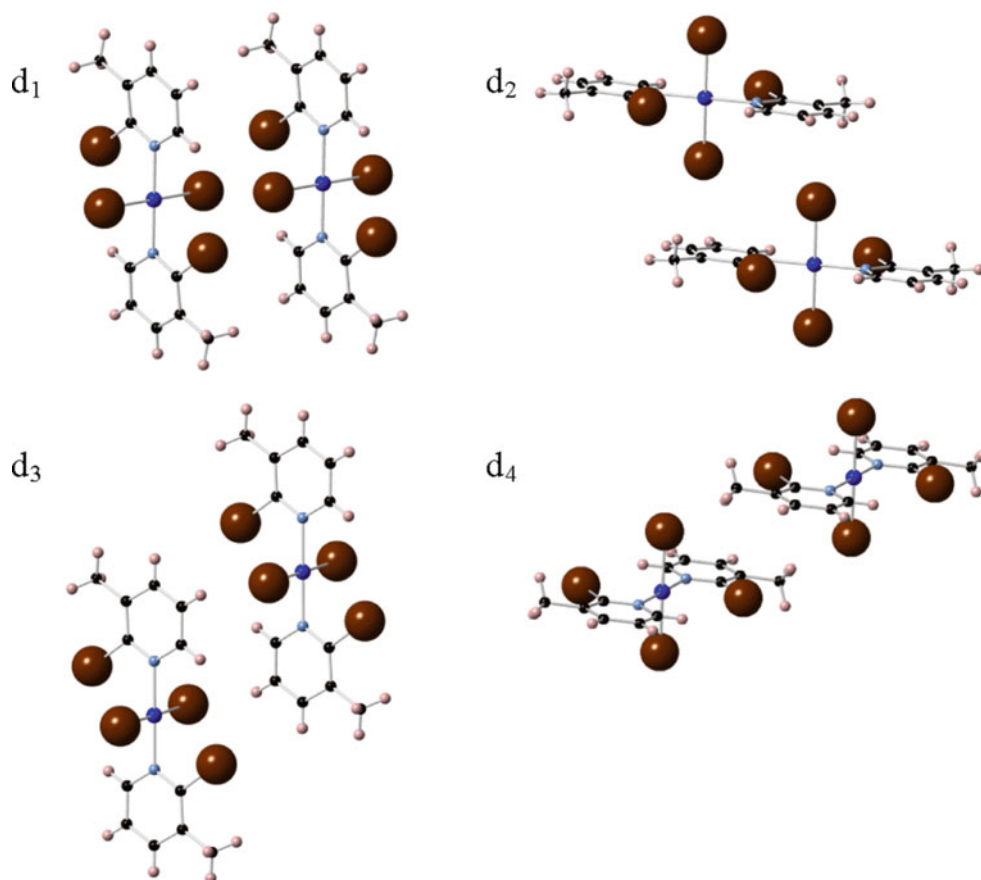

Fig. 5 Unique radical-pairs that present a Cu···Cu distance smaller than 10 Å in **1**

The values of the $J_{AB}(d_i)$ interactions computed at the UB3LYP level for the magnetically relevant radical-pairs in **2** are collected in Table 2 (see Fig. 6 for geometry of pairs). They show the presence of a dominant antiferromagnetic interaction, $(J(d_1) = -2.37 \text{ cm}^{-1})$, which is almost one order of magnitude larger than the three remaining non-negligible magnetic couplings (0.48, −0.29, −0.20 cm^{-1}). These computed $J_{AB}(d_i)$ values compare well with those obtained by fitting the magnetic susceptibility curve of **2** with an isolated dimer model, $J_{dimer} =$

Fig. 6 Unique magnetically non-negligible radical-pairs that present a Cu···Cu distance smaller than 10 Å in **2**
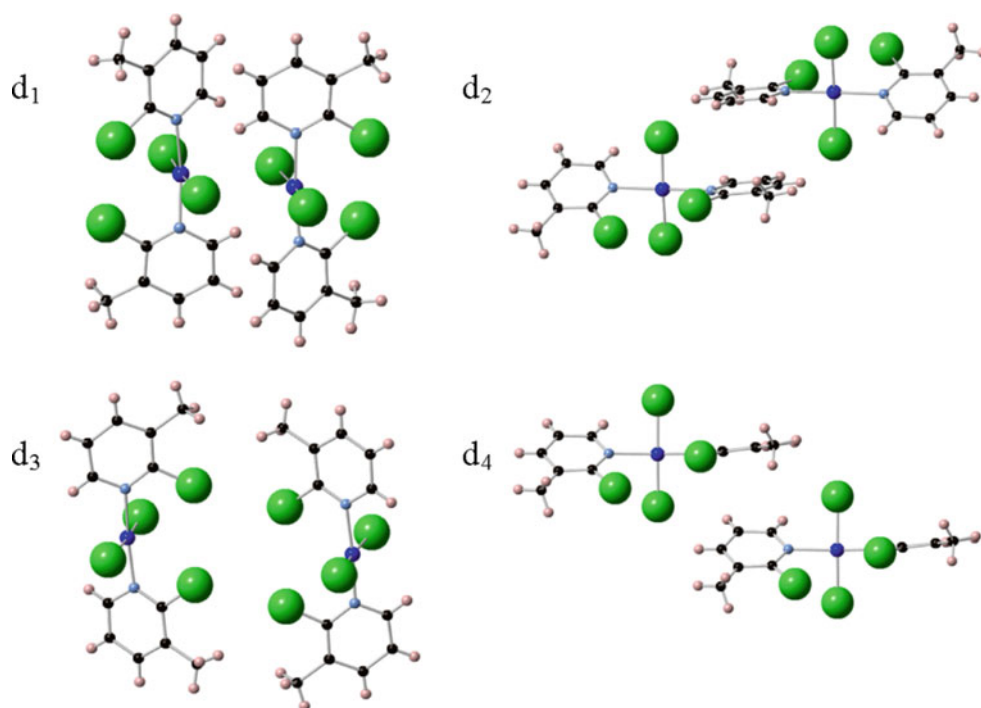

Table 1 Values of the non-negligible magnetic exchange interaction parameter $J_{AB}(d_i)$ computed at the UB3LYP level for all unique radical-pairs present in **1**

Dimer	$d(Cu···Cu)$/Å	$d(Br···Br)$/Å	J_{AB}/cm^{-1}
d_1	6.244	5.253	−0.11
d_2	7.459	4.482	1.16
d_3	7.517	7.089	0.13

The values of the Cu···Cu and the shortest Cu–Br···Br–Cu distances are also given (in Å)

−2.72 cm^{-1} [8][5]. Let us further comment that, according to Fig. 7b, the dominant magnetic coupling (d_1) is a through-space magnetic interaction via a series of exchange pathways, ranging from Cu···Cu to Cu···Cl and Cl···Cl magnetic contacts.

The dominant ferromagnetic exchange observed in compound **1** is unusual as virtually all examples of magnetic superexchange via a two-halide pathway are antiferromagnetic. The reason for that could be the atypical orientation found for Br···Br interactions in **1** (see Fig. 7c, d). In **1**, the Cu–Br···Br angle of the dominant ferromagnetic radical-pair d_2 for this pathway is 106.4°, that is, close to 90°, which prevents any exchange coupling other than Br···Br. Whereas, in **2**, the Cu–Cl···Cl angle in the dominant antiferromagnetic pair d_1 is 62.8°, whose orientation allows Cu···Cu, Cu···Cl, and Cl···Cl magnetic

exchange. The existence of regions of ferromagnetism in the Cu(II)-halide···halide-Cu(II) interactions present in crystals of Cu(II)X$_2$L$_2$ radicals studied here, suggest the need of a detailed study of their magneto-structural properties, but such study it is out of the scope of this paper.

Step 3 Determination of the magnetic topology of the crystal and selection of the appropriate magnetic model.

Figure 8 shows the 2D magnetic topology for the crystals of **1** and **2**. According to Fig. 8a, the network of magnetic interactions in **1** consists of a set of isolated ferromagnetic chains ($J(\mathbf{1}d_2) = 1.16$ cm^{-1}, blue solid lines), which pack forming planes of parallel chains ($J(\mathbf{1}d_1) = -0.11$ and $J(\mathbf{1}d_3) = 0.13$ cm^{-1}, red and green solid lines, respectively). Finally, the planes of chains stack in the third dimension with no magnetic linkages (see Fig. 8a). On the other hand, the magnetic topology of **2** (see Fig. 8b) consists of antiferromagnetic dimers ($J(\mathbf{2}d_1) = -2.37$ cm^{-1} in red solid lines), which then weakly interact to give rise to a stack of magnetically isolated double-decker planes ($J(\mathbf{2}d_2) = -0.29$, $J(\mathbf{2}d_3) = 0.48$ and $J(\mathbf{2}d_4) = -0.20$ cm^{-1} in blue, green, and purple solid lines, respectively). All these results confirm the expectations put forward on the basis of geometrical considerations [8], namely (1) the strongest interactions of **1**, $J(\mathbf{1}d_2)$, are ferromagnetic and run along a chain motif, (2) adjacent chains in **1** are weakly interconnected, although the interchain interactions are not just antiferromagnetic, but both antiferromagnetic $J(\mathbf{1}d_1)$ and ferromagnetic $J(\mathbf{1}d_3)$ in nature, and (3) the strongest

---

[5] Note that $J_{dimer} = -2.72$ cm^{-1} corresponds to an experimentally fitted 2J parameter of ca. −7.824 K.

Fig. 7 Spin density of **a** d_2 in **1** and **b** d_1 in **2**. Local geometry of the (X–Cu–X)···(X–Cu–X) moiety for **c** d_2 in **1** and **d** d_1 in **2**. *Color code*: Cu (*deep-blue*), N (*blue*), C (*black*), Cl (*green*), Br (*brown*)
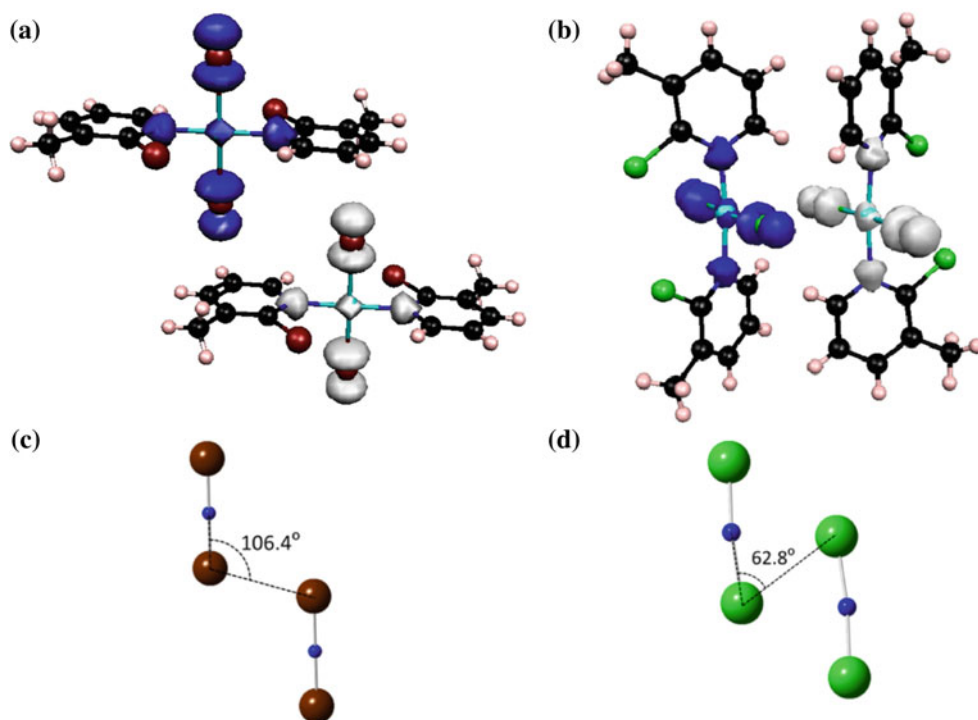

Table 2 Values of the non-negligible magnetic exchange interaction parameter $J_{AB}(d_i)$ computed at the UB3LYP level on all unique radical-pairs present in **2**

Dimer	$d(Cu\cdots Cu)/Å$	$d(Cl\cdots Cl)/Å$	J_{AB}/cm^{-1}
d_1	4.329	3.652	−2.37
d_2	8.071	7.559	−0.29
d_3	8.680	6.762	0.48
d_4	9.519	5.346	−0.20

The values of the Cu···Cu and shortest Cu–Cl···Cl–Cu distances are also given (in Å)

coupling in **2**, $J(\mathbf{2}d_1)$, yields antiferromagnetic dimers, although there are interdimer interactions which are both ferro- ($J(\mathbf{2}d_3)$) and antiferromagnetic ($J(\mathbf{2}d_2)$ and $J(\mathbf{2}d_4)$) that the geometrical considerations did not account for.

Step 4 Calculation of the macroscopic magnetic properties of the crystal.

The macroscopic magnetic susceptibility of **1** and **2** can now be computed from the energy spectrum of the Heisenberg Hamiltonian acting on a given magnetic model space, after substituting the $J_{AB}(d_i)$ interactions by their values in Tables 1 and 2, respectively. It thus follows that one first has to select a proper model space. Analysis of the magnetic topologies of **1** and **2** allows a proper selection of adequate magnetic models (Fig. 9). At this point, let us recall that the larger the magnetic model, the better the agreement between computed and experimental data is

expected to be, but also at the higher computational cost of evaluating the macroscopic properties. The magnetic models were also selected keeping in mind that the ratio of J_i/J_j in the minimal model space should be as close as possible to that found in the full crystal, whatever the i and j radical-pair.

For compound **1**, four magnetic models were selected. The first one was the 2×(2×4) model (Fig. 9a), representing two (2×4)-radical centers from two adjacent planes. The second one, the 1×(2×4) model (Fig. 9b), was chosen to show that the magnetic susceptibility calculated on two planes (2×(2×4) model) is equivalent to that calculated for just one of those planes since they are magnetically isolated, at a reduced computational cost. Finally, the third is a 1×(2×8) model (Fig. 9c), chosen to test the convergence along the crystallographic directions along which the dominant magnetic interactions propagate. Note that this last model allows comparison with the literature data since the fitting model used experimentally was a chain model J(chain) that included a Curie–Weiss-term to account for weak interchain interactions J'(interchain).

Figure 9d–g show the magnetic models selected for compound **2**. Figure 9d shows the minimal magnetic model, the 2×2(2×2) model, which contains 16 spin centers and reproduces, by expansion, the double-decker topology of the full crystal. Similarly to the process followed in compound **1**, we included the 1×2(2×2) model (Fig. 9e) to test whether using a single double-decker model is appropriate or not as no magnetic interactions have been computed between adjacent double-decker

Fig. 8 Computed magnetic topologies for compounds **a 1** and **b 2**. *Color code* (common to **a** and **b**): $J(d_1)$ in *red*, $J(d_2)$ in *blue*, $J(d_3)$ in *green*, $J(d_4)$ in *purple*
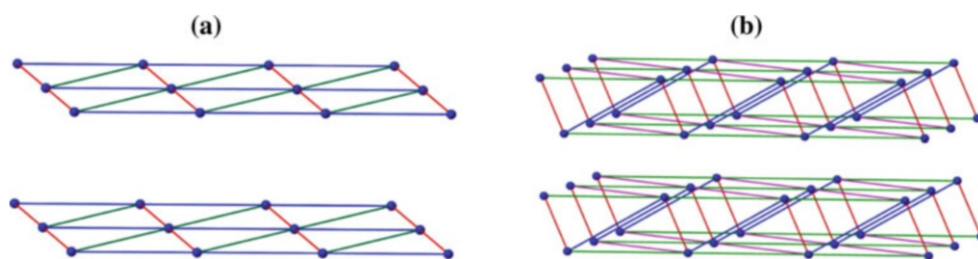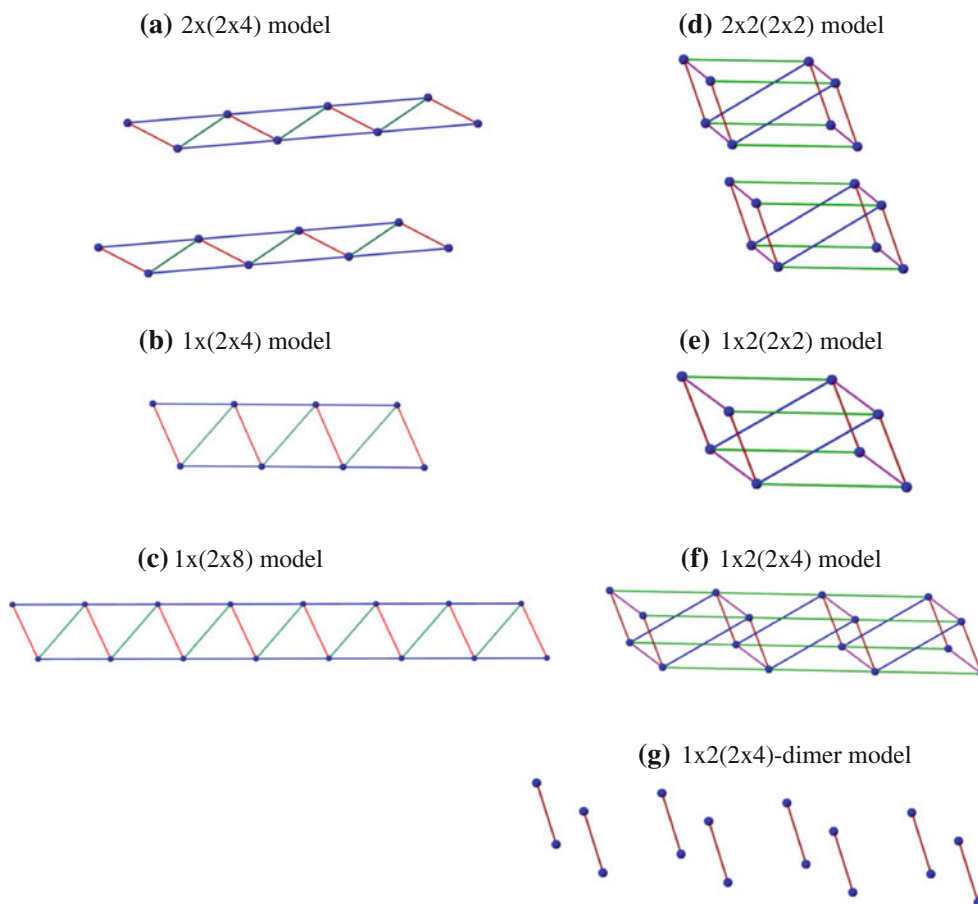

Fig. 9 Magnetic models used to compute the magnetic susceptibility *curves* for **1** (**a–c**) and **2** (**d–g**). *Blue dots* represent the Cu atoms of each radical. *Red*, *blue*, *green* and *purple* lines (see Fig. 8 for *color code*) represent the magnetic exchange interactions $J_{AB}(d_i)$ computed in *Step 2*

planes. The third model studied is the $1\times2(2\times4)$ (Fig. 9f), aimed at getting convergence in each of the planes along one direction. Finally, we included the $1\times2(2\times4)$-dimer model (Fig. 9g) that describes a set of non-interacting dimers, which is the model employed in the literature to fit the experimental J_{dimer} values (this is equivalent to setting the values for $J(d_2)$, $J(d_3)$ and $J(d_4)$ equal to zero).

The magnetic susceptibility $\chi(T)$ curves computed for **1** and **2** are shown in Fig. 10. For **1**, all models agree well with the experimental curve, and the computed $\chi(T)$ data virtually overlap the experimentally measured values. This is consistent with the fact that all models contain the dominant exchange parameter, $J(d_2)$, in exactly the same proportion relative to the crystal as a whole. In other words, the key information about the energy spectrum is

determined by $J(d_2)$ and all remaining non-negligible $J_{AB}(d_i)$ only induce a small perturbation in that spectrum. The results also allow us to conclude that the ferromagnetic character of **1** originates in the ferromagnetic interaction along the two-halide bridge d_2 radical-pair.

A similar trend is observed for the magnetic susceptibility curves of compound **2**. Once again, the $\chi(T)$ curves computed with the $2\times2\times(2\times2)$, $1\times2\times(2\times2)$, $1\times2\times(2\times4)$, and $1\times2\times(2\times4)$-dimer models nearly numerically reproduce the experimental curve. These results allow us to conclude that the antiferromagnetic character of **2** originates in the antiferromagnetic intradimer interaction $J(d_1)$. Once again, the contributions of the additional interdimer interactions are important to describe the magnetic topology of the molecule-based crystal but numerically
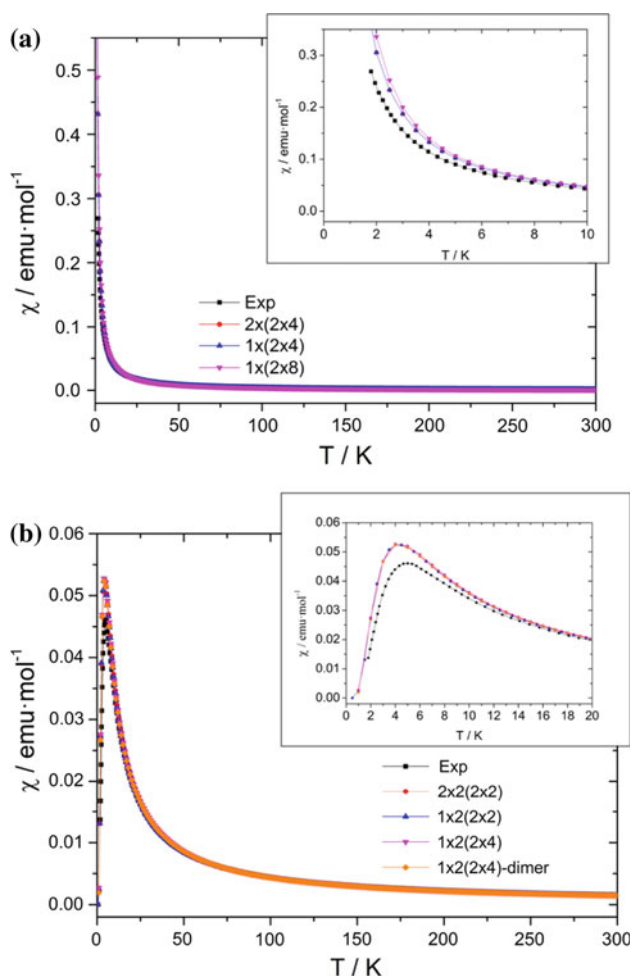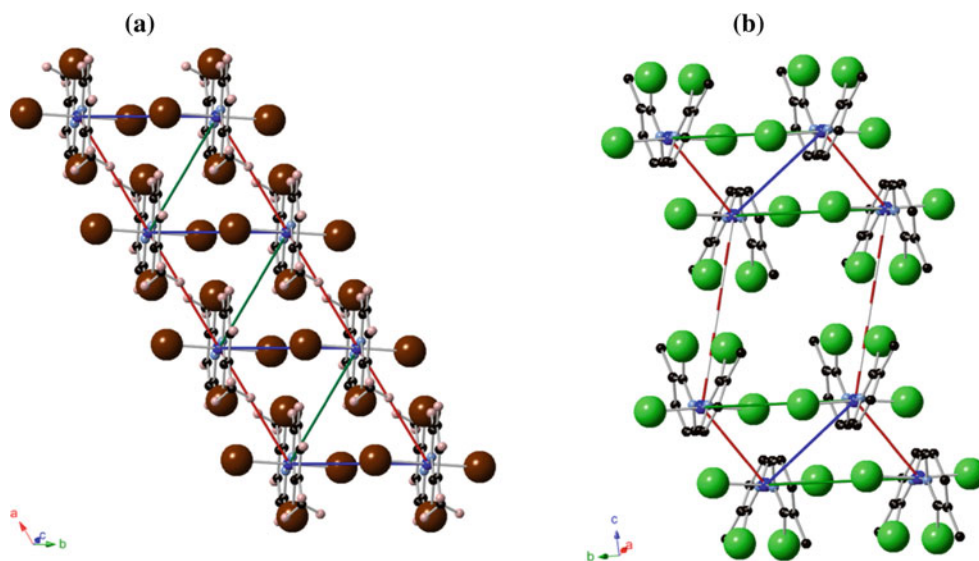
Fig. 10 Magnetic susceptibility curves computed for **a 1** and **b 2**. The *inset* shows a detailed view of the low temperature region of the *curve*. Note that in both compounds, the computed data overlap nearly exactly irrespective of the magnetic model used

negligible to compute $\chi(T)$, which is the magnetic property of interest.

3.3 The effect of conformation on the different magnetic behavior of **1** and **2**

The FPBU study carried out in this work helps in rationalizing the magnetic behavior of **1** and **2**. As it will be discussed, the experimental evidence regarding the magneto-structural relationship for these species is supported by the theoretical calculations presented in this work. A local difference in the conformation of the radical units (see Fig. 1 for *syn-* and *anti*-conformers) leads to a crucial change in the crystal packing and, thus, to the magnetic topology. This can be seen particularly in the d_1 pair that involves the shortest possible Cu···Cu distance between radicals. In **2**, the *syn* arrangement of the (2-Cl-3-Mepy) substituents in a monomer allows a second monomer to approach it along the non-sterically hindered face forming dimers at 4.329 Å (red solid line in Fig. 11b). This fact maximizes the overlap between the SOMO orbitals and enhances the antiferromagnetic interaction between them ($J(d_2) = -2.37$ cm^{-1}). Meanwhile, the closest adjacent monomer to the sterically hindered face is at 8.459 Å and no magnetic exchange interaction is found (red–white striped line in Fig. 11b).

On the other hand, in the brominated counterpart **1**, the *anti*-disposition of the (2-Br-3-Mepy) rings blocks any further coordination at both sterically hindered faces. This local arrangement results in a regular chain with a Cu···Cu distance of 6.244 Å between radicals ($J(d_1) = -0.11$ cm^{-1}, red solid line in Fig. 11a).

This behavior is known. A number of complexes of the general formula Cu(S-py)$_2$X$_2$, where S-py is an

Fig. 11 Superimposed views of the crystal structure and the magnetic topology of **a 1** and **b 2**. Hydrogen atoms have been removed for clarity. *Color code* for atoms: Cu (*deep-blue*), N (*blue*), C (*black*), Cl (*green*), Br (*brown*)

Fig. 12 Geometry of **a 1-opt**, **b 2-opt**, and **c 2**. *Color code* for atoms: Cu (*salmon*), N (*blue*), C (*grey*), H (*white*), Cl (*green*), Br (*brown*)
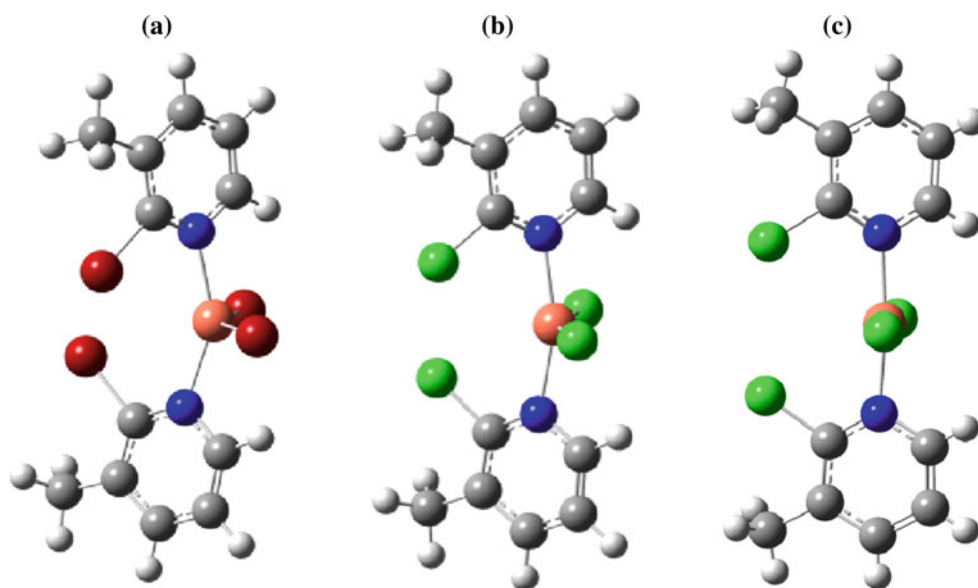
unsymmetrially substituted pyridine and X = Cl, Br, have been prepared and their structures fall into two categories: those which form chain-like structures (similar to **1**) [27–32] and those which form dimeric structures via short X···Cu contacts (similar to **2**) [28, 33, 34]. In all these cases, the relationship between the orientation of the substituents on the pyridine rings (*syn* or *anti*) and the structure holds; dimers are generated by *syn*-conformations and chains by *anti*-conformations. This is not surprising for 2-substituted pyridine ligands, but it also holds for 3-substituted pyridine ligands where the substituents are further from the metal center.

To elucidate whether this behavior can be ascribed to a local effect of the radical unit or to an effect of the crystal packing, we have conducted geometry optimizations of the (2-Cl-3-Mepy)$_2$CuCl$_2$ (**2-opt**) and (2-Br-3-Mepy)$_2$CuBr$_2$ (**1-opt**) monomers arranged in both the *syn*- and *anti*-conformations. In the gas phase, compounds **1-opt** and **2-opt** are more stable in the *syn* arrangement by 7.8 and 1.8 kcal mol^{-1}, respectively, which was expected for **2-opt** as it is the conformation present in the crystal but not for **1-opt**. Although an accurate study should be based on solid-state calculations performed on the crystal structure, a qualitative interpretation can be drawn from gas phase calculations on monomers.

According to the optimized geometries in the gas phase (Fig. 12a, b), the dihedral angle between (2-X-3-Mepy) rings is about 70° for **1-opt** and 54° for **2-opt**, indicating the preference of these monomers to be largely distorted from the planar 0° dihedral angle in the *syn* arrangement (Fig. 12c). The smaller distortion in **2-opt** is reduced to 30° in the crystalline structure of **2** indicating that the crystal packing seems to force the closure of the dihedral angle between rings. The larger distortion in **1-opt** can explain

the absence of a *syn* polymorph as the crystals of **1** cannot accommodate such distortion.

In fact, within the (2-X-3-Mepy)CuX′$_2$ family of compounds, the heterohalide compound (2-Cl-3-Mepy)CuBr$_2$ with Cl atoms attached to the Mepy ring, reported in the same experimental work [8], shows both types of conformations in its polymorphs: *syn*-dimers and *anti*-chains. This indicates that, when X is Cl, the difference in stability between the *syn*-and *anti*-conformers is small enough to obtain a mixture of them, irrespective of X′ being Br. In view of these data, it appears that the key factor responsible for the monomer to adapt to a *syn*-or *anti*-conformation in the solid state is the size of the substituent in the 2-position of the Mepy ring.

4 Conclusions

Within a First-principles Bottom-up (FPBU) strategy, the computed macroscopic magnetic susceptibility χ(T) curves of bis(2-bromo-3-methylpyridine)dibromo copper(II), **1**, and bis(2-chloro-3-methylpyridine)dichlorocopper(II), **2**, agree with the experimentally observed dominant ferro- and antiferromagnetic behavior, respectively.

The computed magnetic topology of **1** consists of ferromagnetic chains ($J(d_2) = 1.16$ cm^{-1}) along the *b*-crystallographic axis, which then weakly interact (−0.11 and 0.13 cm^{-1}) giving rise to magnetic planes. For **2**, the strongest exchange coupling is antiferromagnetic and generates dimers ($J(d_1) = -2.37$ cm^{-1}) that, in turn, weakly interact (0.48, −0.29, and −0.20 cm^{-1}) to form double-decker magnetic planes. In both crystals, the 2D planes pile up along the third dimension showing no magnetic coupling between planes. The ferromagnetic d_2

radical-pair in **1** is a clear through-space Cu-Br⋯Br–Cu interaction, while the antiferromagnetic d_1 in **2** shows a mixture of through-space Cu⋯Cu, Cu⋯Cl, and Cl⋯Cl exchange pathways.

For simulation purposes, only the largest J_{AB} magnetic interactions are required to numerically reproduce the magnetic susceptibility as a function of temperature data. This is in agreement with the fitting models put forward on the basis of geometrical considerations and used to experimentally reproduce the measured $\chi(T)$ data, namely a ferromagnetic chain model with Curie–Weiss interchain corrections for **1** and an antiferromagnetic dimer model for **2**. Therefore, the FPBU analysis quantitatively traces down the origin of the different magnetic behavior of **1** and **2** as due to the change in sign of their dominant magnetic interactions. We have been able to connect such a change in nature of the dominant magnetic interaction with a change in the conformation of the ligands, which converts from *anti* in bis(2-bromo-3-methylpyridine) (**1**) to *syn* in bis(2-chloro-3-methylpyridine) (**2**). The relationship between the orientation of the substituents on the pyridine rings (*syn* or *anti*) and the structure holds; dimers are generated by *syn*-conformations (e.g. **2**) and chains by *anti*-conformations (e.g. **1**). It appears that the key factor responsible for the monomer to adapt to a *syn*- or *anti*-conformation in the solid state is the size of the substituent in the 2-position of the Mepy ring.

References

1. Johnston DC, Johnson JW, Goshorn DP, Jacobson AJ (1987) Phys Rev B Condens Matter Mater Phys 35(1):219
2. Barnes T, Riera J (1994) Phys Rev B Condens Matter Mater Phys 50(10):6817
3. Eccleston RS, Barnes T, Brody J, Johnson JW (1994) Phys Rev Lett 73(19):2626
4. Schwenk H, Konig D, Sieling M, Schmidt S, Palme W, Luthi B, Zvyagin S, Eccleston RS, Azuma M, Takano M (1997) Physica B 237–238:115
5. Eccleston RS, Mutka H, Payen C (1997) Physica B 234–236:895
6. Garrett AW, Nagler SE, Barnes T, Sales BC (1997) Phys Rev B Condens Matter 55(6):3631
7. Garrett AW, Nagler SE, Tennant DA, Sales BC, Barnes T (1997) Phys Rev Lett 79(4):745
8. Herringer SN, Turnbull MM, Landee CP, Wikaira JL (2011) Dalton Trans 40:4242
9. Deumal M, Bearpark MJ, Novoa JJ, Robb MA (2002) J Phys Chem A 106:1299
10. Szabo A, Ostlund NS (1989) Modern quantum chemistry: introduction to advanced electronic structure theory. McGraw-Hill Inc., New York
11. Hohenberg P, Kohn W (1964) Phys Rev 136:B864
12. Clarke CS, Jornet-Somoza J, Mota F, Novoa JJ, Deumal M (2010) J Am Chem Soc 132:17817
13. Deumal M, Giorgi G, Robb MA, Turnbull MM, Landee CP, Novoa JJ (2005) Eur J Inorg Chem 23:4697
14. Vela S, Deumal M, Turnbull MM, Novoa JJ (2012) Polyhedron. doi:10.1016/j.poly.2012.07.085
15. Jornet J, Li L, Turnbull MM, Landee CP, Deumal M, Novoa JJ, Wikaira JL (2007) Inorg Chem 46:11254
16. Novoa JJ, Deumal M, Jornet-Somoza J (2011) Chem Soc Rev 40:3182–3212
17. Noodleman L (1981) J Chem Phys 74:5737
18. Noodleman L, Davidson ER (1986) Chem Phys 109:131
19. Jornet-Somoza J, Deumal M, Turnbull MM, Novoa JJ (2009) Polyhedron 28:1965
20. Parr EG, Yang W (1989) Density functional theory of atoms and molecules. Oxford University Press, New York
21. Becke AD (1988) Phys Rev A 38:3098
22. Becke AD (1993) J Chem Phys 98:5648
23. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785
24. Schafer A, Horn H, Ahlrichs R (1992) J Chem Phys 97:2751
25. Ditchfield R, Hehre WJ, Pople JA (1971) J Chem Phys 54:724
26. Frisch MJ et al (2009) Gaussian 09, Revision B.1. Gaussian, Inc, Wallingford
27. Shortsleeves KC, Dawe LN, Landee CP, Turnbull MM (2009) Inorg Chim Acta 362:1859
28. Awwadi F, Willett RD, Twamley B (2011) Cryst Growth Des 11:5316
29. van Albada GA, Tanase S, Mutikainen I, Turpeinen U, Reedijk J (2008) Inorg Chim Acta 361:1463
30. Awwadi FF, Willett RD, Haddad SF, Twamley B (2006) Cryst Growth Des 6:1833
31. Espallargas GM, van de Streek J, Fernandes P, Florence AJ, Brunelli M, Shankland K, Brammer L (2010) Angew Chem Int Ed 49:8892
32. Lah N, Leban I (2010) Struct Chem 21:263
33. Singh P, Jeter DY, Hatfield WE, Hodgson DJ (1972) Inorg Chem 11:1657
34. Duckworth VF, Stephenson NC (1969) Acta Crystallogr Sect B Struct Crystallogr Cryst Chem 25:2245