The molecular electrostatic potential analysis of solutes and water clusters: a straightforward tool to predict the geometry of the most stable micro-hydrated complexes of β-propiolactone and formamide

Cairedine Kalai, Mohammad Esmaïl Alikhani, Emilie-Laure Zins

The molecular electrostatic potential analysis of solutes and water clusters: A straightforward tool to predict the geometry of the most stable micro-hydrated complexes of β-Propiolactone and Formamide

Cairedine KALAI,[a,b] Mohammad Esmaïl ALIKHANi [a,b] Emilie-Laure ZINS, [a,b*]

a: Sorbonne Universités, MONARIS, UMR 8233, Université Pierre et Marie Curie, 4 Place Jussieu, case courrier 49, F-75252 Paris Cedex 05, France

b: CNRS, MONARIS, UMR 8233, Université Pierre et Marie Curie, 4 Place Jussieu, case courrier 49, F-75252 Paris Cedex 05, France

[*]: Corresponding author, emilie-laure.zins@sorbonne-universite.fr

Abstract

High-resolution microwave spectroscopic studies have recently revealed the structures of six micro-hydrated complexes of β-propiolactone (BPL) as well as five microhydrated complexes of formamide. Complexes containing one to three water molecules were selected as model systems to set a methodological approach based on the study on the isolated partner for the identification of the most stable microhydrated complexes. A four-step study revealed to be particularly straightforward for the identification and characterization of water-solute complexes: i) the topological analysis of isolated partners, to identify complementary interaction sites, ii) the proposal of possible preferable direction of approaches of both partners to allow multiple interactions between complementary interaction sites, iii) the full optimization of initial structures at the MP2 level of theory and iv) the topological analysis of the water-solute intermolecular interactions. It is shown that the segregation of water molecules experimentally observed, the larger complexation energy of $X:(H_2O)_n$ for $X =$ Formamide rather than BPL, and the larger complexation energy of the dimers compared to the monomers and the trimers could be related to the molecular electrostatic potential of isolated partners.

In memoriam of János G. Ángyán: for his in-depth contribution to the description of weak interactions

Graphical abstract
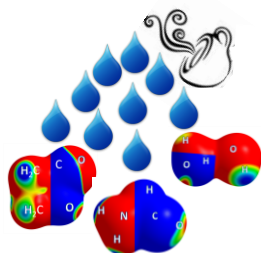
Introduction

Water molecules play a fundamental role in many chemical processes, including biochemistry, (bio-) catalytic processes, atmospheric chemistry, environmental chemistry and exo-biology. From a biochemical point of view, it was shown that the variation of the hydrogen bonding network between water and a protein can directly impact the folding of the latter, and its reacivity.[1]

As far as atmospheric chemistry is concerned, [2] the role of hydrated complexes on the reactivity of air pollutants[3] Criegee intermediates,[4] and small molecules[5] has been the subject of recent studies. The competition between van der Waals type interactions and hydrogen bondings has recently been demonstrated, particularly in the case of ozone hydration.[6]

Despite the wide role of water molecules in many different fields, water-solute interactions, their complementary and competitive role against water-water interactions, are still poorly understood. Studies based on molecular dynamics allow the characterization of interactions between the solute and the solvent molecules. As an example on one of the solute that will be further investigated in the present study, we can cite the QM/MM calculations carried out on the hydration of formamide to shed some light on the hydrolysis of the solute in 2005.[7] More recently, quantum-mechanical charge field-molecular dynamics[8] as well as force-field and Monte Carlo calculations were carried out on this system. [9] At the molecular scale, among other state-of-the art approaches[10], supersonic jet experiments coupled with high-resolution microwave spectroscopy provide valuable structural information in addition with theoretical calculations. Among interesting results recently published with this double experimental and theoretical approach,[11,12-15] two works will be herein discussed: the micro-hydration β-propiolactone (BPL),[16] and formamide.[17,18,19]

In the case of the hydration of BPL, complexes containing up to five water molecules have been identified by means of high resolution microwave spectroscopy. The structures of $BPL:(H_2O)_{n=1-3}$ complexes identified on the basis of spectroscopic features are presented in Figure 1. Globally, the positions of the oxygen atoms of water clusters are only slightly affected by the interaction with the solute. The structural disturbances of the water cluster mainly concern the orientation of hydrogen atoms not involved in the hydrogen bonds network between water molecules. The possible inversion of the hydrogen bonding network between water molecules was also underlined in this experimental study.

In the case of the micro-hydration of formamide, three monohydrated isomers have been characterized (Figure 1). The energy gap between these three isomers has been estimated at around 20 $kJ.mol^{-1}$. One dihydrated as well as one trihydrated isomers have also been experimentally identified. From the structures deduced from the experimental observations, it has been proposed that the $Formamide:(H_2O)_n$ complexes would preferentially adopt structures close to those of the most stable isomers for the $(H_2O)_{n+2}$ complexes (Figure 1).
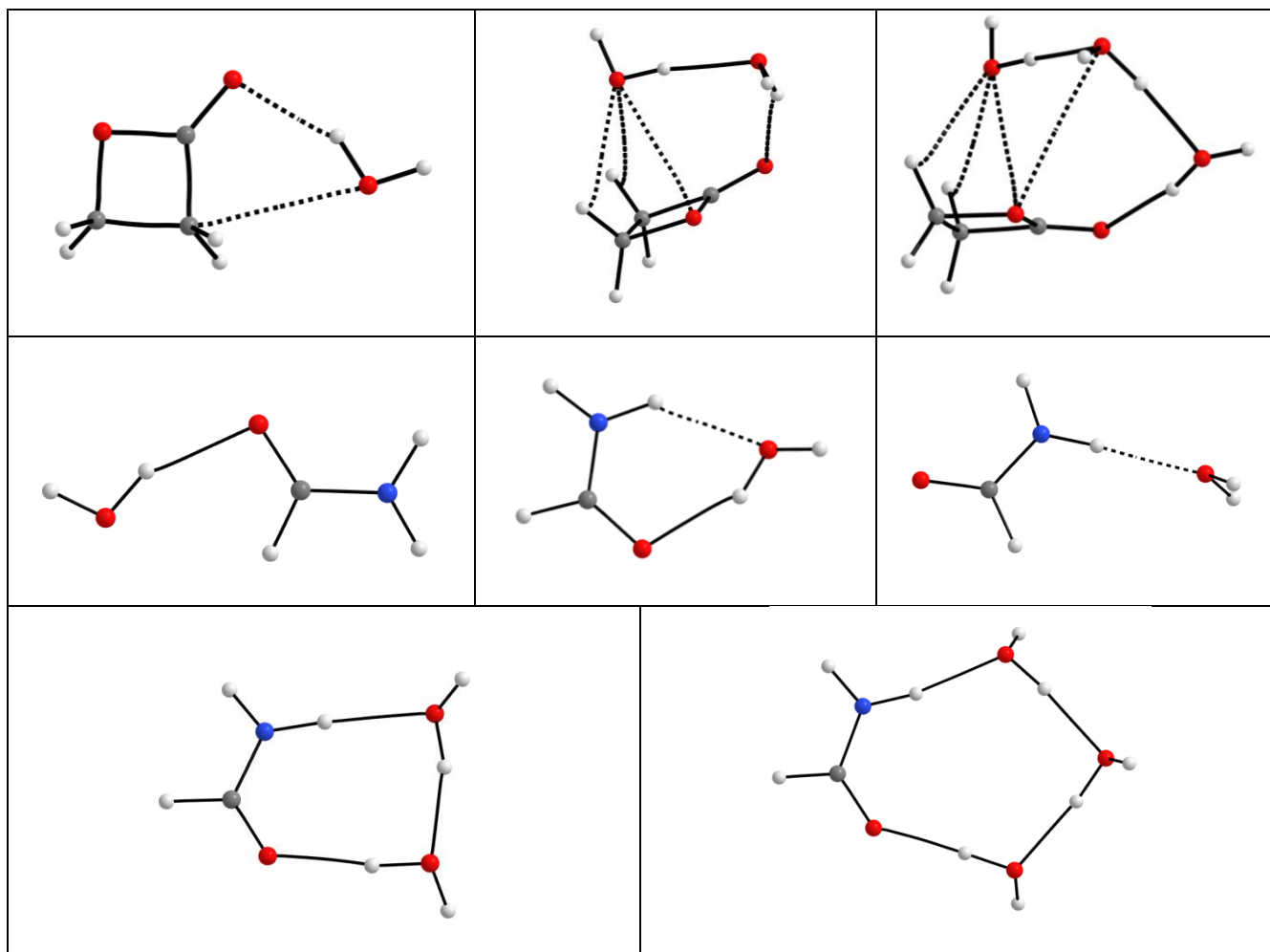
Figure 1: Optimized structures (MP2/avdz) identified by Alonso et al. and Blanco et al.[17,19] on the basis of microwave spectroscopic investigations (Interaction energies are given in Figures 4 and 7, as well as in S.I.1-7 and are discussed in the "Result" part).

Thus, these two in-depth investigations led to the proposal of two slightly different visions of micro-hydration: On the one hand, the solute could be considered as a template on which water clusters in their most stable forms interact with weak structural adaptations in the case of BPL. On the other hand, a solute like formamide would replace two molecules of water in pure water clusters.

The questions arisen from the comparison of these two studies are as follows: Would there be two phenomenologically different pathways of hydration depending on the nature of the solute? If so, how to characterize them? Could we propose a unified theoretical methodology to predict and explain the hydration pattern of both BPL and formamide?

Below will these question be addressed. Several tools based on the quantum chemical calculations were selected for this study at the molecular level.

For several decades, theoretical chemists have relied on the molecular electrostatic potential (MESP), calculated more and more accurately, for the study of intermolecular interactions.[20,21,22] Currently, it is often the MESP calculated from the approach of the quantum theory of atoms in molecules (QTAIM) of Bader[23] which is exploited.[24] Its reliability for the prediction of experimental observables has been demonstrated.[25,26] Thus, Politzer and Murray were able to identify, quantify

3

and compare different classes of intermolecular interactions between the most electrophilic site of a first partner and the most nucleophilic site of a second partner.[27] Their essentially explanatory approach makes it possible to establish a clear phenomenological relationship between chemically different interactions: for example, halogen and hydrogen interactions were found to have a similar physical origin, even if the chemist considers them to be different from nature.[28,29]

Scheiner and collaborators,[30,31,32] de Proft and collaborators,[33,34] and others[35,36] analyzed the MESPs of isolated partners prior to extensively studied non covalent interactions involving lone pairs.

A slightly different approach was developed by Gadre and collaborators, as well as Suresh and collaborators.[37] Indeed, these authors developed an home-made software for geometry optimizations closed to molecular dynamics simulations, based on the topological analysis of the MESP function of the isolated partners.[38,39,40] In such an approach, critical points of the MESP function are analyzed. Complexes are generated by the association of sites characterized by the strongest value of MESP of one partner and the weakest values of MESP of the other partner. Furthermore, such an approach allows the identification of lone pairs by analogy with Lewis structures of the species.[41,42,43]

As we will see in the "Methodological Approach" section, we sought to develop a more general approach, based on the identification of every electrophilic and nucleophilic sites of both partners, in order to identify all possible directions of favorite approaches, in line with the concepts developed by Politzer, Murray, Scheiner, de Proft and collaborators. This approach allows us to propose initial isomeric structures for the complexes. From these proposed structures, geometry optimizations and frequency calculations are carried out at the MP2/aug-cc-pVDZ level of theory in order to identify the minima on the potential energy surface. We will show how this approach makes it possible to find the structures of microhydrates experimentally observed for the two solutes. To explain the origin of the stabilization of isomers experimentally identified, we use two additional topological tools:

- ➤ the topological analysis of the electron density (ρ) within the QTAIM framework,[23,44,45]
- ➤ the topological analysis of the electron localization function (ELF), leading to the identification of valence basins.[46,47,48,49]

The QTAIM analysis mainly leads to the identification of bond paths (BP) and bond critical points (BCP). Atoms involved in the intermolecular interactions can thus be identified. The study of the electron density at the critical points ($\rho(BCP)$) of the intermolecular bonds makes it possible to identify the predominant interactions, characterized by the strongest $\rho(BCP)$ values, as well as secondary interactions characterized by lower $\rho(BCP)$ values.

Topological analysis of the ELF function is also useful to better understand the interactions stabilizing an isomer. This topological analysis makes it possible to locate the electronic valence basins, and thus to locate the electronic pairs in the Lewis sense.[50,51,52] In the case of the study of the micro-hydration process, this approach is particularly relevant:

- ➤ lone electron pairs involved will be identified in isolated partners. This will be helpful to predict the possible hydrogen bond formation between two partners,
- ➤ the identification of electron pairs involved in intermolecular interactions will help in understanding the nature of interactions that stabilizes the complex.

As a first step, we will deepen the MESP study of isolated partners to identify a possible phenomenological difference that could lead to a different micro-hydration of BPL and formamide.

In a second part, we will look for the privileged directions of $(H_2O)_{1-3}$ approaches compared to the BPL, and we will discuss the structures obtained by geometry optimization at the MP2/aVDZ level of theory compared to the experimental results. QTAIM and ELF topological analyses will help in understanding the natures and the strengths of the interactions involved in each isomers. Later on, a similar methodology will be applied to the search for - and characterization of - Formamide:$(H_2O)_{1-3}$ complexes.

In a third part, the natures and the strengths of the water-solute interactions will be discussed to answer the question: "Are the hydration pathways of BPL and Formamide phenomenologically different?".

In a fourth part, results obtained using the methodological approach herein presented will be compared with the predictions of Gadre et al.[53, 54] based on a topological analysis of the MESP function in the case of the micro-hydration of formamide.

Methodological approach

At long distance, two species A and B will orient themselves relative to each other so as to optimize interactions between complementary sites. Hydrogen bondings-, Burgi-Dunitz- and pi-holes-types of interaction being of electrostatic nature, it is quite relevant to study the molecular electrostatic potential of the solute on the one hand and the water complexes on the other hand, in order to determine preferred directional approaches of both partners. The use of molecular electrostatic potential (MESP) makes it possible to identify electrophilic and nucleophilic sites on the two partners A and B (with A = S and B = W_n in the case of formation of the hydrated complexes). This step makes it possible to identify the long-distance interactions between the isolated partners, which lead to the formation of the complex. The partner B will orient itself in such a way as to best coincide its electrophilic sites with nucleophilic sites of partner A and vice versa.

This is the reason why we proposed a methodological approach based on a thorough analysis of the MESPs of the isolated partners to propose guess structures for the complexes. This step leads to the identification of several guessed isomers that are used as initial structures for full geometry optimizations (Scheme 1).
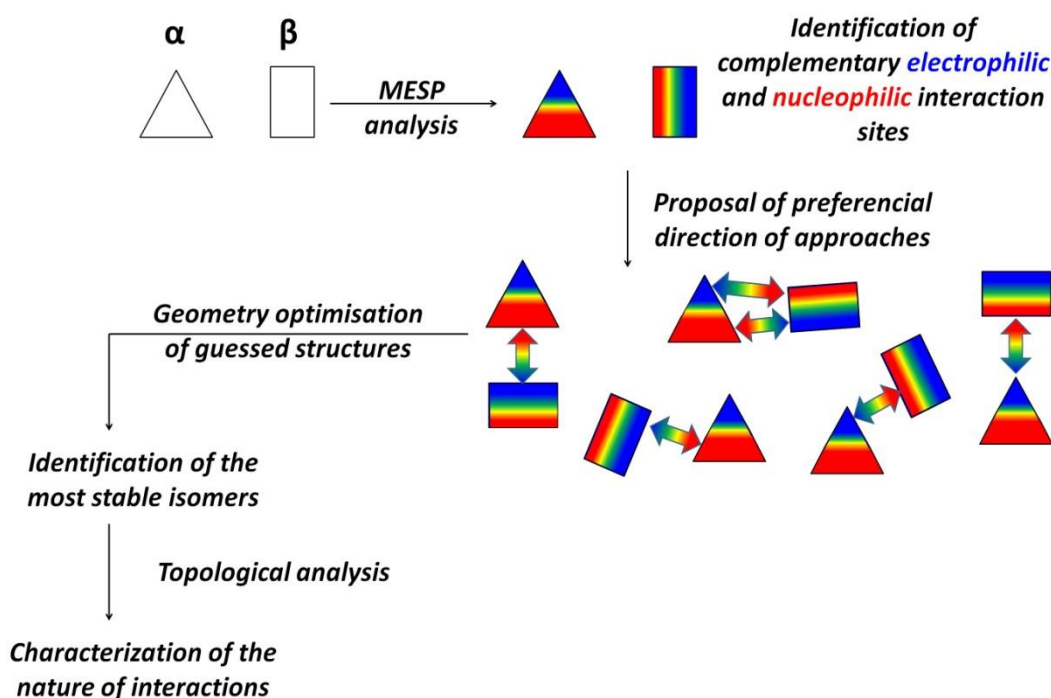
Scheme 1: Methodological approach for search of the most stable complexes formed between α and β species. For sake of simplification, only 2D approaches were considered in this schematic sketch.

To begin with, MP2/aug-cc-pVDZ geometry optimization and frequency calculations[55,56] were carried out for the isolated monomers and water cluster. The notation MP2/aVDZ will be further used for this level of theory. The geometry of the water cluster selected for the present study are those identified by Shields et al. on the basis on CCSD(T) calculations.[57]

Except otherwise mentioned, all the MP2 calculations were carried out with Gaussian09 package [58] and the "Density = Current" option was used to generate the .wfn files further used for the topological analyses. The electron densities thus obtained in the .wfn files are the so-called "relaxed MP2 density" based on the Z-vector formalism.

AIMALL package[59] was used for the generation of the MESP, as well as for the QTAIM analysis of the complexes. In each picture above presented, small green circles represent the bond critical points (BCP), and the black lines indicate bond paths (BP). The MESPs were generated by plotting the total electrostatic potentials on a surface of electronic isodensity ($\rho = 0.001$). For the ELF topological analysis, the Topmod suite package was used.[60] Except otherwise mentioned, all the values are given in a.u. Complexation energies were calculated as follows:

$$De = E_{S(H_2O)_n} - [E_S + E_{(H_2O)_n}]$$

The concept of preferred approach direction applied to the identification of the BPL:(H$_2$O) isomers

Identification of complementary sites of interactions

The MESP analysis of the BPL leads to the identification of seven electrophilic zones: the neighborhood of each of the four hydrogen atoms, two π-holes, perpendicular to the plane of the molecule, a σ-hole , located in the axis of the CO bond (Figure 2).
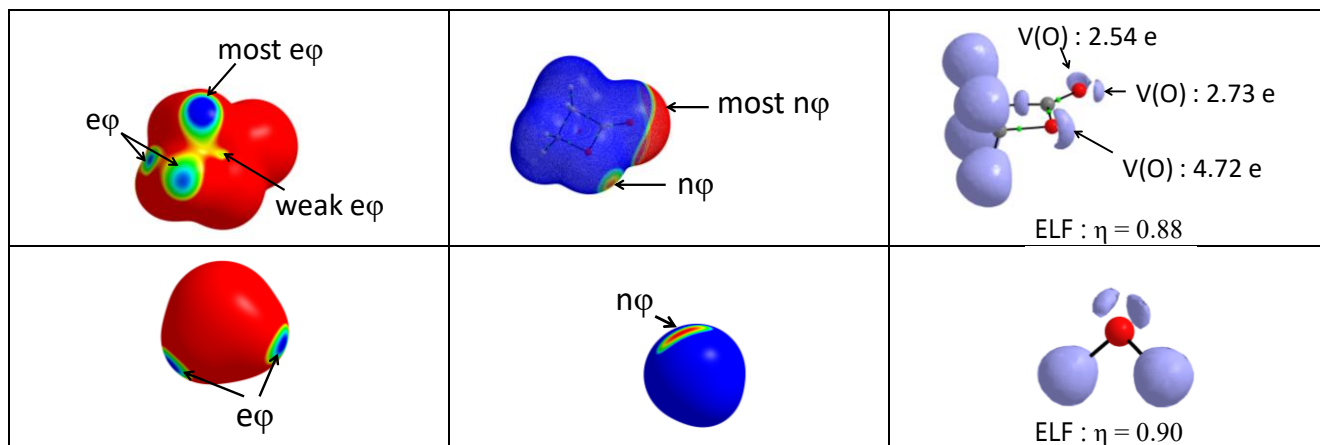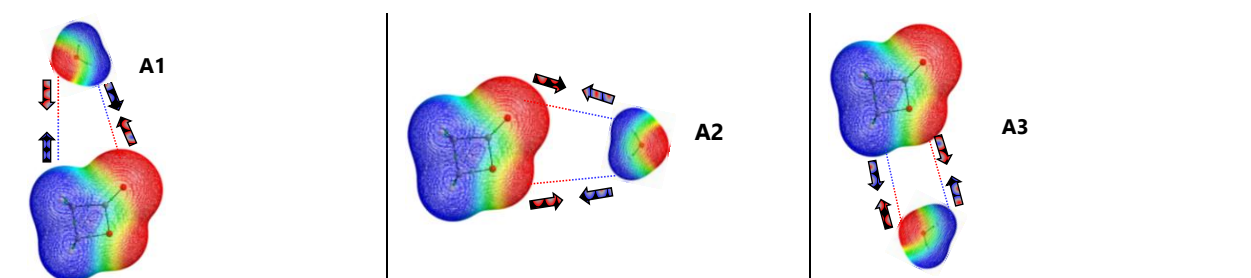
Figure 2: Identification of electrophilic(eφ, left) and nucleophilice (nφ, middle) zones on the BPL and water molecules with the MESP analysis. The ELF valence basins were plotted for : η = 0.88 (BPL) and 0.90 (H₂O, right).

On the contrary, the neighborhood of both oxygen atoms corresponds to a nucleophilic zone. The lone pairs located on the oxygen atoms can be further characterized by the mean of the ELF topological approach. The oxygen atom of the carbonyl group is characterized by two monosynaptic basins, having slightly different populations: 2.54 and 2.73 e. On the other hand, only one monosynaptic basin is located on the oxygen atom of the ether group, with a population of 4.72 e.

The characterization of these three monosynaptic basins give some insight onto the possible interactions of the BPL with a given partner. Indeed, according to Legon and Millen, in the equilibrium structure of a hydrogen bond complex that does not involve π-bonds, the axis of a non-bonding pair of electrons coincides with the axis of the hydrogen-bond donor molecule.[61,62] In that respect, the identification of the location of the lone pairs with the ELF topological approach is particularly relevant for the search for the possible formation of hydrogen bondings.[63,64] Complementarily, the MESP analysis of the water molecule leads to the identification of two electrophilic regions at the level of the hydrogen atoms. On the other hand, a single nucleophilic region is located above the oxygen atom, and the ELF topological analysis is required for the identification of two monosynaptic basins corresponding to the free electron pairs of the oxygen atom.

Based on these MESP and ELF topological characterization ,of both partners, nine possibly favored directions of approach of a water molecule relative to the BPL molecule can be considered (Figure 3).
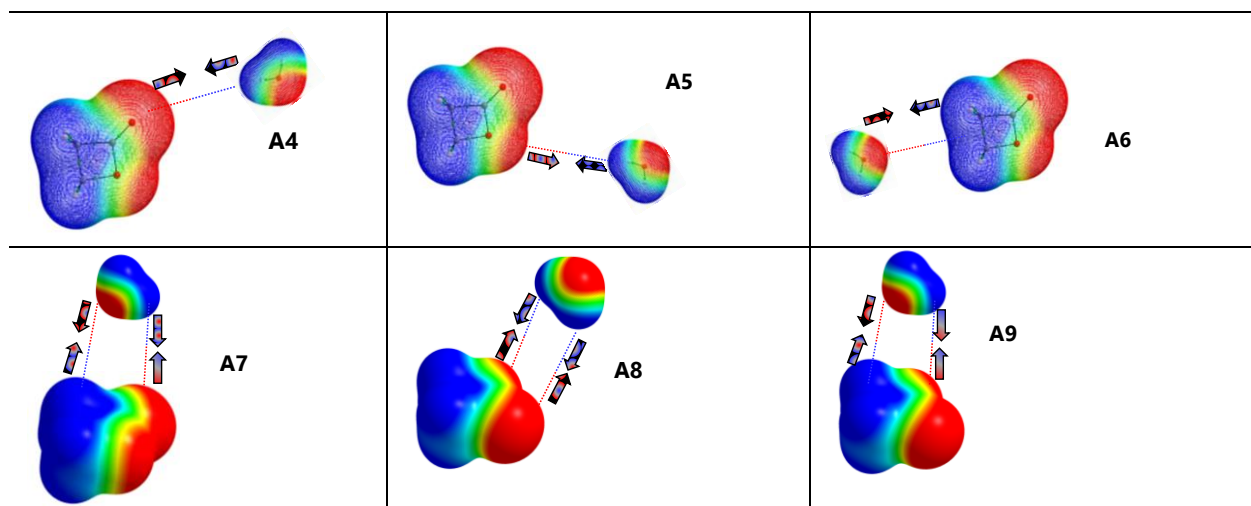
Figure 3: Identification of nine possible preferred approach directions of a water molecule relative to the BPL based on the maximization of interactions between complementary sites. Electrophilic (respectively nucleophilic) regions are colored in blue (respectively red).

These structures were used as guessed-structures for geometry optimizations. From these guessed-structures, five isomers are obtained after geometry optimization at the MP2/aVDZ level of theory. Isomers are ranked by order of their stability, from the most stable one (denoted S1) to the less stable one (denoted S5). The most stable isomer will be denoted S1. The S2 isomer is calculated to be quasi iso-energetic of S1 at the MP2/aVDZ level, only 4.6 kJ.mol^{-1} above the global minimum. It is worth mentioning that this second isomer in order of stability has a three-dimensional structure: the molecule of water is not located in the main plane of the solute but perpendicular to it. The other isomers, S3, S4 and S5 are located 9.5 kJ.mol^{-1}, 10.0 kJ.mol^{-1} and 16.2 kJ.mol^{-1} above the global minimum, respectively. The bond critical points (BCP, QTAIM analysis) as well as the valence basins (ELF analysis) of isomers of the BPL:(H_2O) mono-hydrated complex are shown in Figure 4.

In line with the chemist's intuition, the ELF analysis of the S1 isomer shows two hydrogen bondings: each protonated basin of the CH_2 group interacts with one of the monosynaptic basin of the O_{water} atom. However, the ELF analysis reveals that the monosynaptic basin of the oxygen atom of the carbonyl group involved in this interaction is not the most populated one, but the less populated one. This constitutes a notable deviation from the Legon and Millen rules. Simultaneously, one of the monosynaptic basins of the oxygen atom of the carbonyl function interacts with one of the protonated basins of the water molecule.

In the S2 isomer, the ELF analysis reveals that the oxygen atom of the water molecule acts as a double hydrogen-bonding acceptor: two neighbor hydrogen atoms located on two neighbor carbon atoms interact with the monosynaptic basins of the O_{water} atom. There is an additional interaction between the monosynaptic basin of the O_{ether} atom and a protonated basin of the water molecule. This electrostatically-driven interaction does not correspond to a hydrogen bonding in the sense of the Legon and Millen rules, since the monosynaptic basin of the ether is not aligned on the axis of the protonated basin. This demonstrates that the hydration processes are not uniquely governed by the formation of hydrogen bonding interactions.

A similar interaction between a protonated basin of the water molecule and the monosynaptic basin of the O_{ether} atom is observed in the S3 isomer, in addition with another "classical" hydrogen bond,

involving the monosynaptic basin of the oxygen atom of the carbonyl function and one of the protonated basins of the water molecule.

In the case of the S4 isomer, the ELF analysis leads to the identification of one hydrogen bonding involving a monosynaptic basin of the oxygen atom of the carbonyl function and a protonated basin of the water molecule. This interaction involves the most populated monosynaptic basin of the oxygen atom of the carbonyl group. In the context of the Legon and Millen rules, it is worth mentioning that this isomer is significantly higher in energy than the global minimum, lying 10 kJ.mol^{-1} above the global minimum, according to MP2/aVDZ calculations.

The ELF analysis of S5 suggests the existence of two hydrogen bondings, each of which involving one of the monosynaptic basin of the O$_{water}$ atom and one of the protonated basin of the CH$_2$ group close to the ether function.

In the S1 isomer, the water molecule acts as a hydrogen bonding donor, and the hydrogen bonding acceptor basin is the monosynaptic basin of the O$_{ketone}$ atom with the smallest population (around 2.5 e). However, according to the topological rules for hydrogen-bonded complexes, the monosynaptic basin with the highest population (2.7 e) carried by this O$_{ketone}$ atom should be preferentially engaged in a hydrogen bond. Indeed, according to the Legon and Millen rules, a hydrogen-bonding interaction involving the monosynaptic basin with the highest population should lead to the most stable isomer. Such interaction with the most nucleophilic site of the solute would also be predicted in other models based on the MESP analysis, such as the EPIC model of Gadre.[43,65,41] However, our methodological approach demonstrates that only the S4 isomer that lies 10 kJ/mol above the global minimum involves such an interaction. The only isomer that was identified on the basis of experimental data is the S1 isomer.[16] This underlines the importance of a global analysis of the MESP as an interpretative tool for the prediction and explanation of intermolecular interactions.
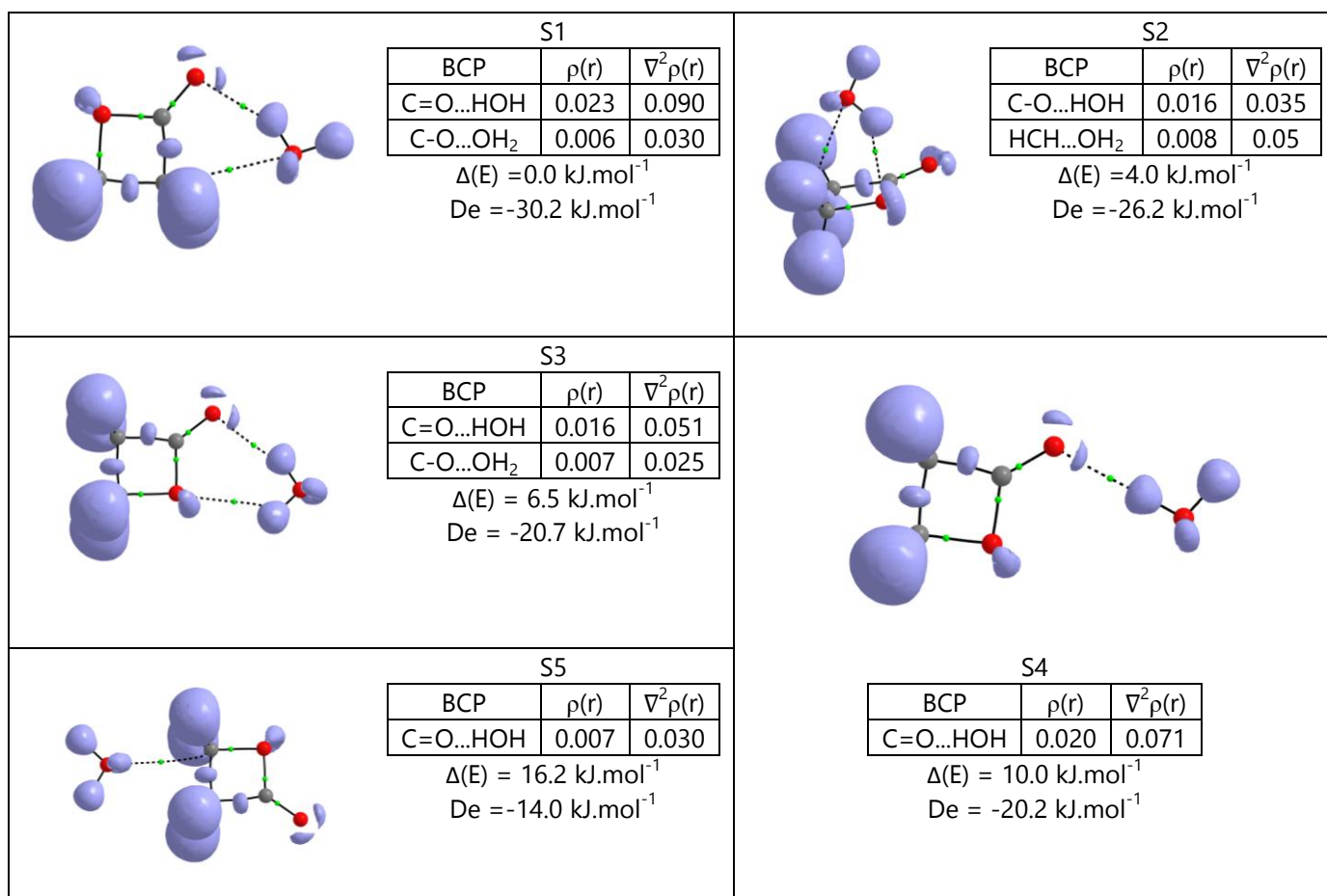
S1		
BCP	$\rho(r)$	$\nabla^2\rho(r)$
C=O...HOH	0.023	0.090
C-O...OH$_2$	0.006	0.030
$\Delta(E) = 0.0$ kJ.mol^{-1}		
De = -30.2 kJ.mol^{-1}		

S2		
BCP	$\rho(r)$	$\nabla^2\rho(r)$
C-O...HOH	0.016	0.035
HCH...OH$_2$	0.008	0.05
$\Delta(E) = 4.0$ kJ.mol^{-1}		
De = -26.2 kJ.mol^{-1}		

S3		
BCP	$\rho(r)$	$\nabla^2\rho(r)$
C=O...HOH	0.016	0.051
C-O...OH$_2$	0.007	0.025
$\Delta(E) = 6.5$ kJ.mol^{-1}		
De = -20.7 kJ.mol^{-1}		

S5		
BCP	$\rho(r)$	$\nabla^2\rho(r)$
C=O...HOH	0.007	0.030
$\Delta(E) = 16.2$ kJ.mol^{-1}		
De = -14.0 kJ.mol^{-1}		

S4		
BCP	$\rho(r)$	$\nabla^2\rho(r)$
C=O...HOH	0.020	0.071
$\Delta(E) = 10.0$ kJ.mol^{-1}		
De = -20.2 kJ.mol^{-1}		

Figure 4: Bond critical points (BCP, QTAIM analysis) in green and valence basins (ELF analysis) of isomers of the BPL:(H$_2$O) mono-hydrated complex.

The concept of preferred approach direction applied to the identification of isomers of the BPL:(H$_2$O)$_2$ isomers

As soon as one is interested in di-hydrated complexes, the question arises of the definition of species a and B. Should we consider a water dimer approaching the solute, or a molecule of water with a monohydrate complex? From an experimental point of view, both are probably relevant, depending on the parameters and the exact assembly. Thus, in a recent study, the Xantheas team used a cell with adjustable pressure to form water clusters by sequential assembly of molecules.

For such an experimental approach, it seems preferable to choose a water molecule for species A and each of the five monohydrate complexes S1 to S5 for species B. However, numerous high resolution experimental studies on polyhydrate complexes rely on the use of a supersonic jet coupled with microwave spectroscopy. Both approaches therefore seem justified in relation to experiments. From a fundamental point of view, for the understanding and the description of the microhydration, both approaches also deserve consideration.

A more precise topological characterization of the water dimer is presented in figure 5. Unsurprisingly, the MESP analysis shows that the electrophilic and nucleophilic sites carried by the two water molecules are different. The most electrophilic sites are located in the OH bond axis of the

10

hydrogen bonding acceptor water molecule. Another electrophilic site is located near the hydrogen pendant of the hydrogen bonding donor molecule. In contrast, the most nucleophilic site is located near the hydrogen atom of the hydrogen bond donor molecule. Another less elecctrophilic site, is located near the oxygen atom of the other water molecule. Three monosynaptic basins available to interact with electrophilic sites are identified by the ELF analysis, as well as three proton basins that can interact with the solute.
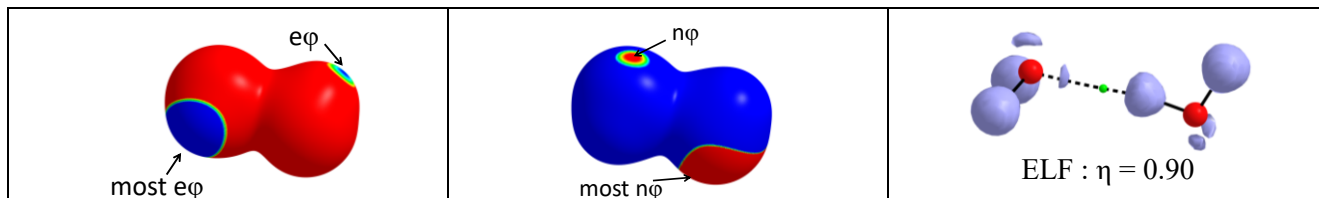

Figure 5: Identification of electrophilic (eφ, left) and nucleophilic (nφ, middle) zones on the (H$_2$O)$_2$ water cluster with the MESP analysis. The ELF valence basins of the water cluster were plotted for : η = 0.90 (right).

The interaction of a dimer of water with the BPL can be done according to five directions of privileged approaches (Figure 6).
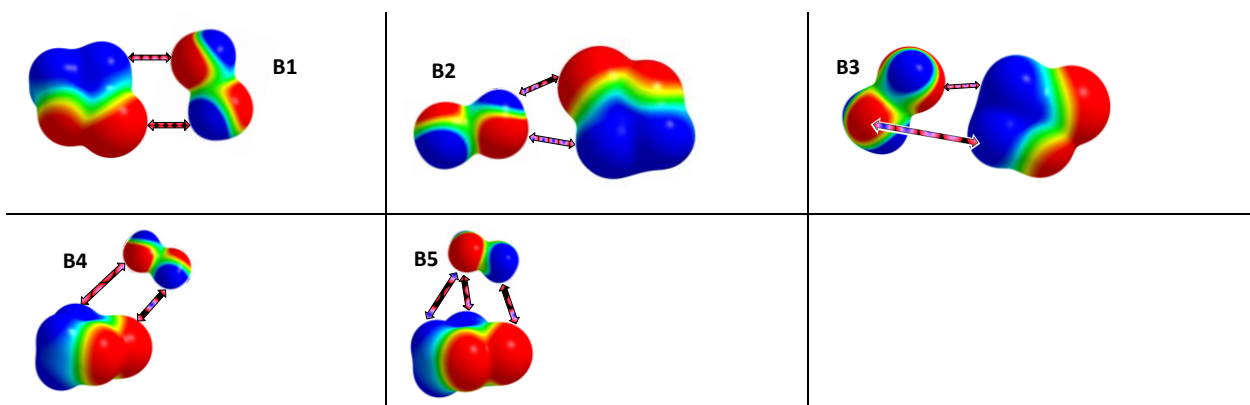

Figure 6: Identification of five possible preferred approach directions of a water dimer relative to the BPL based on the maximization of interactions between complementary sites. Electrophilic (respectively nucleophilic) regions are colored in blue (respectively red).

The geometric optimizations carried out from these five preferred approach directions deduced from the MESP analysis lead to the identification of seven minima on the potential energy surface (Figure 7). Isomers are ranked by order of their stability, from the most stable one (denoted D1) to the less stable one (denoted D7).

QTAIM analysis of the most stable isomer (D1) reveals the formation of four intermolecular interactions between BPL and water molecules: The strongest interaction is identified between the oxygen atom of the carbonyl group and a hydrogen atom of one of the water molecule. The position of the water molecule clearly demonstrate that this interaction cannot be considered as a hydrogen bonding. Instead, this is a relatively strong electrostatically-driven interaction, with ρ(BCP) = 0.022 a.u. It is worth noting that the water molecule dimer is found above the plane of the BPL, and occupies a place comparable to that of the water molecule in the S2 isomer, which is not observed experimentally. This is also the case in the D2 and D3 isomers, that are almost iso-energetic with the D1 global minimum.

A fourth almost iso-energetic isomer of the first three is identified 3.9 kJ.mol^{-1} above the global minimum: it is the D4 isomer (See Figure 7). Its structure is significantly different than that of the three previous isomers. Indeed, in this isomer, the water dimer is almost coplanar with the solute, and two hydrogen bonding interactions are identified by the QTAIM analysis between BPL and water dimer. In the main interaction , BPL acts as a hydrogen bond acceptor with respect to the water molecule that is accepting hydrogen bonding in the water dimer.
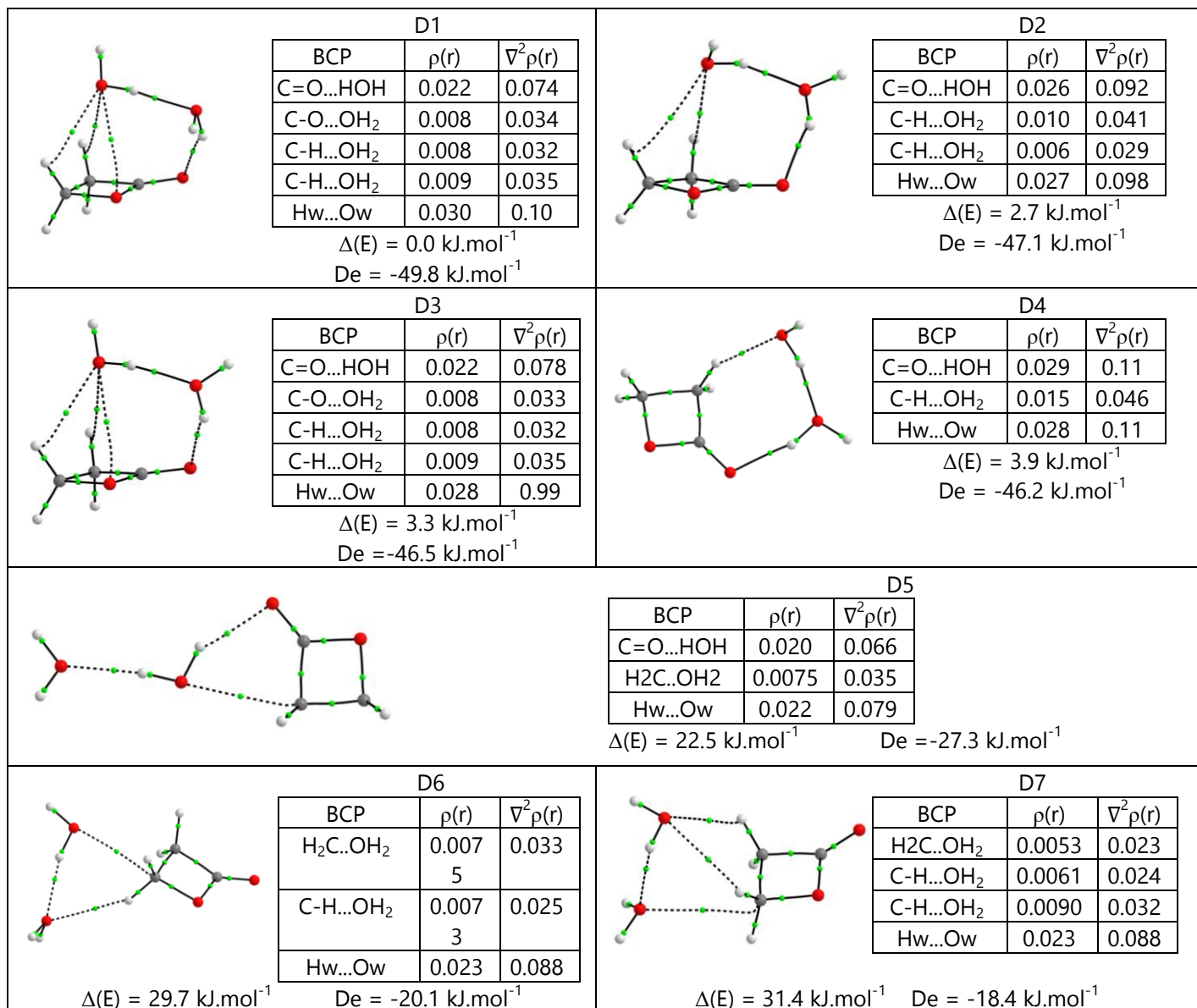

D1

BCP	$\rho(r)$	$\nabla^2\rho(r)$
C=O...HOH	0.022	0.074
C-O...OH$_2$	0.008	0.034
C-H...OH$_2$	0.008	0.032
C-H...OH$_2$	0.009	0.035
Hw...Ow	0.030	0.10

$\Delta(E) = 0.0$ kJ.mol^{-1}
De = -49.8 kJ.mol^{-1}

D2

BCP	$\rho(r)$	$\nabla^2\rho(r)$
C=O...HOH	0.026	0.092
C-H...OH$_2$	0.010	0.041
C-H...OH$_2$	0.006	0.029
Hw...Ow	0.027	0.098

$\Delta(E) = 2.7$ kJ.mol^{-1}
De = -47.1 kJ.mol^{-1}

D3

BCP	$\rho(r)$	$\nabla^2\rho(r)$
C=O...HOH	0.022	0.078
C-O...OH$_2$	0.008	0.033
C-H...OH$_2$	0.008	0.032
C-H...OH$_2$	0.009	0.035
Hw...Ow	0.028	0.99

$\Delta(E) = 3.3$ kJ.mol^{-1}
De =-46.5 kJ.mol^{-1}

D4

BCP	$\rho(r)$	$\nabla^2\rho(r)$
C=O...HOH	0.029	0.11
C-H...OH$_2$	0.015	0.046
Hw...Ow	0.028	0.11

$\Delta(E) = 3.9$ kJ.mol^{-1}
De = -46.2 kJ.mol^{-1}

D5

BCP	$\rho(r)$	$\nabla^2\rho(r)$
C=O...HOH	0.020	0.066
H2C..OH2	0.0075	0.035
Hw...Ow	0.022	0.079

$\Delta(E) = 22.5$ kJ.mol^{-1} De =-27.3 kJ.mol^{-1}

D6

BCP	$\rho(r)$	$\nabla^2\rho(r)$
H$_2$C..OH$_2$	0.0075	0.033
C-H...OH$_2$	0.0073	0.025
Hw...Ow	0.023	0.088

$\Delta(E) = 29.7$ kJ.mol^{-1} De = -20.1 kJ.mol^{-1}

D7

BCP	$\rho(r)$	$\nabla^2\rho(r)$
H2C..OH2	0.0053	0.023
C-H...OH$_2$	0.0061	0.024
C-H...OH$_2$	0.0090	0.032
Hw...Ow	0.023	0.088

$\Delta(E) = 31.4$ kJ.mol^{-1} De = -18.4 kJ.mol^{-1}

Figure 7: QTAIM topological characterization of the most stable BPL:(H$_2$O)$_2$ isomers identified. Green circles correspond to bond critical points (BCP). The values of electron density and of the Laplacian of the electron density at the intermolecular BCP between water molecule and the solute are given in a.u.

Three other isomers, significantly higher in energy, are also obtained from this approach: the D5, D6 and D7 isomer located 22.5 kJ.mol^{-1}, 29.7 kJ.mol^{-1} and 31.4 kJ.mol^{-1} above the global minimum, respectively.

The search for privileged directions of approaches between a BPL:(H$_2$O) complex and an additional water molecule leads to the identification of eight additional structures (S.I.2). For all these complexes, two values of De were calculated:

> a first value taking as reference the isolated BPL and the water dimer. This value, denoted $De_{(BPL+(H_2O)_2)}$, can be compared to De calculated for dihydrated complexes D1 to D7

> a second value taking as reference the isolated BPL in the dimer of water. This value, denoted $De_{(BPL+2(H_2O))}$, can be compared with De calculated for monohydrate complexes S1 to S5

All these structures are located between 16.9 and 36.0 kJ.mol^{-1} above the global minimum D1. In all the complexes obtained, the positions occupied by each water molecules are similar to those identified in the S1 to S5 isomers of the monohydrate complex. We will therefore note these complexes Sx_Sy, where Sx and Sy represent one of the isomers monohydrate. Five isomers of this type are located between 16.9 and 25.9 kJ.mol^{-1} above the global minimum D1: S1_S2, S2_S2, S1_S3, S1_S4 and S1_S5, in order of stability. It is therefore the complexes involving either the most favorable hydration site for a water molecule (S1) in combination with any of the other four identified sites, or the second most favorable hydration site. for a molecule of water (S2), with a symmetrical hydration on both planes of the BPL. The other three complexes of this type are significantly higher in energy: they are the S2_S5, S3S5 and S4S5 complexes located at 31.3, 34.7 and 36.0 kJ.mol^{-1}, respectively, above the global minimum for the dihydrate complex. In all these complexes, which can be seen as double monohydrate complexes, one finds the BCPs identified in the monohydrate complexes, and the values of ρ (BCP) seem little affected, which suggests that the presence of a second water molecule affects little the interaction between the solute and the first molecule of water.

To deepen this point, we compared the De (BPL + 2 (H$_2$O)) values for an Sx_Sy isomer to the sum of the De values for the Sx and Sy monohydrate isomers. The values presented in Table 1 confirm the absence of a cooperative or anticooperative effect for all the complexes, with the exception of the S4-S5 complex which has an anticooperative effect (+8.4 kJ.mol^{-1}).

Monohydrated isomer	De
S1	-30.2
S2	-26.2
S3	-20.7
S4	-20.2
S5	-14

Dihydrated isomer	De(Sx) + De(Sy)	$De_{(BPL+2(H_2O))}$	$\Delta De = De_{(BPL+2(H_2O))} - [De(Sx) + De(Sy)]$
S1_S2	-56.4	-55.0	1.4
S1_S3	-50.9	-50.0	0.9
S1_S4	-50.4	-48.9	1.5
S1_S5	-44.2	-46.0	-1.8
S2_S2	-52.4	-51.5	0.9
S2_S5	-40.2	-40.6	-0.4
S3_S5	-34.7	-37.2	-2.5
S4_S5	-34.2	-25.8	8.4

Table 1: Comparison of the complexation energies of the monohydrated and dihydrated complexes with the BPL, in the case of complexation by separate water molecules.

Study of the BPL:(H₂O)₃ isomers

We have shown in the previous section that the simultaneous interaction of the two water molecules of a water dimer with the solute was energetically more favorable than the interaction of two molecules of water that do not interact with each other. Other examples in the literature have shown that a water cluster forming multiple interactions with a solute is energetically preferred. However, we considered three possible directions of approach for the proposal of guessed-structures:

➢ the approach of a water molecule with respect to a complex BPL:$(H_2O)_2$,
➢ the approach of a water dimer with respect to a complex BPL:(H_2O),
➢ the approach of a water trimer with respect to the solute alone.

This latter case leads to the most stable isomers, and will be presented below. The topological features of the water trimer are highlighted in Figure8. Interestingly, two isoenergetic isomers exist for this trimer, and their topological features are similar: three electrophilic sites and three nucleophilic sites of different strength are identified. This accurate quantitative analysis is important since, in seeking the preferred directions of approaches of the two partners, it is necessary to seek to maximize interactions between complementary sites. The two isomers of the water trimer, which correspond to a tilting of the hydrogen atoms involved in the hydrogen bond network of the water cluster, have similar global topological characteristics, with some variations that must be taken into account for the proposal of guess structures.
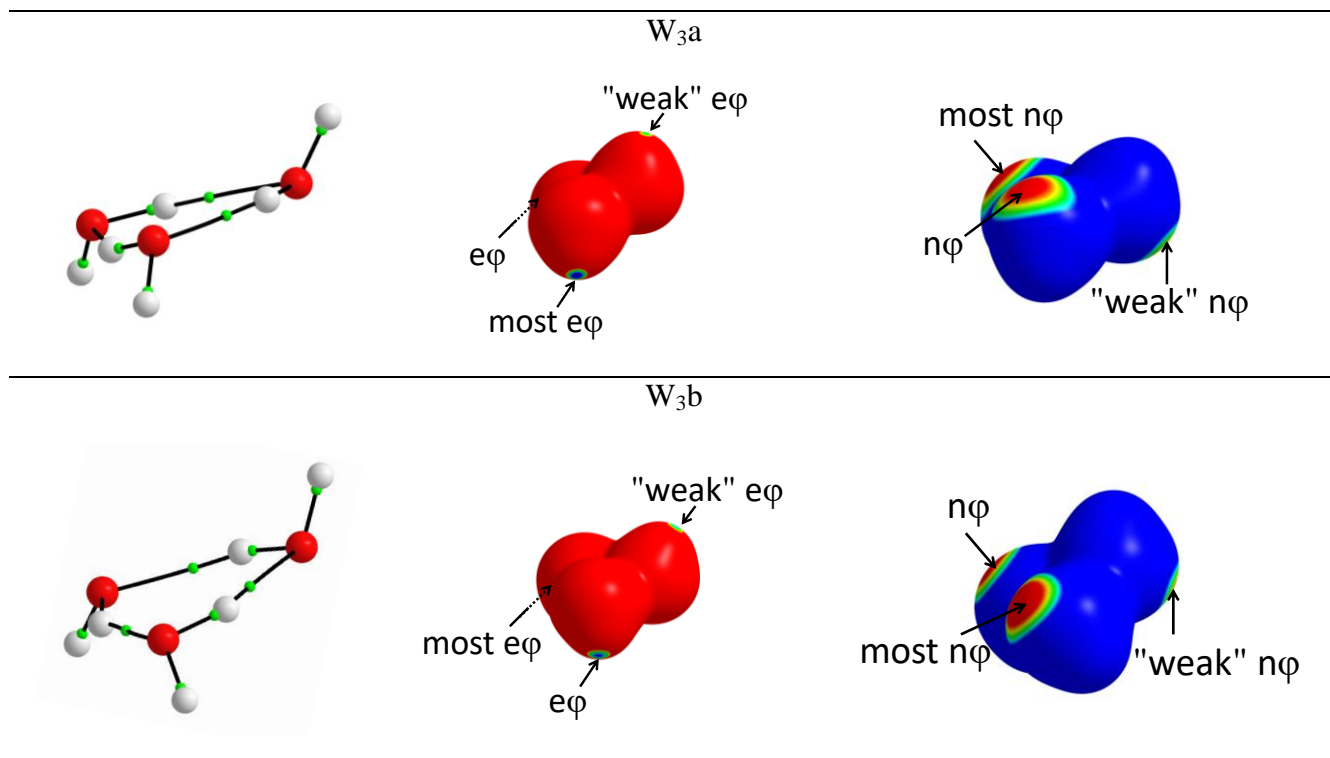

Figure 8: Identification of electrophilic(eφ) and nucleophilice (nφ) zones for both isomers of the (H₂O)₃ water cluster with the MESP analysis.

Whatever the isomer considered for the water trimer, the MESP of the water cluster is quite complementary to the one of the solute, with, on each side, an electrophilic zone, close to the hydrogen atoms not engaged in the hydrogen bonds with the other molecules. of water, and a nucelophilic zone near the oxygen atoms. The preferred directions of possible approaches are presented in Figure 9 in the case of the water trimer W_3a. Similar approaches are also predicted for the W_3b trimer approach to BPL.
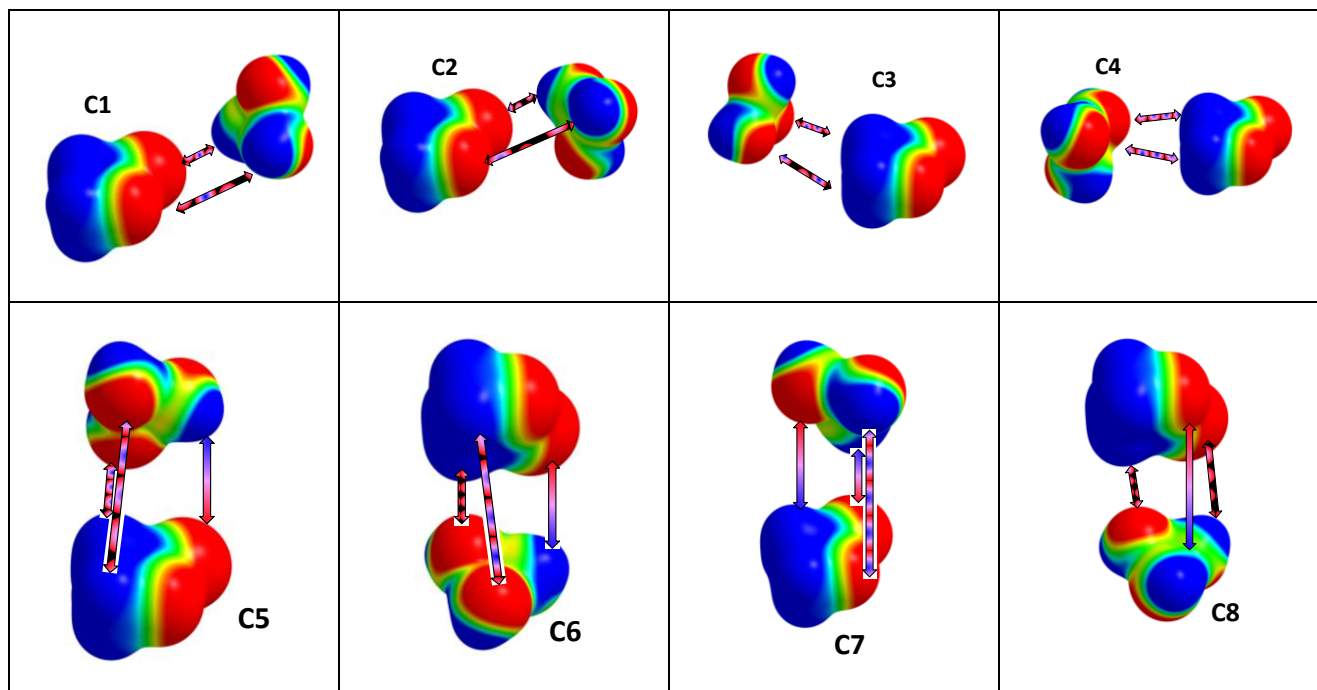
Figure 9: Identification of eight possible preferred approach directions of the T1 isomer of the water trimer relative to the BPL based on the maximization of interactions between complementary sites. Electrophilic (respectively nucleophilic) regions are colored in blue (respectively red).

The most stable complexes obtained from these initial structures are presented in S.I.3: Seven quasi iso-energetic isomers, denoted T1 to T7 in their orders of stability, were identified on the energy potential surface of the BPL:$(H_2O)_3$ complex, with $\Delta E < 5.0$ kJ.mol^{-1} at the MP2/aVDZ level of theory. Even if the uncertainty of the method does not make it possible to discern these six isomers, it is noted that the T1 isomer is that which has been highlighted on the basis of the experimental observables.

Search of the Formamide:(H_2O) isomers

MESP analysis shows that the main nucleophilic center is located near the oxygen atom of the carbonyl function. A weak pi-hole can be located on both sides of the plane of the molecule, above and below the nitrogen atom. The MESP analysis also reveals a fundamental difference in the nature of the hydrogen: the hydrogen atom opposite to the carbonyl function provides the most electrophilic site of the molecule. The second hydrogen atom carried by nitrogen has an electrophilic character. On the other hand, the hydrogen atom adjacent to the oxygen atom is not electrophilic.

As in the case of the study of the previous solute, it is interesting to complete this study by an analysis of the ELF function, to locate and characterize the free pairs carried by the oxygen and

nitrogen atoms, in the Lewis' sense. This study is presented in Figure 10. It can be seen that the two free doublets carried by the nitrogen atom contain less than one electron (0.84 and 0.76 e), suggesting possible interactions with an electrophilic or nucleophilic site. The two oxygen doublets carried by the oxygen atom are equivalent, with a population of 2.65 and 2.76 e.
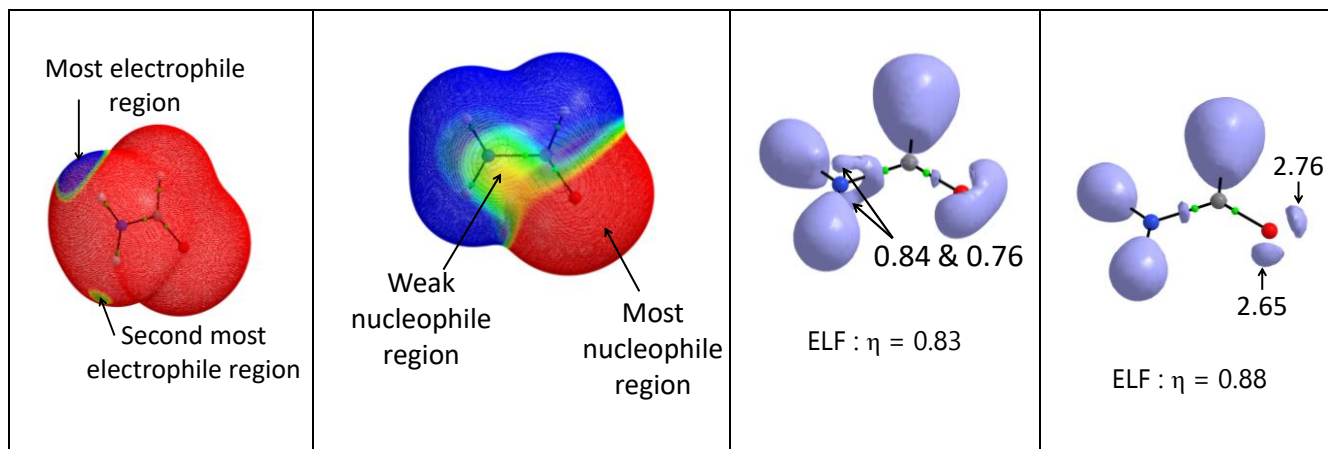

Figure 10: Identification of electrophilic(eφ, left) and nucleophilic (nφ, middle) zones on the formamide molecule with the MESP analysis. The ELF valence basins r were plotted for : η =0.83 and 0.88 (right).

This topological analysis of the solute allows us to consider six privileged directions of approach of a molecule of water (Figure 11).
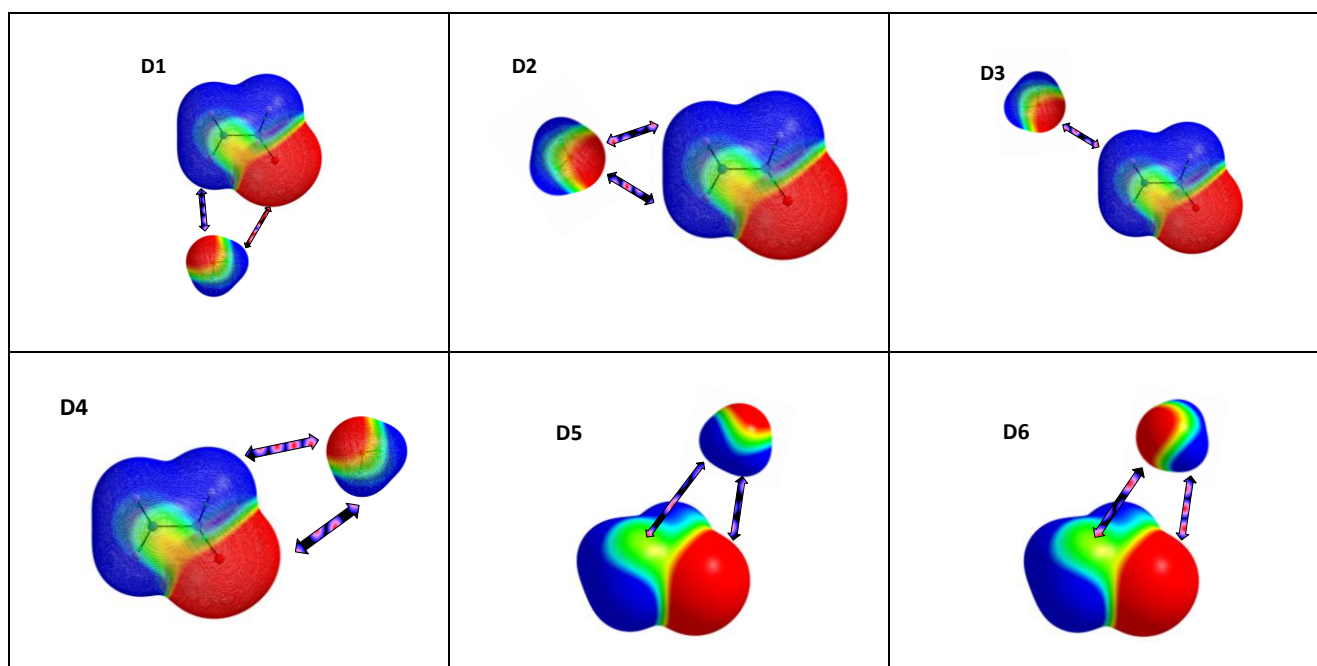

Figure 11: Identification of six possible preferred directions of approach of the water molecule relative to the formamide. Electrophilic (respectively nucleophilic) regions are colored in blue (respectively red).

From all these approaches, four stable structures are obtained after geometry optimization at the MP2/aVDZ level (S.I.4). Three isomers were experimentally identified, despite a significant energy difference of 18.3 kJ.mol^{-1} according to our calculations. When a water molecule interacts with the most nucleophilic site of formamide (Complex S1'), two complementary interactions are established between the solute and the water molecule, as attested by the QTAIM analysis. In agreement with

16

studies previously reported in the literature, our methodological approach suggests that it is the most stable isomer with a complexation energy of 42 kJ.mol^{-1}. Two isomenergetic forms are obtained (S1'a and S1'b) depending on the orientation of the dangling hydrogen atom of the water molecule relatively to the plane of the formamide. We note that only one of these two isomers was experimentally identified; this may be due to a facile inter-conversion between S1'a and S1'b.

Another isomer involving a hydrogen bonding interaction between a hydrogen atom of the water molecule and the oxygen atom of the carbonyl function of the solute was demonstrated on the basis of the experimental data by Blanco et al. it is the S2' complex, located 12.1 kJ.mol^{-1} above the global minimum according to our calculations.

Only a single isomer involving the formation of an interaction with the most electrophilic site of the solute could be obtained using our methodology, in agreement with the data previously reported in the literature for this system: S3' isomer, located 18.3 kJ.mol^{-1} above the global minimum. This approach is therefore proving to be energetically unfavorable. This can be explained by a lower BCP electron density than those identified in the S1' and S2' isomers (ρ (BCP) = 0.021), and the absence of additional intermolecular interaction that could stabilize this complex.

Study of the Formamide:(H$_2$O)$_2$ isomers

As in the case on the di-hydration of the BPL, two different approaches should be considered in search for Formamide:(H$_2$O)$_2$ isomers (Figure 12):

➢ the approach of a water dimer relative to the formamide (approaches E1 to E3)
➢ and the approach of a water molecule relative to a monohydrated complex (approaches E4 and E5).

As already underlined in Figure 2, the MESP of the water dimer and of the formamide presents some similar features. The electrophilic and neighboring nucleophilic sites identified at the surface of the water dimer will preferentially interact with two complementary sites of the solute. This implies that the oxygen atom of the carbonyl function will be engaged in any interaction with the water dimer; additionally a hydrogen atom of the aldehyde function or the amine function will interact with a nucleophilic site of the water dimer. This leads to three possible preferential approaches of a water dimer relative to the formamide solute : the E1, E2 and E3 approaches presented in S.I.5. It is worth underlying that none of these approaches allows a simultaneous interaction of the water dimer with the most electrophilic and the most nucleophilic site of the formamide. Such interaction with the two most reactive sites of the solute are achievable only when the approach of a water molecule relative to a monohydrated isomer is considered. The E4 approach relies on the interaction of a water molecule with the most nucleophilic site of the S1' isomer. Conversely, the E5 approach relies on the interaction of a water molecule with the most electrophilic site of the S1' isomer.
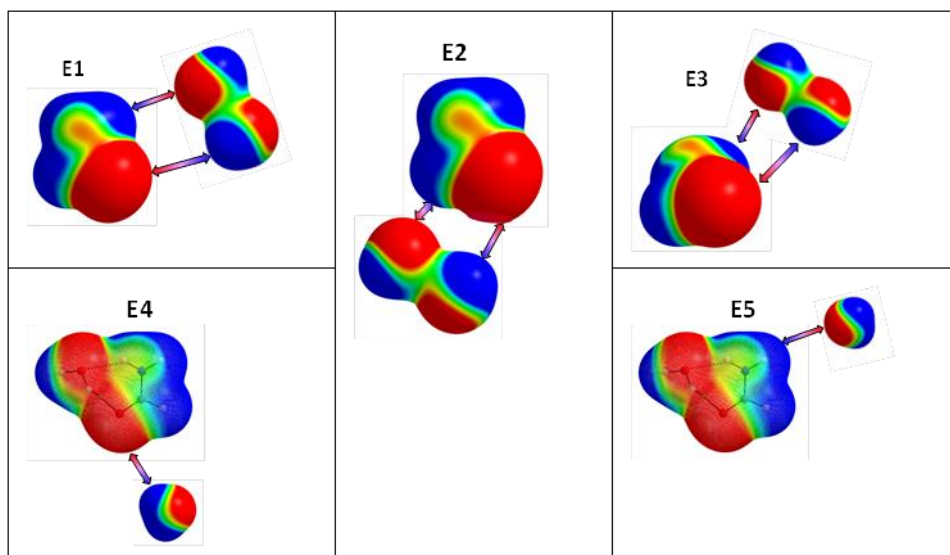
Figure 12: Identification of three possible directions of approach of a water dimer relative to the formamide (E1 - E3) and two possible directions of approach of a water relative to a monohydrated complex (E4,E5) from the MESP analysis of the isolated partners. Electrophilic (respectively nucleophilic) regions are colored in blue (respectively red).

Six stable isomers are obtained after geometry optimization from these guess structures (S.I.5). Isomers are ranked by order of their stability, from the most stable one (denoted D1') to the less stable one (denoted S1'_S3'). The most stable isomer (denoted D1') consists in a water dimer interacting with the solute. The complexation energy assicated with the formation of this complex is as large as 68.2 kJ.mol^{-1}.

Contrary to the case of the BPL:$(H_2O)_2$ isomers, the D'2 isomer for the Formamide:$(H_2O)_2$ complex is found noticeably higher in energy, 18.6 kJ.mol^{-1} above the global minimum

Almost isoenergetic to the D'2 isomer are the S'1_S'2a and S'1_S'2b isomers, located 19.0 kJ.mol^{-1} above the global minimum. Three water-solute interactions are identified in these isomers. These two isomers only differ by the position of the dangling hydrogen atom of the water molecule relatively to the formamide.

In none of these isomers, the most electrophilic site of the solute is interacting with a water molecule. Located 23.6 kJ.mol^{-1} above the global minimum, the S'1_S'3a and S'1_S'3b isomers are the only ones for with a hydrogen bonding involving the most electrophilic hydrogen atom of the amine group is identified within the QTAIM framework. Once again, the S'1_S'3a and S'1_S'3b isomers only differ by the position of the dangling hydrogen atom of the water molecule relatively to the formamide.

The concept of preferred approach directions applied to the identification of the Formamide:$(H_2O)_3$ isomers

Let us now apply the same methodology for the formation of the formamide complex trihydrate. The formation of Formamide:$(H_2O)_3$ complexes involving a water dimer and a water monomer instead of a water trimer were investigated in the literature. It was shown that such isomers are not the most stable ones. This is the reason why we did not reinvestigate these isomers in the present study. These earlier theoretical findings will be further discussed in the Part" Discussion: Comparison with other

theoretical and experimental investigations on the micro-hydration of formamide". From the MESPs of the water trimer on the one hand and the formamide on the other hand, four preferred approach directions can be expected, as depicted in Figure 13. In the F1 approach, the water trimer is located in the same plane as the solute, whereas the water cluster is orthogonal to the solute in the three other approaches. In all the approaches but F3, the oxygen atom of the carbonyl group acts as a hydrogen bond acceptor.
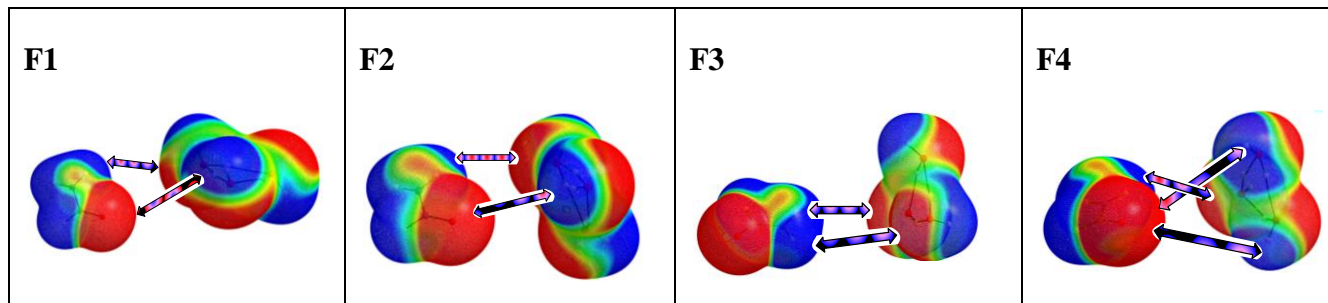

Figure 13: Identification of four possible directions of approach of a water trimer relative to the formamide from the MESP analysis of the isolated partners. Electrophilic (respectively nucleophilic) regions are colored in blue (respectively red).

Experimentally, only one isomer was observed. However, our methodology leads to the identification of five trihydrated isomers within a 0 - 10 kJ.mol^{-1} range above the global minimum S.I.6). Isomers are ranked by order of their stability, from the most stable one (denoted T1') to the less stable one (denoted T5').

Two water-solute interactions are identified in the most stable T1' isomer: the oxygen atom of the carbonyl function acts as a hydrogen bonding acceptor regarding to the hydrogen atom of a water molecule. Conversely, a hydrogen atom of the amine group acts as a hydrogen bonding donor regarding to an oxygen atom of another water molecule. The cycle between the water trimer is broken. On the other hand, all three molecules of water and formamide form another cycle, greater in the image of the water pentamer, as emphasized by Blanco et. al. A homodromic network of hydrogen bonds is formed.

The T2' isomer is located 2.6 kJ.mol^{-1} above the global minimum at the MP2/aVDZ level of theory. An opening of the water trimer is also observed in this trimer. However, each water molecule forms a hydrogen bonding with the solute, and another hydrogen bonding with one or two water molecules.

In the T3'isomer, both hydrogen atoms of the amine group are involved in hydrogen bondings with water molecules. An early stage of a water-crown is formed around the solute. This isomer is located 7.4 kJ.mol^{-1} above the global minimum at the MP2/aVDZ level of theory.

The T4' isomer shares some of the characteristics of the T1' isomers, with the formation of two water-solute interactions. The oxygen atom of the carbonyl function acts as a hydrogen bonding acceptor regarding to the hydrogen atom of a water molecule. Conversely, a hydrogen atom of the amine group acts as a hydrogen bonding donor regarding to an oxygen atom of another water molecule. The water ring is preserved. This isomer is located 8.2 kJ.mol^{-1} above the global minimum at the MP2/aVDZ level of theory.

The T5' isomer shares some of the characteristics of the T2' isomers, with the formation of three water-solute interactions: each water molecule forms a hydrogen bonding with the solute, and another hydrogen bonding with one or two water molecules.

On the basis of spectroscopic features, the T1' isomer was the only one that was identified by Blanco et al. This indeed corresponds to the most stable isomer predicted using our methodological approach.

Discussion: On the strength of the hydrogen bonding network in water clusters compared to water-solute hydrogen bonding network

For all the complexes considered ($S:(H_2O)_{1-3}$, S = BPL or Formamide), different stable isomers can be identified by means of the analysis of the electrostatic potential of the isolated partners. In all the cases considered, the most stable isomers obtained using our methodology correspond to the isomers experimentally identified by high resolution microwave spectroscopy. This suggests that long-distance electrostatic interactions between isolated species:

➢ either play a critical role in the formation and stability of micro-hydrated complexes,
➢ or are a suitable descriptors for helping the chemist in identifying the most relevant isomers.

For monohydrated complexes, the isolated species to be considered should be the solute and the water molecule. On the other hand, for species $S:(H_2O)_n$, different isolated partners can be considered: S and $(H_2O)_n$ or $S:(H_2O)_m$ and $(H_2O)_{n-m}$. For the two solutes considered here, the initial structures constructed from the study of the MESP of S and $(H_2O)_n$ make it possible to obtain, after optimization of geometry, the most stable isomers. This phenomenon was referred to as "segregation of water molecules" in the literature. However, depending on the nature of the solute, it cannot be ruled out that, in some cases, the approach of a water dimer to a monohydrate complex, or the approach of a water molecule to a complex dihydrate, lead to more stable isomers than approaching a water trimer on isolated solute.

From an energetic point of view, the complexation energy is between 30 and 70 kJ.mol^{-1} for the most stable isomers of the $S:(H_2O)_n$ complexes. Furthermore, for a given number of water molecules, the complexation energy in the most stable isomer is systematically greater with formamide than with BPL (Figure 14). In addition, for a given solute, the complexation energy with the water dimer is greater than in the trimer. How to explain the segregation of water molecules, the particular stability of dihydrated complexes, and the increased stability of Formamide:$(H_2O)_n$ complexes compared to their BPL:$(H_2O)_n$ analogues?
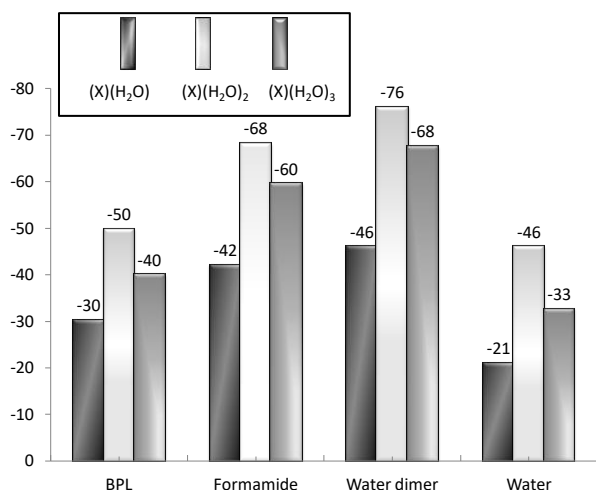
Figure 14: Complexation energies of the most stable S:$(H_2O)_n$ isomers, S = BPL or Formamide, n = 1 - 3, calculated relative to the isolated solute and the most stable $(H_2O)_n$ clusters.

To increase the set of complexes herein studied, $(H_2O):(H_2O)_n$ and $(H_2O)_2:(H_2O)_n$ complexes were also considered. A caveat is necessary in respect of the structures of these complexes: $(H_2O):(H_2O)_n$ and $(H_2O)_2:(H_2O)_n$ complexes do not necessarily adopt the geometry of the most stable $(H_2O)_{n+1}$ and $(H_2O)_{n+2}$ isomers, respectively. As an example, the $(H_2O):(H_2O)_3$ complex will lead to a pyramidal geometry, whereas the $(H_2O)_2:(H_2O)_2$ complex presents the square planar structure of the most stable water tetramer isomer (Table 2). From the isolated partners, the MESP analysis led to the identification of the most nucleophilic and electrophilic sites, and the initial structures of the $(H_2O):(H_2O)_n$ and $(H_2O)_2:(H_2O)_n$ complexes were built up as was detailed for the formamide and BPL.
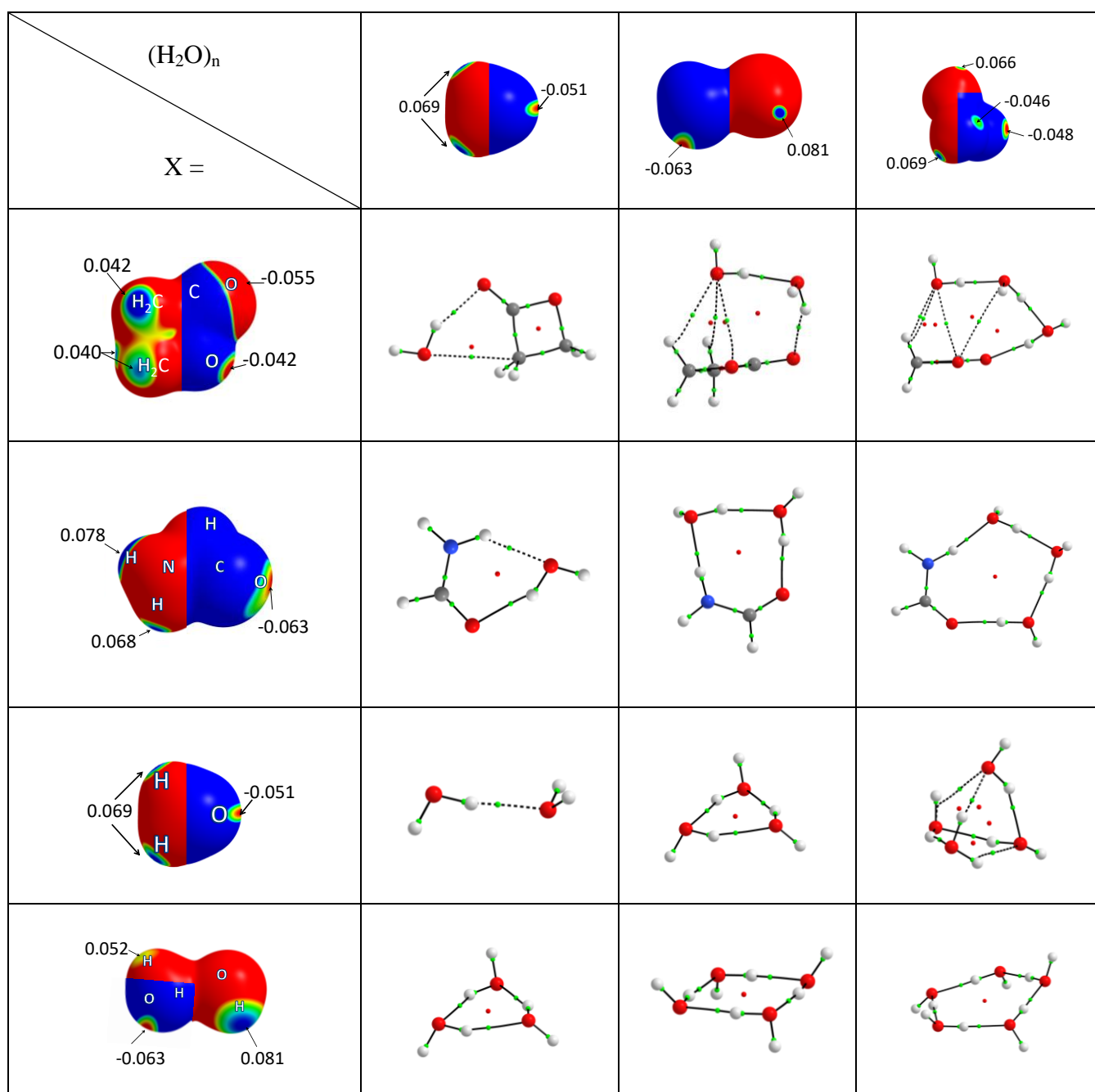
(H₂O)ₙ / X =			

Table header: $(H_2O)_n$ ... $X =$

Row 1 (top, $(H_2O)_n$ MESP images):
- 0.069, −0.051
- −0.063, 0.081
- 0.066, −0.046, −0.048, 0.069

Row 2:
- 0.042, −0.055, H₂C, C, O; 0.040, H₂C, O, −0.042

Row 3:
- 0.078, H, H, N, C, O; 0.068, −0.063

Row 4:
- H, −0.051, 0.069, O, H

Row 5:
- 0.052, H, O, H, O, H; −0.063, 0.081

Table 2 : Combined MESP pictures of the (H2O)₁₋₃ and X = BPL, Formamide, (H₂O) and (H₂O)₂ partners as well as the most stable isomers of the X: (H₂O)₁₋₃ complexes. The values of the MESP are given in a.u.

Finally, the complexation energies of the $(H_2O):(H_2O)_n$ and $(H_2O)_2:(H_2O)_n$ complexes are presented in Figure 14, in addition to the complexation energies of the $BPL:(H_2O)_n$ and $Formamide:(H_2O)_n$ complexes. In all the cases, the complexation energy is greater with the water dimer, and smaller with the water monomer. Furthermore, for a given value of n, the complexation energy is smaller for $X = H_2O$ and BPL compared to $X = (H_2O)_2$ and Formamide. Following the above presented methodology, we were interested in the detailed analysis of the MESPs of all the involved partners. Would the MESP analysis be useful for a quantitatively prediction of the strengths of the interactions, and not only a qualitative tool ? In order to test for that possibility, Table 2 presents

combined pictures with the most electrophilic and the most nucleophilic sites gathered into a single picture for each considered species, and the values of the electrophilic and nucleophilic sites leading to intermolecular interactions are presented in Table 3.

n =	X =	$n\varphi(X) \to e\ \varphi(H_2O)$	$e\varphi(X) \to \varphi(H_2O)$	$E_{complexation}$
	H_2O	$-0.051 \to 0.0686$		21
1	BPL	$-0.055 \to 0.0686$	$0.041 \to -0.051$	30
	Formamide	$-0.063 \to 0.0686$	$0.0675 \to -0.051$	42
	H_2O	$-0.051 \to 0.081$	$0.0686 \to -0.0635$	46
2	BPL	$-0.055 \to 0.081$	$0.041 \to -0.0635$ $0.0396 \to -0.0635$	49
	Formamide	$-0.063 \to 0.081$	$0.0675 \to -0.0635$	68
	BPL	$-0.055 \to 0.069$	$0.041 \to -0.048$ $0.0396 \to -0.046$	40
3	H_2O	$-0.051 \to 0.069$	$0.0686 \to -0.048$	51.5
	Formamide	$-0.063 \to 0.069$	$0.0675 \to -0.048$	59

Table 3 : Quantitative identification of the electrophillic and nucleophilic sites of the partners leading to intermolecular interactions in the X:(H$_2$O)$_n$ complexes (in a.u.), as well as the complexation energy in kJ.mol^{-1}.

For the $(H_2O)_n$ species, these combined pictures reveal that the water dimer has simultaneously more nucleophilic and more electrophilic sites than the water trimer and water monomer. The most nucleophilic site and the most electrophilic site of the water dimer are simultaneously involved in the most stable X:(H$_2$O)$_2$ isomers. The electrophilic and nucleophilic sites available on the water molecular surface are weaker than the ones of the water dimer, and this may explain why the complexation energies of the X:(H$_2$O) species are smaller than the ones of their X:(H$_2$O)$_2$ counterparts. The most nucleophilic site available on the molecular surface of the water trimer is weaker than the ones on the water and water dimer surfaces. Furthermore, the most electrophilic available on the water dimer molecular surface is as strong as the one of the water molecule. This may explain why the complexation energies of the X:(H$_2$O)$_3$ species are smaller than the ones of their X:(H$_2$O)$_2$ counterparts. On the other hand, multiple interaction sites of the water trimer are simultaneously involved in the intermolecular interaction with the X species, and this may explain why the complexation energy for the X:(H$_2$O)$_3$ complexes are greater than the ones for the corresponding X:(H$_2$O) species.

As far as the nature of the X partner is concerned, Figure 14 shows that, for a fixed value of n, the complexation energies increases in the following order: (H$_2$O) < BPL < Formamide < (H$_2$O)$_2$. Yet in Table 3, the strengths of the electrophilic and nucleophilic sites available on the molecular surface increases from (H$_2$O) to BPL to Formamide and to (H$_2$O)$_2$. Thus, the strengths of of the electrophilic and nucleophilic sites available on several A and B partners can serve as a guidance to evaluate which combination of A:B partners may lead to the isomer displaying the strongest complexation energy.

For these isomers, the electron density at the intermolecular, water-solute BCPs linearly decreases with the increase of the intermolecular distance (Figure 15). It is interesting to note that such linear regressions could be further connected with topological and energetic features of the hydrogen bondings.[66, 67]
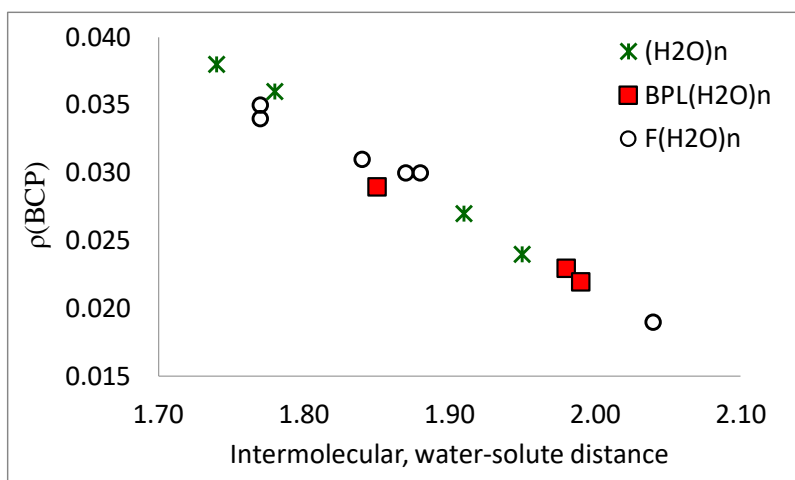

Figure 15: Illustration of the linear decrease of the electron density at the intermolecular, water-solute BCP as the function of the intermolecular distance.

Discussion: Comparison with other theoretical and experimental investigations on the micro-hydration of formamide

As already pointed out in the introduction, the micro-hydration of formamide was theoretically investigated using an alternative method based on the topological analysis of the MESP function in 2000 and 2004 by Gadre and collaborators.[53,54] Herein, we would like to compare their results, experimental finding, and the results of the present study.
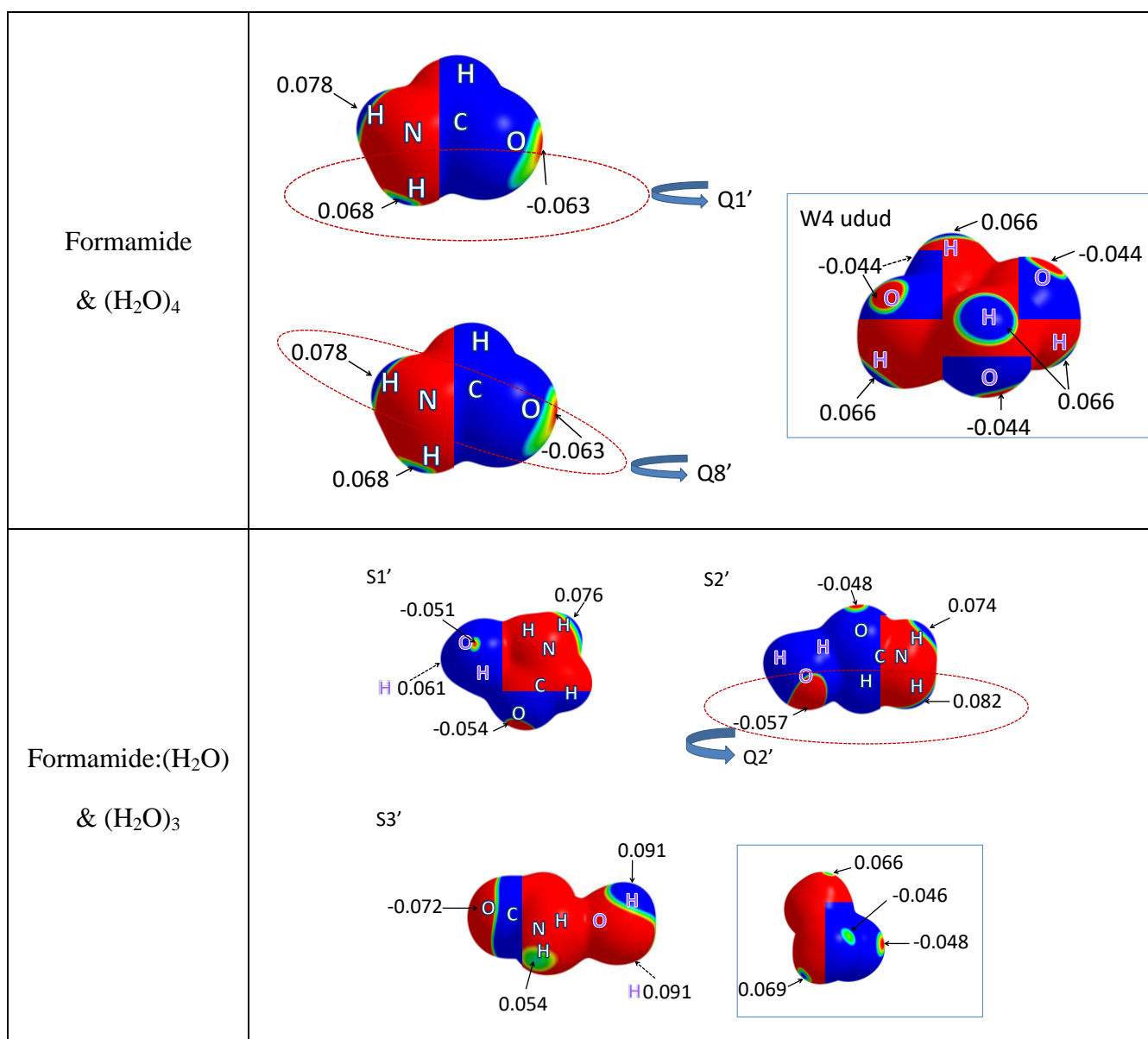
Formamide:(H_2O)		Formamide:(H_2O)$_2$		Formamide:(H_2O)$_3$	
S'1	Gadre: 0.0 Exp: Identified This w.:0.0	D'1	Gadre: 0.0 Exp: Identified This w.: 0.0	T''1	Gadre: 0.0 Exp: Identified This w.: 0.0
S'2	Gadre: 12.1 Exp: Identified This w.: 12.1	D'2	Gadre: N.P. Exp: This w.: 18.6	T'2	Gadre: N.P. Exp: N.S. This w.: 2.6
S'3	Gadre: N.P. Exp: Identified This w.: 18.3	S'1S'2	Gadre: 16.3 Exp: N.S. This w.: 19.0	T'3 - T'5	Gadre: N.P. Exp: N.S. This w.: 7.4 → 9.5
		S'1S'3	Gadre: N.P. Exp: N.S. This w.: 23.6	D'1 S'3	Gadre: 7.1 Exp: N.S. This w.: N.S.
				D'1 S'2	Gadre: 10.8 Exp: N.S. This w.: N.S.

Table 4 : Comparison of previous theoretical identification of micro-hydrated complexes of formamide, experimental findings,[17,19] and the present results. Relative energies were calculated at the RHF/6-31G** (Gadre et al., reference 53) and MP2/aVDZ (This work, denoted "This w.") levels, and are given in kJ.mol^{-1}. N.P. and N.S. stand for "not predicted" and "not studied", respectively.

In the case of the monohydrated complex, the experimental investigations led to the identification of three isomers,[17] in full agreement with the results we obtained. Two of these isomers were also predicted by the MESP approach of Gadre et al.[53] From experimental investigations, only one isomer (denoted D'1 in Table 4) was identified for the Formamide:(H_2O)$_2$ complex. This D'1 isomer was also theoretically identified as being the most stable by Gadre et al.[53] as well as in the present study. We note in passing that we identified an additional isomer (D'2) in which a water dimer is interacting with the solute. The only isomer experimentally identified for the Formamide:(H_2O)$_3$ complex (denoted T'1 in Table 4) is also predicted as being the most stable isomer by Gadre et al.[53] as well as by our methodological approach. Additionally, we identified four other isomers in which a water trimer interacts with the solute. We specifically note that the T'2 isomer lies only 2.6 kJ/mol above the global minimum. Furthermore, Gadre et al. demonstrated by their approach that isomers in which a water dimer and a water monomer are interacting with the solute are slightly less stable than these two most stable isomers. Indeed, the D'1 S'3 and D'1 S'2 isomers (denoted V$_2$ and V$_3$, respectively, by Gadre et al.) lie 7.1 and 10.8 kJ/mol above the global minimum (7.7 and 10.5 kJ/mol, respectively at the MP2/aVDZ level including the ZPE correction).

We thus note that both the Gadre's approach and our methodological approach lead to the identification of the most stable structures for Formamide:$(H_2O)_{1-3}$ complexes. Furthermore, our approach is simple, rapid, can be carried out using available softwares, and lead to the identification of all the isomers experimentally identified to date.

To further compare the results obtained using the Gadre's approach and our own method, we re-investigate the possible structures for Formamide:$(H_2O)_4$ complexes, for which no experimental data are available yet. We tried to find all the most stable isomers, by considering Formamide, $(H_2O)_4$, Formamide:(H_2O), $(H_2O)_3$, Formamide:$(H_2O)_2$, $(H_2O)_2$ as well as Formamide:$(H_2O)_3$ and (H_2O) as initial "building blocks". The identification of possible directions of approaches of the water clusters relative to the building blocks containing the formamide molecule are identified in Figure 16 in each case.
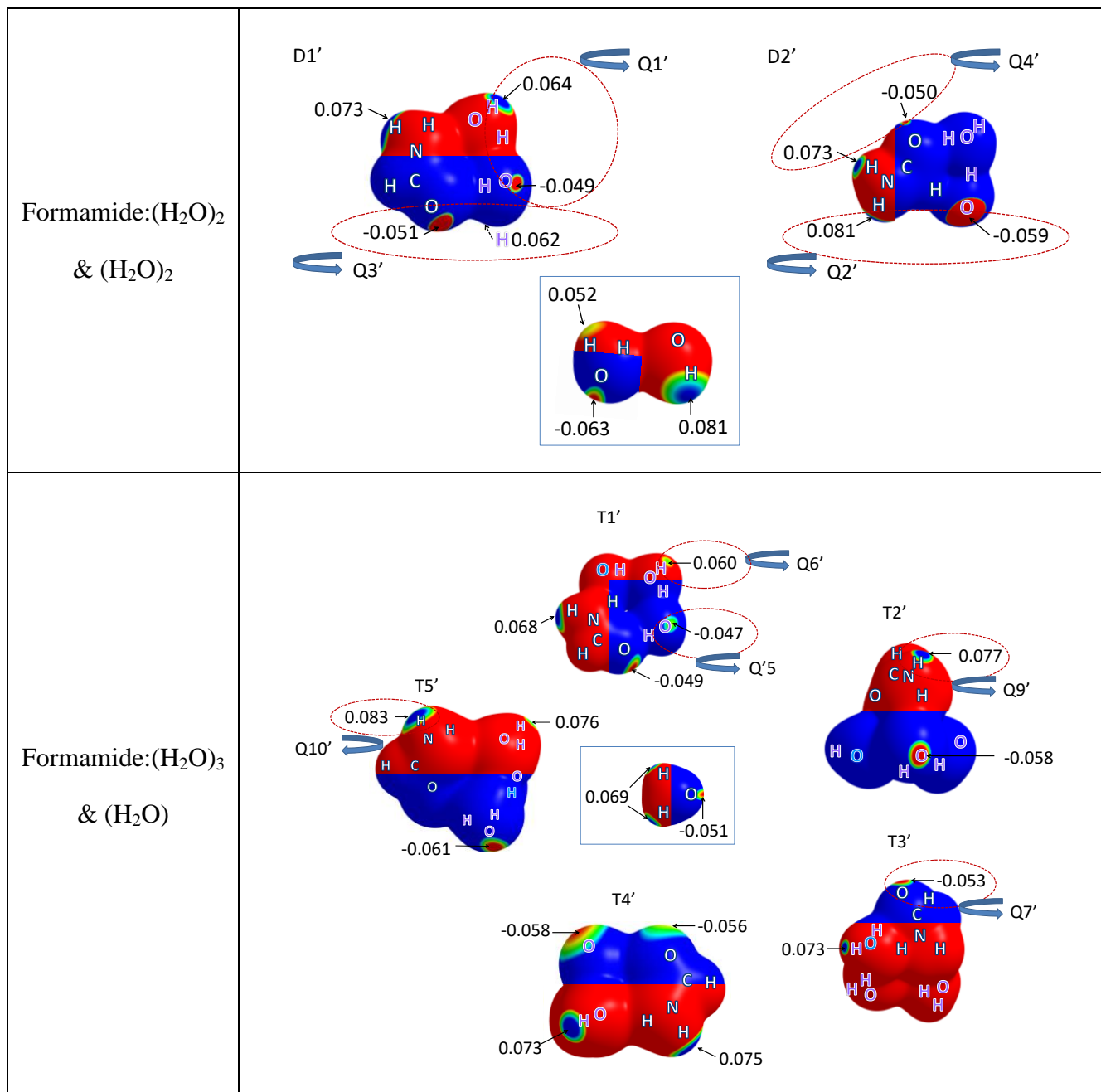
Formamide & $(H_2O)_4$	
Formamide:(H_2O) & $(H_2O)_3$	

Figure 16: Identification of possible directions of approach of a $(H_2O)_m$ building block relative to a Formamide:$(H_2O)_n$ building block, with m + n = 4. Electrophilic (respectively nucleophilic) regions are colored in blue (respectively red).

Owing to the size of the systems, and to reduce computation time, geometry optimizations were carried out using the density functional theory (DFT), with the Grimme empirical dispersion correction with Becke–Johnson damping (GD3BJ).[68,69] The LC-ωPBE[70,71,72] long-range corrected functional in combination with the 6-311++G(d,p) basis set was chosen for this study. The reliability of the LC-ωPBE/6-311++G(d,p) level of theory with the GD3BJ empirical dispersion correction to correctly describe water-solute interactions was previously pointed out.[73] At this level of theory, ten isomers were identified for the Formamide:$(H_2O)_4$ complex, and the features of the intermolecular, water-solute BCPs are given in S.I.7. Isomers are ranked by order of their stability, from the most

stable one (denoted Q1') to the less stable one (denoted Q10'). For all the isomers, the energy relative to the global minimum (denoted $\Delta(E)_{DFT}$) was calculated. For the isomers Q2' to Q6', $\Delta(E)_{DFT}$ was calculated to be less than 25 kJ.mol^{-1}, and MP2 calculations were carried out to further characterize the relative energy of the complexes. Thus, for the Q1', Q2', Q3', Q4', Q5' and Q6' isomers, the relative energy and the relative free energies calculated at the MP2/aVDZ level (denoted $\Delta(E)_{MP2}$ and (denoted $\Delta(G)_{MP2}$, respectively) are given in S.I.7. In all the optimized isomers, both the amine and carbonyl groups are involved in hydrogen bondings with water molecules. In their study, Gadre et al. identified the isomer herein denoted Q1'. We note however that two other isomers are almost isoenergetic with the global minimum at the levels of theory herein chosen: Q2' and Q3'. In these three isomers, four water-water BCPs are identified, in addition with water-solute BCPs. Further experimental investigations looking for these isomers might be very much interesting.

In the Q4' isomer, two separate water dimers are interacting with the solute. This isomer is notably higher in energy than the global minimum, with $\Delta(E)_{DFT}$=15.7 kJ.mol^{-1}, $\Delta(E)_{MP2}$=14.0 kJ.mol^{-1}, and $\Delta(G)_{MP2}$=19.5 kJ.mol^{-1}.

In the Q5' and Q6' isomers, a water trimer and a water molecule are separately interacting with the solute. It is interesting to note that the entropic term is particularly high for both of these cases, with $\Delta(E)_{MP2}$=17.5 and 23.0 kJ.mol^{-1} and $\Delta(G)_{MP2}$=6.1 and 6.0 kJ.mol^{-1} for the Q5' and Q6' isomers, respectively.

The Q8', Q9' and Q10' isomers also consist of a water trimer and a water molecule separately interacting with the solute. These isomers are lying almost 30 kJ.mol^{-1} above the global minimum at the LC-ωPBE/6-311++G(d,p) level of theory with the GD3BJ correction.

The Q7' isomer consists of a water tetramer interacting with the solute but it is notably higher in energy than the Q1' isomer.

Finally, our results are in agreement with the ones previously obtained by Gadre et al. But additional structures that may be experimentally identified are predicted using our methodological approach.

Conclusions and perspectives

The identification of the most stable isomers of hydrated complexes involving a small solute X is not an easy task: from a theoretical point of view, the reliability of structures obtained depends (among others) on the structures initially considered for the X:(H$_2$O)$_n$ complexes, and many isomers can often be considered, even for a small value of n. Herein, we have shown that a thorough analysis of the molecular electrostatic potential (MESP) of isolated partners X and (H$_2$O)$_n$ constitutes a helpful guidance for predicting the most favorable isomers. Indeed, a quantitative analysis of the MESP on the molecular surfaces of X and (H$_2$O)$_n$ allows to identify all the electrophilic and nucleophilic sites available. Taken into account the geometry of the partners, it is possible to propose several guessed structures combining some of the most electrophilic sites of one partner with some of the most mucleophilic sites of the other partner, and vice versa. It is interesting to note that, depending on geometric considerations the most electro- and/or nucleo-philic site of one of the partner may not be involved in the intermolecular interactions in the most stable isomers. Geometry optimisation was carried out on the guessed structures at the MP2/aVDZ level of theory. Five isomers for the

BPL:(H$_2$O) complex, four particularly stable isomers of the BPL:(H$_2$O)$_2$ complex, and six almost iso-energetic isomers for the BPL:(H$_2$O)$_3$ complex were identified using our methodological approach. For the BPL:(H$_2$O), BPL:(H$_2$O)$_2$, BPL:(H$_2$O)$_3$, Formamide:(H$_2$O), Formamide:(H$_2$O)$_2$ and Formamide:(H$_2$O)$_3$ complexes, the most stable isomers found using our methodological approach correspond to the one experimentally identified. In all the complexes, complementary intermolecular interactions simultaneously involving electrophilic and nucleophilic sites of both partner were identified using the QTAIM topological analysis. It was found that the complexation energy is larger for the Formamide:(H$_2$O)$_2$ complex than for the Formamide:(H$_2$O)$_{1,3}$ complexes. Similarly, the complexation energy is larger for the BPL:(H$_2$O)$_2$ complex than for the BPL:(H$_2$O)$_{1,3}$ complexes, for all the X:(H$_2$O)$_{1-3}$ complexes. This increased complexation energy with the water dimer was also identified for X = H$_2$O and (H$_2$O)$_2$, and was explained by the simultaneous involvement of particularly strong electrophilic and nucleophilic sites of the water dimer in the intermolecular interactions with the X partner. Our methodological approach could be further used to predict the geometry of micro-hydrated complexes prior to their experimental identification. The results obtained using the methodological approach herein described are in perfect agreement with the experimental data available. Furthermore, the present results are also in agreement with, and more complete than the ones obtained using other theoretical approaches.

1 Derat E, Shaik S, Rovira C, Vidossich P, Alfonso-Prieto M (2007) J Am Chem Soc 129:6346-6347

2 Vaida V (2011) J Chem Phys 135:020901

3 Li C et al. (2017) Environ Sci Technol 51:5043-5051

4 Long B, Bao J L, Truhlar D G (2016) J Am Chem Soc 138:14409-14422

5 Gerber R B, Varner M E Hammerich A D, Riikonen S, Murdachaew G, Shemesh D, Finlayson-Pitts B J (2015) Acc Chem Res 48:399-406

6 Anglada J M, Hoffman G J, Slipchenko L V, Costa M., Ruiz-Lopez M F, Francisco J S (2013) J Phys Chem A, 117:10381-10396

7 Gor, L, Asensio A, Tuñón I & Ruiz-López M F (2005) Chem Eur J 11:6743-6753

8 Weiss A K, Hofer T S, Randolf B R, Bhattacharjee A & Rode B M (2011) Phys Chem Chem Phys 13:12173-12185

9 Kiss B, Fábián B, Idrissi A, Szőri M & Jedlovszky P. (2017) J Phys Chem B 121:7147-7155

10 Tanabe K, Miyazaki M, Schmies M, Patzer A, Schütz M, Sekiya H, Sakai M, Dopfer O & Fujii, M (2012) Angew Chem Int Ed 124:6708-6711

11 Alonso J L, Cocinero E J, Lesarri A, Sanz M E, López J C (2006) Angew Chem Int Ed 45:3471-3474

12 Pérez C, Krin A, Steber A L, López J C, Kisiel Z, Schnell M (2015) J Phys Chem Lett 7:154-160

13 Mata S, Cortijo V, Caminati W, Alonso J L, Sanz M E, López J C, Blanco S (2010) J Phys Chem A 114:11393-11398

14 Vaquero V, Sanz M E, Peña I, Mata S, Cabezas C, López J C, Alonso J L (2014) J Phys Chem A 118:2584-2590

15 López J C, Sánchez R, Blanco S, Alonso J L (2015) Phys Chem Chem Phys 17:2054-2066

16 Pérez C, Neill J L, Muckle M T, Zaleski D P, Peña I, Lopez J C, Alonso J L, Pate B H (2015) Angew Chem 127:993-996

17 Blanco S, López J C, Lesarri A, Alonso J L (2006) J Am Chem Soc 128:12111-12121

18 Blanco S, Pinacho P, López J C (2016) Angew Chem 128:9477-9481

19 Blanco S, Pinacho P, López J C (2017) J Phys Chem Lett 8:6060-6066

20 Politzer P, Laurence P R, Jayasuriya K (1985) Environ Health Perspect 61, 191

21 Weiner P K, Langridge R, Blaney J M, Schaefer R, Kollman P A (1982) Proc Nat Acad Sci 79:3754-3758

22 Buckingham A D, Fowler P W (1983) J Chem Phys 79:6426-6428

23 Bader R F W, Nguyen-Dang T T, Tal Y (1981) Reports on Progress in Physics, 44:893

24 Dimitrova V, Ilieva S, Galabov B (2002) J Phys Chem A, 106:11801-11805

25 Murray J S, Brinck T, Grice M E, Politzer P (1992) THEOCHEM, 256:29-45

26 Murray J S, Politzer P (1991) J Org Chem 56:6715-6717

27 Brinck T, Murray J S, Politzer P (1992) Int J Quantum Chem 44:57-64

28 Politzer P, Murray J S, Clark T (2010) Phys Chem Chem Phys 12:7748-7757

29 Shields Z P, Murray J S, Politzer P (2010) Int J Quantum Chem 110:2823-2832

30 Adhikari U & Scheiner S (2014) Chem Phys 440:53-63

31 Scheiner S (2016) Chem Eur J 22:18850-18858

32 Scheiner S (2017) Faraday discussions 203:213-226

33 Geboes Y, Nagels N, Pinter B, De Proft F, & Herrebout W A (2014) J Phys Chem A 119:2502-2516

34 De Vleeschouwer F, Denayer, M, Pinter B, Geerlings P & De Proft F (2018) J Comput Chem 39:557-572

35 Rezaei Z, Solimannejad M & Esrafili M D (2015) Comput Theor Chem 1074:101-106

36 Zins E L & Alikhani M E (2016) Mol Phys 114:1317-1325

37 Mohan N & Suresh C H (2014) J Phys Chem A 118:1697-1705

38 Gadre S R, Babu K & Rendell A P (2000) J Phys Chem A 104:8976-8982

39 Gadre S R & Pingale S S (1998) Current Science 75:1162-1166

40 Pingale S S, Gadre S R, & Bartolotti L J (1998) J Phys Chem A 102:9987-9992

41 Kumar A, Gadre S R, Mohan N & Suresh C H (2014) J Phys Chem A 118:526-532

42 Mohan N, Suresh C H, Kumar A & Gadre S R (2013) Phys Chem Chem Phys 15:18401-18409

43 Gadre S R & Kumar A (2015) Understanding lone pair-π interactions from electrostatic viewpoint. In Noncovalent Forces (pp. 391-418). Springer, Cham.

44 Bader RFW (1990) Atoms in Molecules - A Quantum Theory, Oxford University Press, Oxford

45 Bader R F (2005) Chem Mon136: 819-854

46 Becke A D, Edgecombe K E (1990) J Chem Phys 92:5397-5403

47 Savin A, Nesper R, Wengert S, Fässler T F (1997) Angew Chem Int Ed 36:1808-1832

48 Silvi B, Savin A (1994) Nature, 371:683-686

49 Silvi B, Fourré I, Alikhani M E (2005) Chem Mon 136:855-879

50 Silvi B (2015) The Relevance of the ELF Topological Approach to the Lewis, Kossel, and Langmuir Bond Model. In The Chemical Bond II (pp. 213-247) Springer, Cham.

51 Andrés J, Berski S, Silvi B (2016) Chem Comm 52:8183-8195

52 Silvi B (2017) Theor Chem Acc 136:106_1-106_6

53 Gadre, S. R., & Kulkarni, A. D. (2000). Indian J Chem 39A:50-59

54 Kulkarni, A. D., Babu, K., Gadre, S. R., & Bartolotti, L. J. (2004) J Phys Chem A 108:2492-2498

55 Dunning TH Jr (1989) J Chem Phys 90:1007-1026

56 Kendall RA, Dunning TH, Harrison RJ (1992) J Chem Phys 96:6796-6805

57 Shields R M, Temelso B, Archer K A, Morrell T E, Shields G C (2010) J Phys Chem A, 114:11725-11737

58 Frisch MJ, Trucks GW, Schlegel HB et al (2009) Gaussian 09, revision D.01. Gaussian, Inc., Wallingford

59 AIMAll (Version 17.11.14), Todd A. Keith, TK Gristmill Software, Overland Park KS, USA, 2017 (aim.tkgristmill.com)

60 Noury S, Krokidis X, Fuster F, Silvi B (1999) Comput Chem 23:597-609

61 Legon A C, Millen D J (1987) Chem Soc Rev 16:467-498

62 Legon A C (1995) Chem Phys Lett 237:291-298

63 Fuster F, Silvi B (2000) Chem Phys 25:279-287

64 Alikhani M E, Fuster F, Silvi B (2005) Structur Chem 16:203-210

65 Singh G, Nandi A, Gadre S R, Chiba T, Fujii A (2017) J Chem Phys 146:154303

66 Ziółkowski, M., Grabowski, S. J., & Leszczynski, J. (2006) J Phys Chem A 110:6514-6521

67 Espinosa, E., Molins, E., & Lecomte, C. (1998) Chemical Physics Letters 285:170-173

68 Grimme S, Ehrlich S & Goerigk L (2011) J Comput Chem 32:1456–1465

69 Grimme S & Steinmetz M (2013) Phys Chem Chem Phys 15:16031

70 Vydrov O A & Scuseria G E (2006) J Chem Phys 125:234109

71 Vydrov O A, Heyd J, Krukau A & Scuseria G E (2006) J Chem Phys 125:074106

72 Vydrov O A, Scuseria G E & Perdew J P (2007) J Chem Phys 126:154109

73 Kalai C, Zins E L & Alikhani M E (2017) Theor Chem Acc 136:48