UNIVERSITÀ DEGLI STUDI DI MILANO

FACOLTÀ DI SCIENZE E TECNOLOGIE

DIPARTIMENTO DI CHIMICA
Doctoral School in Chemistry - XXXII Cycle

HALOGEN BONDING: A DFT AND VB INVESTIGATION

Rosario RUSSO

Tutor: Prof. Maurizio SIRONI
Coordinator: Prof. Emanuela LICANDRO

— Academic year 2019/2020 —

Contents

Introduction

During my Ph.D. I focused my attention on the study of properties of the halogen bonding from a computational point of view. Due to the growing attention towards this kind of interaction, it is important to have some computational and theoretical models able to explain and reproduce its features. Halogen bonding, indeed, has been demonstrated to be a powerful tool due to the large number of applications in different fields, ranging from biological macromolecules [1–4] to supramolecular chemistry (such as assemblies with nonlinear optical properties [5–8]), from nanomaterials and crystal engineering (like the self-assembly control of host-guest solids, with applications in liquid-crystalline [9–12], porous [13, 14], magnetic [15] and organic phosphorescent materials [16]) to materials science (such as the development of solid-state materials with peculiar electronic properties [17]). Though a plethora of theoretical studies have been carried out on this interaction, there are still some open questions to be solved.

I have investigated halogen bonding by using mainly two computational approaches, that is the Density Functional Theory (DFT) and the Valence Bond Spin-Couped Theory. The first one is implemented in a number of widely used computational chemistry software programs, while the second one has been implemented by a restricted number of research groups, according to different strategies. In this thesis, after an introduction on the main properties of halogen bonding, I will briefly describe the Spin-Coupled Theory formulated on the Slater determinants [18], adopted for the present studies. The related software, developed by my supervisor Prof. M. Sironi, was able to treat up to 11 valence electrons. During my Ph.D. thesis, I worked in order to overcome this limitation: the software now runs on systems with up to 14 valence electrons. Finally, I will show the issues tackled and discuss the results obtained.

Chapter 1

The halogen bonding

Halogen bond is a non-covalent interaction [19] that occurs between an electrophilic region, associated with a halogen atom in a molecule, and a nucleophilic region in another, or the same, molecule: R−X$\cdots$B. The main features of halogen bonding can be understood making a comparison with hydrogen bonding:

R−H$\cdots$B hydrogen bonding
R−X$\cdots$B halogen bonding

H is the hydrogen atom involved in hydrogen bonding, and X is the halogen atom involved in halogen bonding. In both cases, B acts as Lewis base (hydrogen/halogen bond acceptor) towards the electron poor species R−H or R−X (acting as Lewis acid, the hydrogen/halogen bond donor). So the halogen bond donor accepts electrons while the halogen bond acceptor donates electrons. A simple model, formulated by Politzer, can explain the origin of such kind of interaction [20]: the R−X electronic structure can be described as $s^2p_x^2p_y^2p_z^1$ assuming z in the bond direction (see Fig. 1.1).

Compared to the spherical distribution of electron density around the isolated halogen atom, a reduction in the electron density distribution is observed outwards the covalent R−X region, so this region can interact with the halogen bond acceptor along the R−X direction. The electrostatic potential (ESP) plotted on an isosurface of electron density shows a positive spot along the extension of the R−X bond (the so-called σ hole) and a negative belt around the halogen, perpendicular to the R−X bond. Such anisotropy of the ESP explains the strong dependence of halogen bond on the R−X$\cdots$B angle.

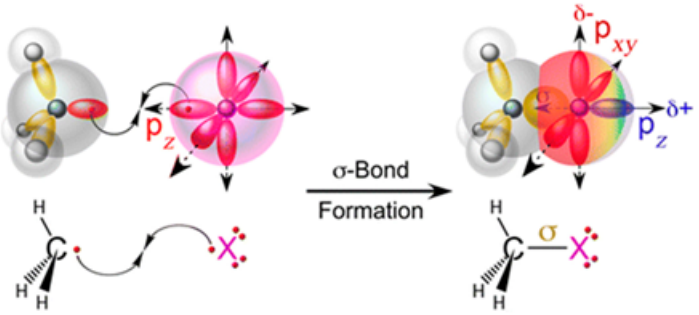

Figure 1.1: The Politzer model to explain the origin of the halogen bond.

A minimum in the potential energy (maximum interaction energy) appears at 180°, thus imposing a linear geometry for the interaction [21] (Fig. 1.2).

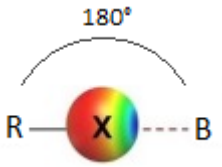

Figure 1.2: Geometrical disposition of atoms in a halogen bonding.

Halogens participating in halogen bonding include iodine, bromine, chlorine, and in rare cases (when R is a strongly electron acceptor group) fluorine. The strength of the interaction decreases with decreasing the polarizability and increasing the electronegativity of the halogen, according to the following order: I > Br > Cl ≫ F.

Due to the anisotropic distribution of electron density, a halogen atom could act as both Lewis acid and base. Therefore, two or more halogen centers could form a halogen···halogen (X···X) bond [22]. Analysis of the Cambridge Structural Database (CSD) [23] reveals two main approach geometries for such interaction, named type I ($\theta_1 \approx \theta_2$) and type II ($\theta_1 = 180°$, $\theta_2 = 90°$), where θ_1 and θ_2 are the two R−X···B angles (see Fig. 1.3) [24]. It is however to be noted that only type II geometries are indicative of a true halogen bond, that is an electrostatic interaction between a halogen atom acting as donor (that on the left in Fig. 1.3) and a halogen atom acting as acceptor (that on the right). The σ hole of the former points toward the negative belt of the latter. Type I interactions are instead dispersive inter-

actions, more frequently observed in the case of the lighter halogens (F and Cl) [25].

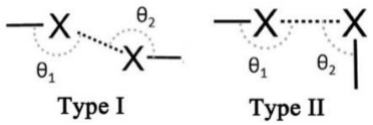

Figure 1.3: The stable geometries for the X$\cdots$X interactions: type I and type II.

Single-crystal X-ray played a key role in the identification of atoms involved in halogen bonds. The R$-$X bond distance afforded by X-ray analyses, indeed, is an informative indication of the halogen bonding presence. This distance is typically longer in the halogen-bonded adduct than in the pure halogen bond donor [26]. Other features afforded by X-ray analyses used as structural information to assess the halogen bonding presence are the donor$\cdots$acceptor separation and the angle between the covalent and non-covalent bonds around the halogen atom [27,28]. Other information afforded by X-ray analyses on the halogen bond acceptor shows that the σ hole on the electrophilic halogen atom enters preferentially along the local symmetry axis of a π-orbital when double or triple bonds or aromatic systems are the electron density donors [29,30]. These analyses allow the ranking of halogen bond donor sites mentioned (I $>$ Br $>$ Cl $\gg$ F). In order to have an indicator of the interaction strength, the normalized contact N_c is defined. This quantity is the ratio $\frac{D_{XB}}{r_X+r_B}$, where D_{XB} is the experimental distance between the halogen bond donor atom X and the halogen bond acceptor atom B and r_X and r_B are the corresponding vdW radius for the halogen bond donor and vdW radius for the halogen bond acceptor, respectively. N_c allows distances between different interacting sites to be compared, thus being more useful than the halogen bond distance itself. The N_c values of some halogen-bonded crystal structures involving dihalogens and interhalogens [31–36] appear in Table 1.1.

I_2	0.80		BrI	0.68
Br_2	0.75	0.97	ClI	0.66
Cl_2	0.82	0.94	ClF	0.88

Table 1.1: The typical N_c values of some halogen-bonded crystal structures involving dihalogens I_2, Br_2, Cl_2 and interhalogens BrI, ClI, ClF.

Chapter 2

Valence Bond Spin-Coupled Theory

2.1 Introduction

The energy levels of a molecular system of N electrons are predicted solving the Schrödinger equation

$$\hat{H}\Psi = E\Psi \tag{2.1}$$

where, within the Born-Oppenheimer approximation, $\hat{H}$ is the Hamiltonian operator of the electronic system, Ψ is the wavefunction and E represents the corresponding energy. We assume that the electronic hamiltonian $\hat{H}$ does not contain spin terms, that is, it has the following structure:

$$\hat{H} = -\sum_{i=1}^{N}\left[\frac{\hbar^2}{2m_i}\nabla_i^2 + \sum_{\alpha=1}^{A}\frac{Z_\alpha e^2}{r_{i\alpha}}\right] + \frac{1}{2}\sum_{ij=1}^{N\,i\neq j}\frac{e^2}{r_{ij}} \tag{2.2}$$

where m_i is the mass of the electron i and e its charge, A represents the number of nuclei of the molecule, ∇_i the Laplace operator acting on the coordinates of the electron i, Z_α the charge of the nucleus α, $r_{i\alpha}$ the distance between the electron i and the nucleus α and r_{ij} the distance between the electrons i and j.

Let us focus our attention on the wavefunction Ψ. We require that Ψ is antisymmetric under exchange of two particles. Furthermore $\hat{H}$ has no spin

terms, so it commutes with the total spin angular momentum operator $\hat{S}^2$ and its projection along z direction $\hat{S}_z$:

$$\left[\hat{H}, \hat{S}^2\right] = \left[\hat{H}, \hat{S}_z\right] = \left[\hat{S}^2, \hat{S}_z\right] = 0 \tag{2.3}$$

The eigenvalue equations for $\hat{S}^2$ and $\hat{S}_z$ are

$$\begin{cases} \hat{S}^2 \Theta^N_{S,M} = S(S+1)\hbar^2 \Theta^N_{S,M} \\ \hat{S}_z \Theta^N_{S,M} = M\hbar \Theta^N_{S,M} \end{cases} \tag{2.4}$$

where $\Theta^N_{S,M}$ represents the spin-eingenfunctions, $S(S+1)\hbar^2$ and $M\hbar$ are the corresponding eigenvalues.

The equation (2.3) means that Ψ is a common eingenfunction of $\hat{H}$, $\hat{S}^2$ and $\hat{S}_z$, that is, the wavefunction obeys to the same equations for the spin operators (2.4):

$$\begin{cases} \hat{S}^2 \Psi_{S,M} = S(S+1)\hbar^2 \Psi_{S,M} \\ \hat{S}_z \Psi_{S,M} = M\hbar \Psi_{S,M} \\ \hat{H} \Psi_{S,M} = E \Psi_{S,M} \end{cases} \tag{2.5}$$

where the indices S and M were added to remember that $\Psi_{S,M}$ is eigenfunction of the spin operators. We have to build the general structure of the wavefunction $\Psi_{S,M}$ in order to satisfy the equations 2.5 and the antisymmetry condition.

2.2 The Spin-Coupled wavefunction

The spin-coupled wavefunction for a system of N electrons in the S spin state has the following form [37]:

$$\Psi_{S,M} = \sum_{k=1}^{f^N_S} c_{S,k} \mathcal{A} \left[\phi_1 \phi_2 \cdots \phi_N \Theta^N_{S,M,k}\right] \tag{2.6}$$

where f^N_S describes the different possible modes of coupling the individual spins, $\mathcal{A}$ is an antisymmetry operator, ϕ_i are a set of non-orthogonal mono-electronic orbitals (the spin-coupled orbitals) and $c_{S,k}$ are the spin-coupling coefficients: they determine the weight that the structure described by the eigenfunction $\Theta^N_{S,M,k}$ has in the molecular electronic structure.

Each spin-coupled orbital is expanded on a basis set of dimension M: $\phi_i = \sum_{\mu=1}^{M} b_{\mu,i}\chi_\mu$. Normally atomic or molecular orbitals are adopted as basis set, but in principle any type of function can be used.

The $c_{S,k}$ and $b_{\mu,i}$ coefficients are determined by variational analysis, minimizing the expression of the molecular energy corresponding to 2.6.

We want to stress that if we choose the spin-coupled orbitals in singlet state

$$\begin{cases} \phi_1 = \phi_2 = \tilde{\phi}_1 \\ \phi_3 = \phi_4 = \tilde{\phi}_2 \\ \vdots \\ \phi_{N-1} = \phi_N = \tilde{\phi}_{\frac{N}{2}} \end{cases} \tag{2.7}$$

and if we choose only the first spin-coupling coefficient different from zero

$$\begin{cases} c_{S,k} = 1 \\ c_{S,k} = 0, \quad k = 2, \ldots, f_S^N \end{cases} \tag{2.8}$$

the spin-coupled wavefunction coincides with the Hartree-Fock wavefunction: $\Psi_{S,M} = \Psi_{HF}$.

Among the various possibilities for $\Theta_{S,M,k}^N$, the Rumer eigenfunctions are generally adopted due to their easy chemical interpretability.

For example, let us consider the case of benzene in singlet state:

$N = 6$, $S = M = 0$

the spin-coupling factor f_S^N is equal to five, so we have five structures, that means five Rumer eigenfunctions. We can represent each of them, like is shown in Figure 2.1.

The choice of the Rumer eigenfunctions as basis set for the spin space is convenient in order to represent the spin-coupled wavefunction in terms of Slater determinants: we can write $\Theta_{S,M,k}^N$ as a linear combination of N_d primitive spin functions θ_i

$$\Theta_{S,M,k}^N = \sum_{i=1}^{N_d} d_{i,k}^S \theta_i, \quad k = 1, \ldots, f_N^S. \tag{2.9}$$

Indeed, for a system of N spins of which N^α are α and N^β are β, N_d represents the ways to arrange N^α spins α in a product of N functions:

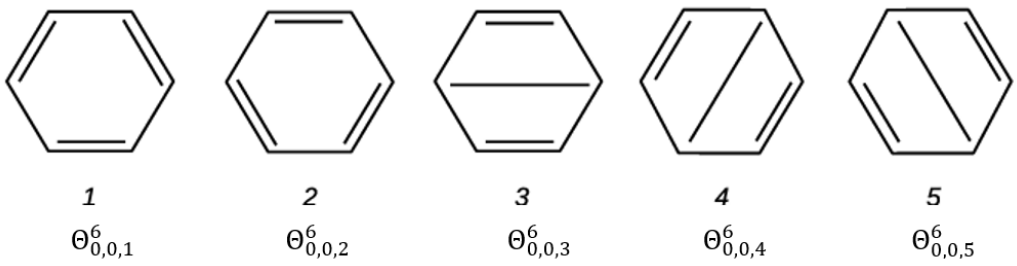

Figure 2.1: The Rumer eigenfunctions for the benzene and the corresponding structures.

$$\binom{N}{N-N^\alpha} = \binom{N}{N^\beta} = N_d. \tag{2.10}$$

This means that N_d is the number of independent primitive spin functions.

The latter ones can be written as $\theta_i = \prod_{j=1}^{N} \omega_j^i$, where $\omega_j^i = \alpha, \beta$. In the Rumer basis set we can only have $d_{i,k}^S = 0, \pm 1$, so the space-orbitals product $\phi_1\phi_2\cdots\phi_N$ in 2.6 becomes a spin-orbitals product $\varphi_1^i\varphi_2^i\cdots\varphi_N^i$, with $\varphi_j^i = \phi_j\omega_j^i$. But the antisymmetry operator $\mathcal{A}$ acting on a spin-orbitals product generates a Slater determinant; this means that the 2.6 takes the following form:

$$\Psi_{S,M} = \sum_{i=1}^{N_d} b_{S,i} U_i \tag{2.11}$$

where we defined the coefficients $b_{S,i} = \sum_{k=1}^{f_S^N} c_{S,k} d_{i,k}^S$ and U_i is the expression of the i-th Slater determinant of the spin-orbitals φ_j^i: $U_i = |\varphi_1^i\varphi_2^i\cdots\varphi_N^i|_i$ [18].

2.3 The expression of the energy functional

In this section we are going to show the expression of the energy corresponding to 2.6. For sake of simplicity, we use the atomic units: $e = m = \hbar = 1$. The hamiltonian 2.2 becomes simply

$$\hat{H} = -\sum_{i=1}^{N}\left[\frac{1}{2}\nabla_i^2 + \sum_{\alpha=1}^{A}\frac{Z_\alpha}{r_{i\alpha}}\right] + \frac{1}{2}\sum_{ij=1}^{N\, i\neq j}\frac{1}{r_{ij}} = h + \frac{1}{r_{1,2}}. \tag{2.12}$$

The energy is given by

$$E = \frac{\langle\Psi_{S,M}|\hat{H}|\Psi_{S,M}\rangle}{\langle\Psi_{S,M}|\Psi_{S,M}\rangle}. \tag{2.13}$$

Using 2.12 and the expression of the wavefunction in terms of Slater determinants 2.11, the energy becomes

$$E = \frac{1}{\mathcal{D}}\left[\sum_{ij=1}^{N}\langle\phi_i|h|\phi_j\rangle\,\mathcal{D}(i|j) + \frac{1}{2}\sum_{ijml=1}^{N}\langle\phi_i\phi_j|\phi_m\phi_l\rangle\,\mathcal{D}(ij|lm)\right] \tag{2.14}$$

where $\langle\phi_i|h|\phi_j\rangle$, $\langle\phi_i\phi_j|\phi_m\phi_l\rangle$ are the one- and two-electron integrals, $\mathcal{D}$ is the normalization integral and $\mathcal{D}(i|j)$ and $\mathcal{D}(ij|lm)$ are called *super-cofactors*: they are elements of the one- and two-electron density matrices in the spin-coupled orbitals basis set; $\mathcal{D}$ can be defined as a super-cofactor too:

- $\mathcal{D} = \sum_{pq=1}^{N_d} b_{S,p}b_{S,q}D_{pq}$ zero-order super-cofactor;
- $\mathcal{D}(i|j) = \sum_{pq=1}^{N_d} b_{S,p}b_{S,q}I_{ij}^{pq}D_{pq}(i|j)$ first-order super-cofactor;
- $\mathcal{D}(ij|ml) = \sum_{pq=1}^{N_d} b_{S,p}b_{S,q}I_{im}^{pq}I_{jl}^{pq}D_{pq}(ij|ml)$ second-order super-cofactor;

where I_{ij}^{pq} represents the spin integral: its value is $+1$ if the spin of the spin-orbital φ^i in the p-th Slater determinant is equal to the spin of the spin-orbital φ^j in the q-th Slater determinant, 0 otherwise; Furthermore

$$D_{pq} = \det S_{pq} = \langle U_p|U_q\rangle \tag{2.15}$$

$$D_{pq}(i|j) = \det S_{pq}(i|j) \tag{2.16}$$

$$D_{pq}(ij|lm) = \begin{cases} +\det S_{pq}(ij|lm) & \text{if } i<j \text{ and } m<l \text{ or } i>j \text{ and } m>l \\ -\det S_{pq}(ij|lm) & \text{if } i<j \text{ and } m>l \text{ or } i>j \text{ and } m<l \end{cases} \tag{2.17}$$

S_{pq} is the matrix of the overlap integrals between the spin-orbitals of the Slater determinants p and q. $S_{pq}(i|j)$ is the $N-1 \times N-1$ matrix obtained deleting the i-th row and the j-th column of S_{pq}; similarly $S_{pq}(ij|lm)$ is the $N-2 \times N-2$ matrix obtained deleting the rows i,j and the columns l,m of S_{pq}.

In order to obtain the spin-coupling coefficients $c_{S,k}$ in 2.6 we have to minimize the energy 2.13, that is, we have to calculate the energy derivatives with respect to the orbitals and spin variables. But the energy functional is now expressed in terms of one- and two-electron integrals and super-cofactors, so the computation of the latter ones with respect ϕ_k and $c_{S,k}$ is required.

In order to make it possible, we define the following quantities:

- $\mathcal{D}(\mu_1 \dots \mu_r|\nu_1 \dots \nu_r) = \sum_{pq=1}^{N_d} b_{S,p} b_{S,q} I^{pq}_{\mu_1\nu_1} \cdots I^{pq}_{\mu_r\nu_r} D_{pq}(\mu_1 \dots \mu_r|\nu_1 \dots \nu_r)$
  r-order super-cofactor;
- $\mathcal{D}_k(\mu_1 \dots \mu_r|\nu_1 \dots \nu_r) = \sum_{pq=1}^{N_d} b_{S,p} d^S_{q,k} I^{pq}_{\mu_1\nu_1} \cdots I^{pq}_{\mu_r\nu_r} D_{pq}(\mu_1 \dots \mu_r|\nu_1 \dots \nu_r)$
  r-order indexed super-cofactor;
- $\mathcal{D}''_k(\mu_1 \dots \mu_r|\nu_1 \dots \nu_r) = \mathcal{D}_k(\mu_1 \dots \mu_r|\nu_1 \dots \nu_r) + \mathcal{D}_k(\nu_1 \dots \nu_r|\mu_1 \dots \mu_r)$
  r-order indexed symmetrized super-cofactor

where $d^S_{q,k}$ is the weight of the q-th Slater determinant to the k-th spin eigenfunction.

A very useful recurrence relation allows to obtain all the r-order super-cofactors we need from $(r+1)$-order ones:

$$\mathcal{D}(\mu_1 \dots \mu_r|\nu_1 \dots \nu_r) = \sum_{t=1}^{N} \langle \phi_t|\phi_v \rangle \, \mathcal{D}(\mu_1 \dots \mu_r t|\nu_1 \dots \nu_r v). \tag{2.18}$$

We want to stress that the 2.18 contains only one sum from 1 to N (with $t \neq \mu_i$, $v \neq \nu_j$) and not a double sum as in the definition of super-cofactor.

Using the recurrence relation 2.18, it can be demonstrated that the derivatives of the r-th order super-cofactor with respect to the spin-coupled orbitals and the spin-coupling coefficients are given by

$$\frac{\partial}{\partial \langle \phi_k |} \mathcal{D}(\mu_1 \dots \mu_r \, | \nu_1 \dots \nu_r \rangle = \sum_s | \phi_s \rangle \, \mathcal{D}(\mu_1 \dots \mu_r k \, | \nu_1 \dots \nu_r s \rangle \tag{2.19}$$

$$\frac{\partial}{\partial c_{S,k}} \mathcal{D}(\mu_1 \dots \mu_r | \nu_1 \dots \nu_r) = \mathcal{D}'_k(\mu_1 \dots \mu_r | \nu_1 \dots \nu_r) \tag{2.20}$$

Equation 2.19 tells us that the derivative of r-order super-cofactors with respect to the spin-coupled orbitals involves $(r+1)$-order super-cofactors; equation 2.20 instead reveals us that the derivative with respect to the spin-coupling coefficients of a r-order super-cofactor involves indexed symmetrized super-cofactor of the same order.

In the spin-coupled method is necessary to calculate the second-order derivatives of the energy, in order to minimize it. Looking at the equations 2.19 and 2.20, this means that we need to compute the super-cofactors up to the fourth-order. More in detail, denoting with $\mathcal{D}^{(r)}$ the r-order super-cofactor, the derivatives involved and the corresponding super-cofactors are

- $\frac{\partial}{\partial \phi_k} E \to \mathcal{D}^{(3)}$;
- $\frac{\partial}{\partial c_{S,k}} E \to \mathcal{D}^{(2)}$, $\mathcal{D}^{(2)}_k$

for the first-order derivatives;

- $\frac{\partial^2}{\partial \phi_k \partial \phi_h} E \to \mathcal{D}^{(4)}$;
- $\frac{\partial^2}{\partial c_{S,k} \partial \phi_h} E = \frac{\partial^2}{\partial \phi_h \partial c_{S,k}} E \to \mathcal{D}^{(3)}$, $\mathcal{D}^{(3)}_k$;
- $\frac{\partial^2}{\partial c_{S,k} \partial c_{S,h}} E \to \mathcal{D}^{(2)}_{kh}$

for the second-order derivatives.

But the calculation of the two-indexes super-cofactors $\mathcal{D}^{(2)}_{kh}$ is computationally very expensive, so the spin-spin derivatives are calculated evaluating 2.13 on the basis set of the structures $\Omega_k = \mathcal{A}\left[\phi_1 \dots \phi_N \Theta^N_{S,M,k}\right]$. In this way the computation of the derivatives can be done without involving the two-indexes super-cofactors.

2.4 The steps of the Spin-Coupled software

After the description of the mathematical tools on which the Spin-Coupled Theory is based, in this section we are going to briefly explain, without entering in the programming details, the steps followed by the Spin-Coupled software, that are:

- choice of a guess for the spin-coupled orbitals and the spin-coupling coefficients in order to have a starting spin-coupled wavefunction;
- transformation of the one- and two-electron integrals from the starting basis set $\{\chi_\mu\}_{\mu=1}^{M}$ to the spin-coupled basis set $\{\phi_i\}_{i=1}^{N}$;
- computation of the energy with respect to the variational parameters;
- calculation of the super-cofactors;
- construction of the first- and second-order derivatives;
- correction of the coefficients.

Being an iterative method, the software will repeat the steps mentioned until convergence is reached.

Chapter 3

The extremely localized molecular orbitals

As it is well known, the molecular orbitals (MOs) theory to solve the Schrödinger equation is a very helpful approach to investigate molecular systems. However, such method suffers from the delocalized nature of the MOs, which are not centered on one or few nuclei but spread out on the full molecule. This feature prevents their transfer between molecules and it does not provide an intuitive interpretation of chemical bonds and other molecular properties.

Many efforts have been made to recover the concept of locality leading to the development of methods based on localized molecular orbitals (LMOs). The main drawback of the latter ones is that they show orthogonalization 'tails' beyond the localization region (Fig. 3.1a); even if the coefficients associated to these tails are small, their effect on the energy is not negligible. Furthermore the presence of tails makes their transfer difficult. The approach based on the extremely localized molecular orbitals (ELMOs) overcomes these limitations.

ELMOs are a set of molecular orbitals strictly localized only on a few atoms of a molecule, i.e. they do not show any tail (Fig. 3.1b); this characteristic makes them perfect candidates for transfer [38].

There is not a unique way to select fragments that define ELMOs, i.e. there can be different localization schemes. Generally a fragment is defined for each bond employing atomic orbitals (AOs) centered only on the atoms involved in that bond, while the ELMOs corresponding to lone pairs or core electrons belonging to a given atom use only the basis functions located

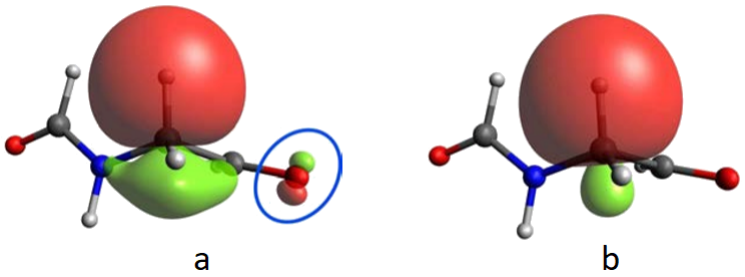

Figure 3.1: Localization of a C−H bond orbital of N-ethanalmethanamide using (a) LMOs, the orthogonalization 'tail' is highlighted in the circle; (b) ELMOs, there is no 'tail'.

on it [39]. We want to stress that different fragments can share the same AOs; for example, in order to describe bonds and lone pairs, it is possible to 'delocalize' ELMOs allowing them to use AOs centered on the neighbours, as it is shown in Figure 3.2.

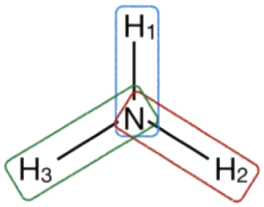

Figure 3.2: A possible localization scheme for ammonia, that is, an example of how the three H−N fragments share the same AOs of nitrogen atom.

ELMOs are defined partitioning a given basis set into subsets $\{\xi_i^f\}_{i=1}^{n_f}$ built-up with n_f functions of the total basis set, being f the number of fragments. Each subset corresponds to a 'partial' basis set for a localization center or fragment (bond, lone pair, atom). Since each ELMO is expanded on the corresponding subset, the tails are excluded. The expression for the m-th ELMO is [40]

$$\varphi_m^f = \sum_{j=1}^{n_f} c_{m,j}^f \xi_j^f \tag{3.1}$$

where the coefficients $c_{m,j}^f$ are obtained through the minimization of the energy: ELMOs are determined using the variational principle. Due to the

way they are defined, it follows that ELMOs of the same fragment are orthogonal to each other but this is no more true when ELMOs come from different fragments.

Chapter 4

Halogen bonding at work: investigated systems and results

4.1 DFT study of the anisotropy of halogen bonding

While most theoretical studies on halogen bonding are focused on the evaluation of energetic and geometric properties of the interaction at the minimum of the potential energy surface (PES), that is the R−X$\cdots$B linear approach, X-ray crystal structures often reveal a variable deviation from this geometry, as a consequence of the concomitant effect of packing forces and other specific interactions (for example, hydrogen bonds). It is therefore of interest to investigate the reliability of DFT methods in describing the regions outside the energy minimum and in particular the angular dependence of the interaction energy of halogen bonding.

To this aim, we have chosen a set of simple systems composed by two molecules interacting through halogen bonding and we have evaluated at DFT level the binding energy E_b and anisotropy of halogen bonding under distortion of the geometry. To assess the reliability of the DFT approaches we need some reference calculations, i.e., calculations performed at higher level of theory. The ideal choice would be represented by Coupled Cluster calculations, but we observed that the Møller-Plesset (MP) perturbation theory can provide results of the same quality as those provided by the Coupled Cluster method at cheaper computational cost. In particular, we have adopted the MP2.X approach as reference. In the latter one, the binding energy is

computed combining the results obtained by second- and third-order MP perturbation theory (MP2 and MP3 level, respectively) with two different basis sets according to the following formula:

$$E_b(\text{MP2.X}) = E_b(\text{MP2}|hl) + c\,[E_b(\text{MP3}|ll) - E_b(\text{MP2}|ll)] \qquad (4.1)$$

where hl designates the high level calculation (i.e. performed with the largest basis set), ll indicates the low level calculation (i.e. performed with the smallest basis set) and c is a scaling parameter which was optimized for each basis set to minimize the root-mean-square (RMS) errors in a dataset of non-covalent interactions [41].

We selected eight halogenated systems to study, which can be grouped as follows:

1. $FBr\cdots NH_3$, $FI\cdots NH_3$
2. $HCCBr\cdots NH_3$, $HCCI\cdots NH_3$
3. $FBr\cdots HCCH$, $FI\cdots HCCH$
4. $FBr\cdots C_6H_6$, $FI\cdots C_6H_6$

where the halogen bond occurs between I or Br and either the lone pair of ammonia nitrogen or (only for I and Br) the π-electron system of acetylene or benzene, the latter in T-shape configuration. We studied the variation of the binding energy with the R−X$\cdots$B angle (θ) (Fig. 4.1, 4.3) where B is either the N atom of ammonia or the center of mass of acetylene or benzene. Moreover, for the last two groups, also the translation of RX along the symmetry axis of acetylene (Fig. 4.2) or, for benzene, along the axis through the center of benzene and one of the carbon atoms of the ring (Fig. 4.4) has been investigated.

4.1.1 Computational details

All the quantum mechanical (QM) calculations were performed using the Gaussian16 quantum chemistry software [42] *in vacuo*. All the geometries were optimized using second-order Møller-Plesset (MP2) perturbation theory with aug-cc-pVTZ basis set. The binding energies have been computed using:

- restricted Hartree-Fock (RHF), MP2, MP3 methods with aug-cc-pVxZ basis set (x = D, T, Q) for groups 1-4;

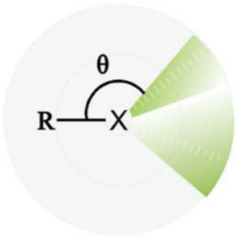

Figure 4.1: Deformation of the halogen bond geometry by variation of bond angle θ.

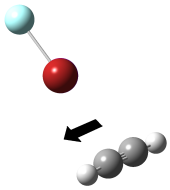

Figure 4.2: Deformation of the halogen bond geometry by translation of FBr along the HCCH molecular axis.

- Coupled Cluster (CCSD(T)) with aug-cc-pVTZ basis set for groups 1-3;
- DFT with aug-cc-pVDZ basis set for groups 1-4.

In particular, the MP2 and MP3 levels were used to obtain the binding energies at MP2.X level [41], set as reference. The CCSD(T) method was used as check of the MP2.X results. The iodine atom was described by means a pseudopotential based aug-cc-pVxZ-PP basis set. To correct the inherent basis set superposition error (BSSE), the BSSE-corrected geometry optimizations and the BSSE-corrected binding energies were obtained with the counterpoise (CP) procedure proposed by Boys and Bernardi [43].

A total of 42 DFT functionals have been chosen, that is: BHandHLYP, HCTH407, B3P86, PBEPBE, BLYP, THCTH, THCTHHYB, BMK, PW91PW91, MPW1PW91, TPSSTPSS, TPSSH, TPSSVWN, M06HF, B97D, B97D3, B3LYP, B3PW91, PBE1PBE, APF, APFD, WB97X, WB97XD, M052X, M05, M06, M062X, M11, B2PLYP, MPW2PLYP, B2PLYPD, B2PLYPD3, MPW2PLYPD, CAM-B3LYP, M06L, LC-BLYP, M08HX, MN15, MN15L, DSDPBEP86, PBE0DH, PBEQIDH.

The binding energy of the investigated systems was obtained by means of

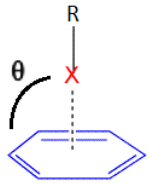

Figure 4.3: Deformation of the halogen bond geometry by variation of bond angle θ.

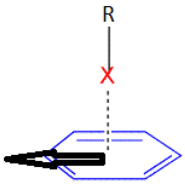

Figure 4.4: Deformation of the halogen bond geometry by RX translation along the axis through the center of benzene and one of the carbon atoms of the ring.

'rigid' energy scans, i.e. performing single point energy calculations, starting from the optimized equilibrium geometry and varying the bonding angle θ from 180° to 90° in step of 3°. For groups 3 and 4, the translation of RX along the molecular axis of acetylene or benzene was also analyzed, starting from the optimized equilibrium geometry, in 10 steps of $\delta = 0.060$ Å for the complexes with acetylene and $\delta = 0.139$ Å for the complexes with benzene. Using such step sizes, the final geometries correspond to have RX vertically placed on a carbon atom of acetylene or benzene, respectively.

The CCSD(T) values reported in Table 4.2, obtained by using the aug-cc-pVQZ and aug-cc-pV5Z basis sets and in the limit of the complete basis set (CBS), result from the following extrapolation formulas:

$$E_b(\text{CCSD(T)}/hl) = E_b(\text{MP2}/hl) + [E_b(\text{CCSD(T)}/ll) - E_b(\text{MP2}/ll)] \quad (4.2)$$

where hl designates the high level calculation, while ll indicates the low level calculation [41];

$$E_b(\text{CCSD(T)}/CBS) = E_b(\text{CCSD(T)}(med)) + F(med)(\varepsilon_{corr}^{MP2/CBS} - \varepsilon_{corr}^{MP2/med}) \quad (4.3)$$

in which

- $E_b(\text{CCSD(T)}(med))$ denotes the energy difference between the CCSD(T) and MP2 methods evaluated with a *medium* basis set;

- $F(med) = (\varepsilon_{corr}^{CCSD(T)/med} / \varepsilon_{corr}^{MP2/med}) - 1$;
- $F(med) = \frac{\varepsilon_{corr}^{CCSD(T)/med}}{\varepsilon_{corr}^{MP2/med}} - 1$

  where $\varepsilon_{corr}^{CCSD(T)/med}$ and $\varepsilon_{corr}^{MP2/med}$ are the correlation energies evaluated at the CCSD(T) and MP2 level, respectively, with the *medium* basis set, instead $\varepsilon_{corr}^{MP2/CBS}$ is the correlation energy of the MP2 evaluated at CBS limit, obtained as the difference between the MP2/CBS value and the RHF value computed with the aug-cc-pVQZ basis set, assumed as the value of the RHF method at the CBS limit [44].

4.1.2 Results

The molecular complexes investigated

1. $FBr\cdots NH_3$, $FI\cdots NH_3$
2. $HCCBr\cdots NH_3$, $HCCI\cdots NH_3$
3. $FBr\cdots HCCH$, $FI\cdots HCCH$
4. $FBr\cdots C_6H_6$, $FI\cdots C_6H_6$

were fully optimized and CP corrected without constraints. The intramolecular and intermolecular equilibrium distances for each system are presented in Table 4.1.

CP optimized parameters for Br-systems			CP optimized parameters for I-systems		
R−Br···B	d(R−Br)	d(Br···B)	R−I···B	d(R−I)	d(I···B)
$FBr\cdots NH_3$	1.82432	2.32851	$FI\cdots NH_3$	1.98166	2.50124
$HCCBr\cdots NH_3$	1.78972	3.02303	$HCCI\cdots NH_3$	2.00023	3.01437
$FBr\cdots HCCH$	1.78348	2.72279	$FI\cdots HCCH$	1.95225	2.84086
$FBr\cdots C_6H_6$	1.76646	3.26692	$FI\cdots C_6H_6$	1.93724	3.40590

Table 4.1: The intramolecular and intermolecular equilibrium distances (Å) of the systems studied.

In order to prove the goodness of MP2.X as reference, we calculated the binding energy of the $FBr\cdots NH_3$ system at equilibrium geometry at MP2.X and CCSD(T) level. When used with large basis sets, the last one is considered the "golden standard" of quantum chemistry calculations. The

comparison between the results provided by the two methods can be seen in Table 4.2. As shown, the value (in red) of the binding energy provided by the MP2.X method using the aug-cc-pVTZ and aug-cc-pVQZ basis sets reproduces very well the value obtained through CCSD(T) calculations in the limit of CBS (in red too) without involving calculations over the third-order of MP, so at cheaper computational cost.

	aug-cc-pVDZ	aug-cc-pVTZ	aug-cc-pVQZ	aug-cc-pV5Z
RHF (raw)	-10.60	-10.20	-9.98	-9.94
RHF (CP)	-9.86	-9.84	-9.93	-9.925
MP2 (raw)	-16.90	-18.04	-18.47	-18.35
MP2 (CP)	-14.30	-16.11	-16.96	-17.15
MP3 (raw)	-14.05	-15.06	-15.51	-15.44
MP3 (CP)	-11.43	-13.27	-14.22	-14.45
CCSD(T) (raw)	-15.09	-16.06	-16.6	-16.57
CCSD(T) (CP)	-12.25	-14.19	-15.26	-15.52
CCSD(T)/QZ	-14.90			
CCSD(T)/5Z	-15.23			
CCSD(T)/CBS	-15.57			
MP2.X (QZ and DZ)	-15.46			
MP2.X (QZ and TZ)	-15.54			

Table 4.2: Binding energy (kcal/mol) for $FBr\cdots NH_3$ at the equilibrium geometry calculated at different levels of theory.

An example of the angular dependence of the binding energy as described by the reference MP2.X method and some DFT functionals is presented in Figure 4.5 for the $FBr\cdots NH_3$ system.

Due to the large number of functionals investigated and, for each functional, the large number of binding energy evaluations, a statistical approach is required to quantify the performance of the different functionals in describing the anisotropy of halogen bonding. To this aim, for each of the complexes studied and for each of the 42 functionals examined, we calculated the global relative mean error (GRME), defined as:

$$\overline{\Delta} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i - y_i}{\overline{y}}\right| \cdot 100 \tag{4.4}$$

where x_i denotes the MP2.X binding energies, y_i indicates the values of

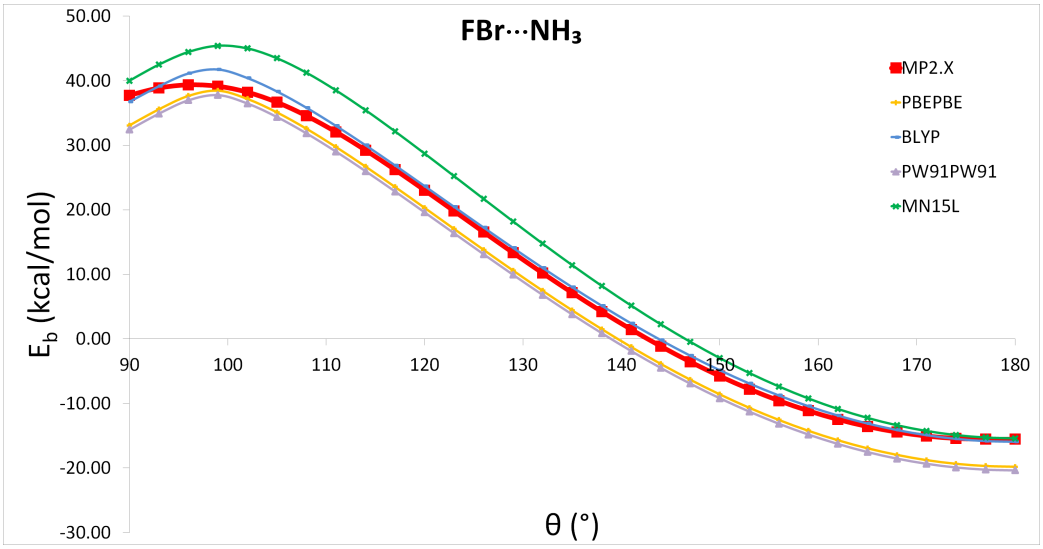

Figure 4.5: Binding energy as function of the bond angle θ for the FBr$\cdots$NH$_3$ system as determined at MP2.X and DFT levels, using the PBEPBE, BLYP, PW91PW91 and MN15L functionals.

binding energy provided by each DFT functional and $\overline{y}$ is the arithmetic average value of DFT binding energies: $\overline{y} = \sum_{i=1}^{N} y_i$. A threshold equal to 10% was chosen to establish the goodness of a functional. Namely, if a functional shows a discrepancy with respect to the reference greater than 10%, it is to be considered not good enough to describe the correct behavior of the binding energy for that complex. In Table 4.3 we have reported, for each of the systems studied, the functionals with a GRME $< 10\%$ and the functionals that provide the worst description of the binding energy for that complex (thus, with the largest GRME): among the first ones, those red colored are the functionals that show the smallest GRME.

As it can be seen from the table, for HCCBr$\cdots$NH$_3$ and HCCI$\cdots$NH$_3$ systems there is not any functional able to reproduce the reference MP2.X within a GRME equal or smaller than 10%, while for FI$\cdots$NH$_3$ and FBr$\cdots$HCCH systems, considering the angular dependence, there is only a functional with a GRME under the threshold. Instead, by analyzing the dependence from the lateral translation in the FBr$\cdots$HCCH complex, a higher number of functionals (19) are considered able to reproduce the reference. The reason of this result could be found in the cylindrical symmetry of the HCCH elec-

	The functionals with GRME $< 10\%$	The worst functional
$FBr\cdots NH_3$	HCTH407, BLYP, MPW1PW91, B3PW91, APF, M06	PW91PW91
$FI\cdots NH_3$	B3PW91	PW91PW91
$HCCBr\cdots NH_3$	none	B97D3
$HCCI\cdots NH_3$	none	PW91PW91
$FBr\cdots HCCH$: θ	M08HX	TPSSVWN
$FI\cdots HCCH$: θ	B3P86, THCTH, B97D, PBE1PBE, APF, APFD, WB97X, WB97XD, M052X, M062X, B2PLYPD3, DSDPBEP86, PBE0DH, PBEQIDH	B97D3
$FBr\cdots HCCH$: *translation*	MPW1PW91, M06HF, B97D, B3PW91, APF, WB97X, WB97XD, M052X, M05, M062X, B2PLYPD3, MPW2PLYPD, LC-BLYP, M08HX, MN15, MN15L, DSDPBEP86, PBE0DH, PBEQIDH	TPSSVWN
$FI\cdots HCCH$: *translation*	BMK, MPW1PW91, B97D, B3PW91, APF, WB97XD, M05, M062X, M11, B2PLYPD, B2PLYPD3, MPW2PLYPD, LC-BLYP, DSDPBEP86, PBEQIDH	TPSSVWN
$FBr\cdots C_6H_6$: *translation*	M06HF, WB97X, WB97XD, M052X, M06, M11, B2PLYPD, B2PLYPD3, MPW2PLYPD, M06L, DSDPBEP86	TPSSVWN
$FI\cdots C_6H_6$: *translation*	WB97X, WB97XD, M052X, M11, B2PLYPD, B2PLYPD3, MPW2PLYPD, M06L, DSDPBEP86, PBEQIDH	TPSSVWN
$FBr\cdots C_6H_6$: θ	B97D, WB97X, M06, M11, M06L, DSDPBEP86	BHandHLYP
$FI\cdots C_6H_6$: θ	WB97X, WB97XD, M052X, M06, M11, B2PLYPD, B2PLYPD3, MPW2PYLPD, M06L, DSDPBEP86	BHandHLYP

Table 4.3: The functionals with a GRME $< 10\%$ (on the left) and the functionals with the largest GRME (on the right) for each of the complexes investigated; the functionals that exhibit the smallest GRME are red colored.

tronic cloud and in the strength shown by the bromine halogen bonds, that is lower than that of iodine halogen bonds. Looking at the plot of the binding energy described by MP2.X and the DFT functionals (Fig. 4.6), it can be seen a smooth decrease of the energy from the configuration in which FBr lies in the middle of the C≡C bond towards the configuration in which the FBr is aligned with a C atom, due to the reduction of the overlap between the electronic clouds of the monomers. Thanks to this feature, the interaction can be easily reproduced by a large number of functionals. A similar result does not occur for the same system when changing the θ bond angle. This can be understood remembering that the halogen bond is a strongly directional interaction. To find a DFT functional able to correctly reproduce the trend of the binding energy along the whole rotation (Fig. 4.7), from the equilibrium configuration ($\theta = 180°$) to the repulsive one ($\theta = 90°$), is more difficult and not all the DFT functionals, among those selected, are able to correctly describe the energy trend, as evidenced for the $HCCBr\cdots NH_3$, $HCCI\cdots NH_3$ and $FBr\cdots HCCH$ systems. Another behavior is exhibited by the π-systems $FBr\cdots C_6H_6$ and $FI\cdots C_6H_6$, for which there are about a dozen functionals under the threshold value and, in particular, most of them are able to describe the binding energy behavior both under translation and under variation of the bond angle θ. The reason could be the symmetry of the benzene electronic distribution: the translation of RX (FBr or FI) from the center of the ring towards a carbon atom leads to an increase of the overlap between the electronic clouds of the monomers and, consequently, to a smooth increase of the binding energy while the bond angle variation involves a situation in which the fluorine atom approach to benzene and this leads to a repulsive interaction (positive value of the binding energy).

The GRME is, as the name says, a *global* quantity: it allows us to discriminate among the functionals for a fixed threshold and an *average* behavior, but it does not take into account the *local* behavior of a DFT functional. For example, a given functional could presents a large GRME but, in some crucial points of the binding energy curve, could reproduce very well the behavior of the reference. It follows that a description based only on the GRME is not exhaustive: the choice of which functional to use to have the best reproduction of the halogen bond in a molecular system must be made taking carefully into account not only the minimum energy geometry but also the regions somewhat outside the minimum, which could be more likely observed in the real systems.

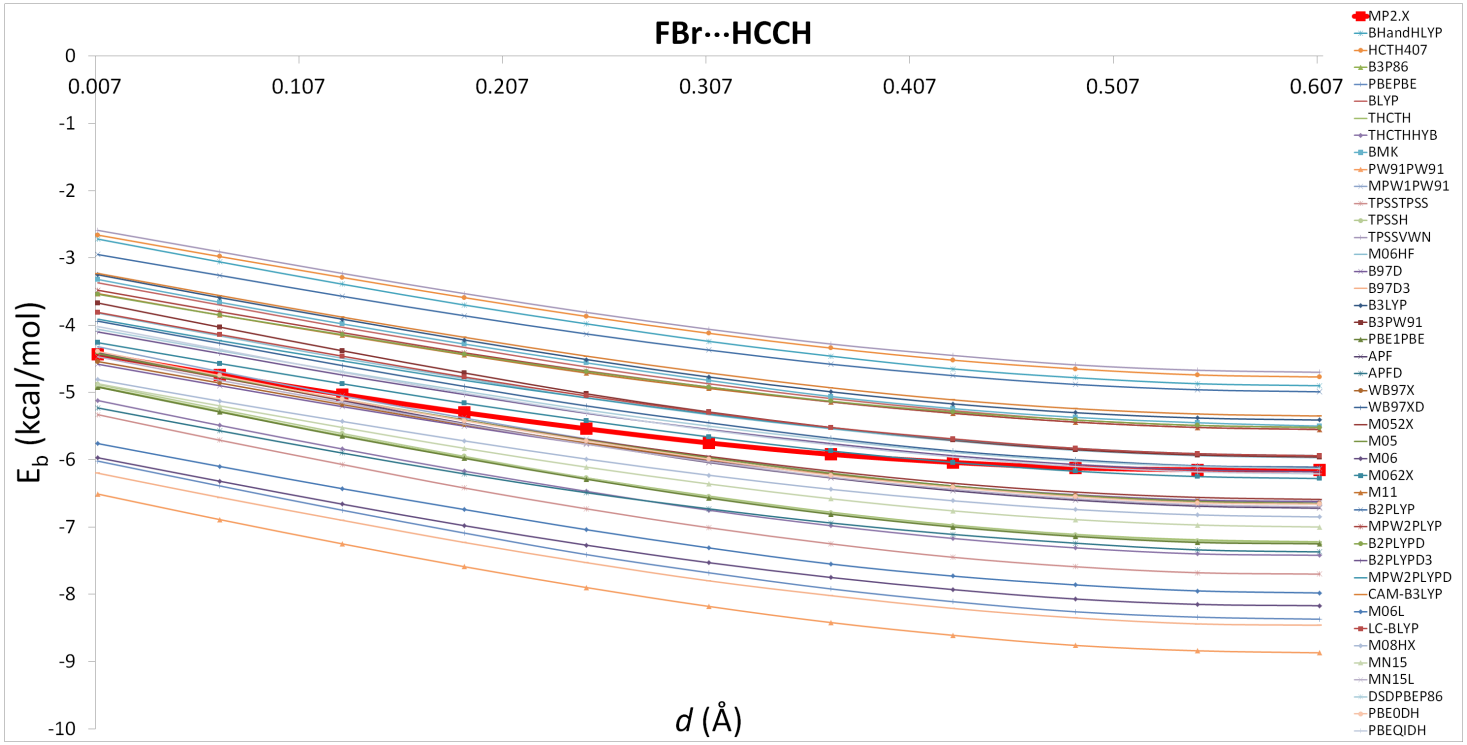

Figure 4.6: The binding energy of the FBr···HCCH system under translation obtained at MP2.X (the thick red line) and DFT levels. All the DFT functionals are reported.

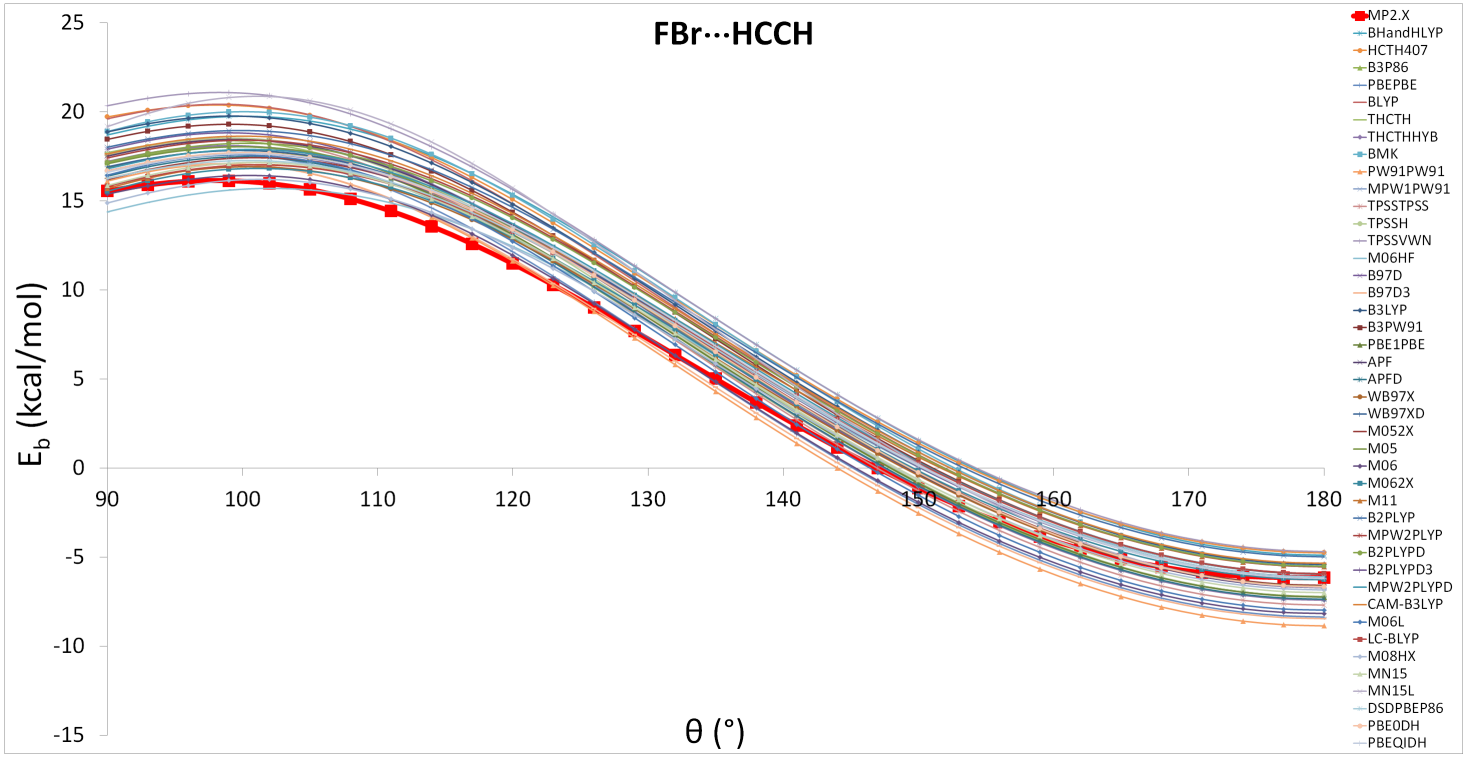

Figure 4.7: The binding energy of the FBr···HCCH system when changing the θ bond angle; the MP2.X level (the thick red line) and all the DFT functionals are reported.

4.2 Interdependence of halogen and hydrogen bonds

According to Voth and coworkers [45], halogen and hydrogen bonds in protein-ligand complexes can be considered as geometrically perpendicular and energetically independent on each other: they proposed that orthogonal halogen and hydrogen bonds are analogous to the orthogonal protecting groups in chemical synthesis, which can be added or removed 'independently' of each other. We wanted to check if this result could be observed also for halogen/hydrogen bonding acceptors when the electronic charge of the halogen/hydrogen bonding donor comes from the π-electron system of benzene. To this aim, we have first considered the following dimers (Fig. 4.8):

1. $NCBr{\cdot}C_6H_6$;
2. $H_2O{\cdot}C_6H_6$.

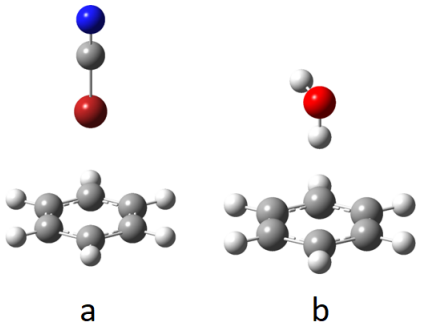

Figure 4.8: The system (a) (1), (b) (2).

Both dimers are at T-shape geometry: NCBr and (H)OH are on an axis perpendicular to the plane of benzene, pointing towards the center of the ring with Br or H at distance r from it (Fig. 4.9).

We determined the variation of the binding energy with r for both systems using some appropriated DFT functionals. Then, in order to investigate the interdependence between halogen and hydrogen bonding, we built-up systems (1') and (2') by adding a water molecule to (1) and a NCBr molecule to (2) on the opposite site of the benzene ring with respect to the NCBr or water molecule, respectively, keeping the T-shape geometry (see Figure 4.10). We

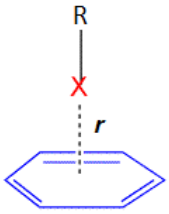

Figure 4.9: The T-shape geometry adopted.

also included the systems (1") and (2"), obtained respectively by adding a NCBr molecule to (1) and a water molecule to (2) (Fig. 4.11). The following trimers have been therefore investigated:

1'. $(NCBr{\cdot}C_6H_6){\cdot}H_2O$;

2'. $(H_2O{\cdot}C_6H_6){\cdot}NCBr$;

1". $(NCBr{\cdot}C_6H_6){\cdot}NCBr$;

2". $(H_2O{\cdot}C_6H_6){\cdot}H_2O$.

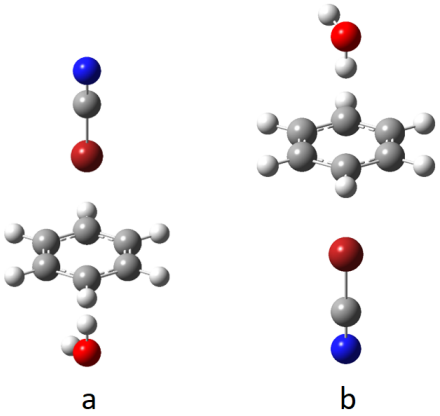

Figure 4.10: The system (a) (1'), (b) (2').

4.2.1 Computational details

The Gaussian16 software [42] was used to perform all the QM calculations *in vacuo*. We chose the M06-2X, wB97X and M11 functionals for DFT

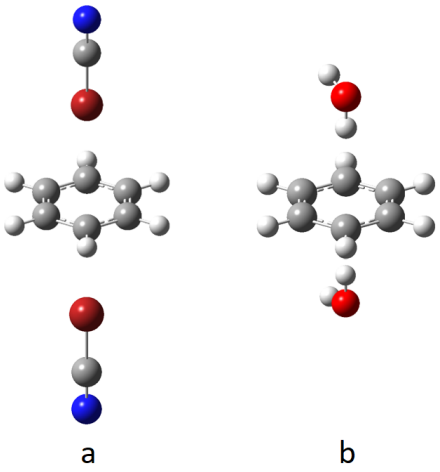

Figure 4.11: The system (a) (1"), (b) (2").

calculations, owing to their good performance in describing halogen bonding with benzene, using the aug-cc-pVTZ basis set.

The geometries of the monomers (NCBr, H_2O, C_6H_6) were optimized at each level of theory mentioned. The binding energies (of dimers and trimers) were CP corrected using the procedure of Boys and Bernardi [43]. The dimers binding energy curves were obtained by varying r over the range from 1.8 Å to 6.0 Å with a step size $\delta = 0.1$ Å. For each step we performed a single point energy calculation (i.e., a rigid scan of energy). Only for the dimers, the step size was reduced to $\delta = 0.02$ Å around the minimum energy, in order to better localize the equilibrium value r_{eq} (Table 4.5). The same procedure has been followed for the binding energies of the trimers: the range and the step size are the same as the dimers, but the evaluations were done adding a H_2O molecule to (1) and a NCBr molecule to (2) on the opposite side of the benzene with respect to the NCBr or H_2O molecule, respectively; we fixed both molecules added at the equilibrium distance as determined from the previous calculations on the dimers. The same was done for the systems (1") and (2") (Fig. 4.12).

The approximation to use a rigid PES scan instead of a relaxed one, where monomers are optimized for each intermolecular distance, was checked by computing a curve of binding energy with both methods, using the wB97X functional. At the equilibrium distance both methods provide the same result, as shown in Table 4.4 and Figure 4.13. The differences are, indeed,

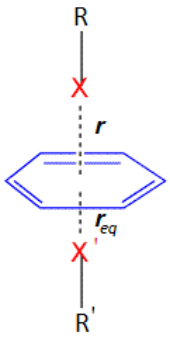

Figure 4.12: The configuration adopted for the trimers.

well below 1 kcal/mol, therefore negligible. The same is true for the region around the minimum and on the tail but, of course, not in the repulsive region where monomers are expected to undergo a substantial rearrangement in their geometry (Fig. 4.13).

wB97X	$E_b(r_{eq})$	$E_b(r_{sr})$	$E_b(r_{lr})$
Rigid PES scan	-3.670	-1.811	-1.436
Relaxed PES scan	-3.678	-1.830	-1.439

Table 4.4: Comparison between the binding energy (kcal/mol) provided at wB97X/aug-cc-pVTZ level by the rigid PES scan and the relaxed one at the equilibrium distance, $r_{eq} = 3.40$ Å, in the short-range region ($r_{sr} = 3.0$ Å) and in the long-range region ($r_{lr} = 4.5$ Å) of (1).

In the pictures reported in the section 4.2.2, for (1), (1') and (1") the range shown is reduced to $[2.7, 6.0]$ Å, due to the huge repulsive behavior of these systems in the range $[1.8, 2.7[$ Å that masks the attractive region around the equilibrium distance.

The ESP maps were plotted on a 0.001 a.u. isosurface of electron density and in the $[+6, -6]$ a.u. range, in order to emphasize the differences among them.

4.2.2 Results

First, we calculated the binding energies E_b of the dimers (1) and (2) with all the functionals chosen varying the distance r between NCBr or H_2O and

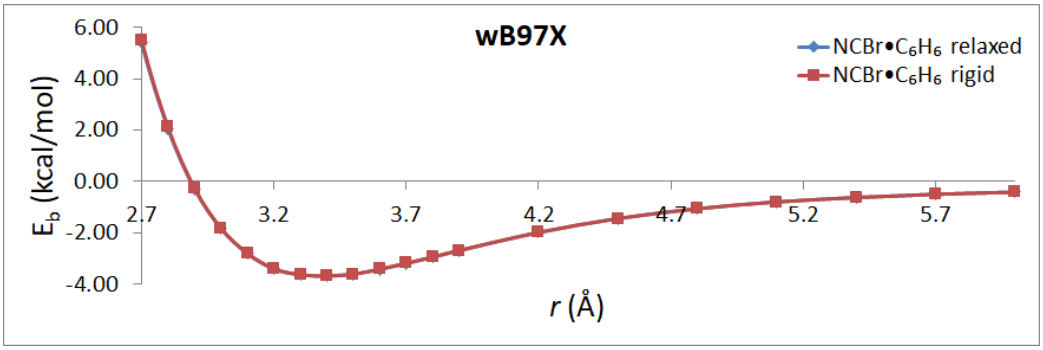

Figure 4.13: Comparison between the binding energy curves provided at wB97X/aug-cc-pVTZ level by the rigid PES scan and the relaxed one for (1). The curves are almost overlapped.

the center of benzene ring in the range mentioned (see 4.2.1 section) and determining the equilibrium distance r_{eq} and the corresponding energy for both the systems (see Table 4.5).

	$NCBr{\cdot}C_6H_6$		$H_2O{\cdot}C_6H_6$	
	r_{eq}	$E_b(r_{eq})$	r_{eq}	$E_b(r_{eq})$
wB97X	3.40	-3.67	2.42	-3.14
M06-2X	3.24	-3.83	2.32	-3.44
M11	3.32	-3.48	2.34	-3.01

Table 4.5: The equilibrium distances (Å) and the corresponding energies (kcal/mol) for the dimers (1) and (2) for each examined functional, using the aug-cc-pVTZ basis set.

Once the minima were known, adding a H_2O molecule to (1) or a NCBr one to (2) at their r_{eq}, we determined the binding energies of the trimers (1') and (2'). By comparing the energy behavior of systems (1) and (2) with that of (1') and (2'), respectively, we can get insight on the relative interdependence between halogen and hydrogen bonding. A perfect orthogonality of the two interactions, in fact, should lead to perfectly overlapped E_b curves for (1) and (1') and, similarly, for (2) and (2').

First we will focus our attention on the results obtained with the wB97X functional. The other two functionals provide similar results and they will presented further on.

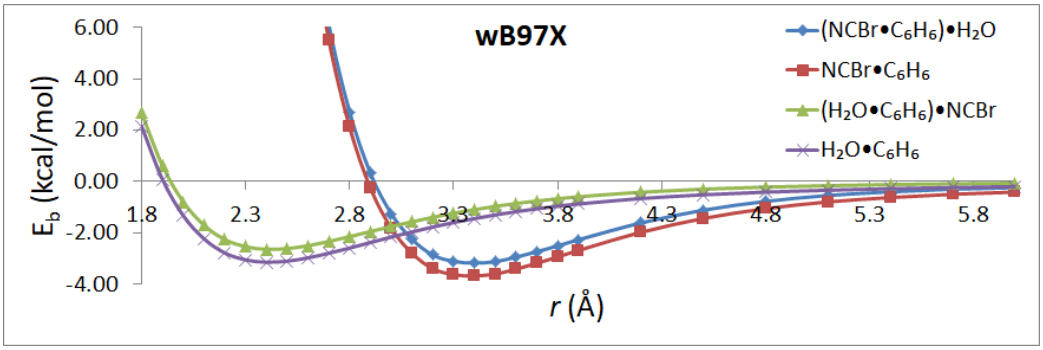

Figure 4.14: Comparison between the binding energy curves obtained at wB97X/aug-cc-pVTZ level for (1), (1'), (2) and (2').

As shown in Figure 4.14, the presence of the additional molecule increases E_b in the full range of r. Evaluating the difference in the minimum r_{eq}, the increase is 0.5 kcal/mol for (1'); for (2') the evaluation was done at $r_c = 2.40$ Å, that is the closest point to the minimum, and it leads to the same results obtained for (1'), i.e. 0.5 kcal/mol. Such values represent 16% and 19%, respectively, of the binding energy for the two systems, indicating that the interactions cannot be considered orthogonal. On the other side, it should be noted that the value of r_{eq} is not affected by the presence of the other molecule. In other words, from a geometric point of view, halogen and hydrogen bonds can be considered orthogonal.

In order to complete the analysis, we determined the binding energies of the trimers (1") and (2") obtained adding a NCBr molecule to (1) and a H_2O molecule to (2) at their equilibrium distances (Fig. 4.11). We found an analogous behavior to that of the trimers (1') and (2'), that is, the presence of the added molecule increases the binding energy, as shown in Figure 4.15. At the energy minimum the increase amounts to 0.75 kcal/mol for (1") and 0.39 kcal/mol for (2") compared to (1) and (2), respectively.

By comparing the behavior of the systems (1') and (1") we observe that (1") is more destabilized compared to (1') (Fig. 4.16). Comparison between (2') and (2"), instead, reveals that (2') exhibits a larger energy increase with respect to (2") (Fig. 4.17).

By considering the results provided by the other two functionals, that is M06-2X and M11, the closest points r_c to the minima in which to evaluate the energy differences between dimers and trimers are

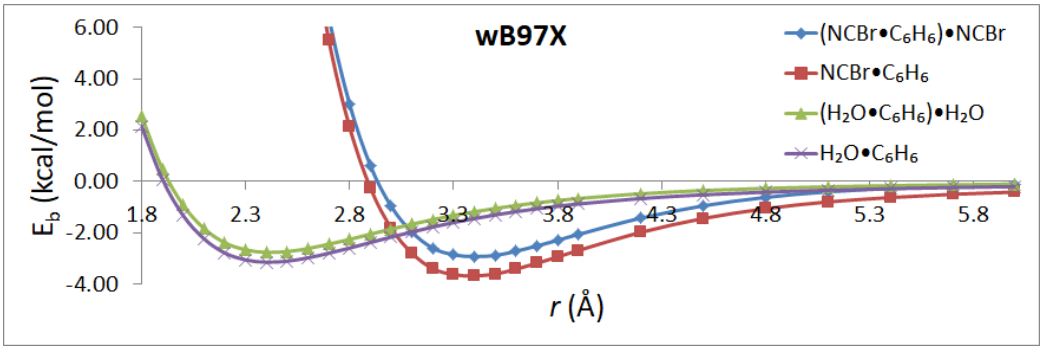

Figure 4.15: Comparison between the binding energy curves obtained at wB97X/aug-cc-pVTZ level for (1), (1"), (2) and (2").

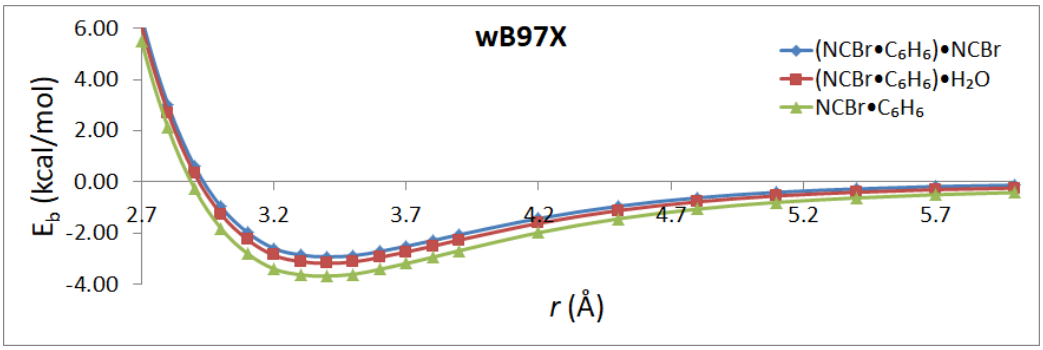

Figure 4.16: Comparison between the binding energy curves obtained at wB97X/aug-cc-pVTZ level for (1), (1') and (1").

- M06-2X, M11: 3.30 Å for (1);
- M06-2X, M11: 2.30 Å for (2).

The corresponding energies are shown in Table 4.6.

The comparison among the binding energy trends of the dimers and all the trimers analyzed reveals that (1") and (2') are more destabilized compared to (1') and (2"), respectively (Fig. 4.18, 4.20, 4.19 and 4.21).

The adoption of M06-2X and M11 functionals then confirms the results determined by using the wB97X functional.

Further details can be discussed analyzing the behavior of the three DFT functionals for the same dimer. Let us consider (1) first. As shown in Figure 4.22, the wB97X curve of binding energy is less attractive than the other two

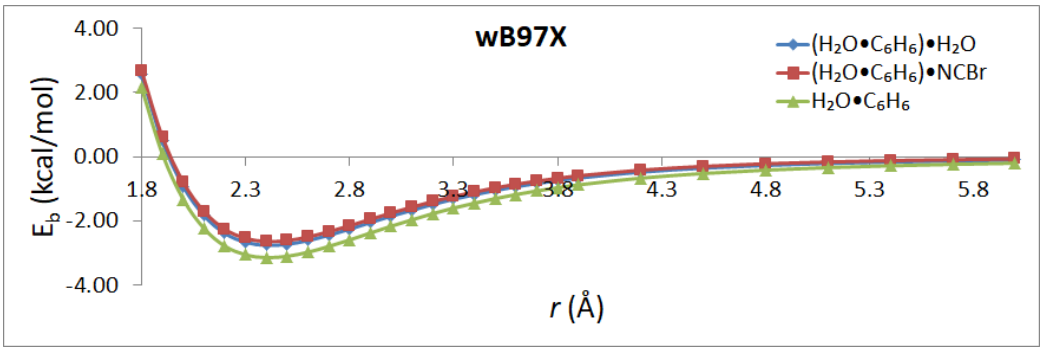

Figure 4.17: Comparison between the binding energy curves obtained at wB97X/aug-cc-pVTZ level for (2), (2') and (2").

	$NCBr{\cdot}C_6H_6$		$H_2O{\cdot}C_6H_6$	
	r_c	$E_b(r_c)$	r_c	$E_b(r_c)$
M06-2X	3.30	-3.80	2.30	-3.43
M11	3.30	-3.48	2.30	-3.01

Table 4.6: The distances r_c (Å) at which the evaluations of the energy differences between dimers and trimers were done for the functionals M06-2X and M11 and the corresponding energies (kcal/mol).

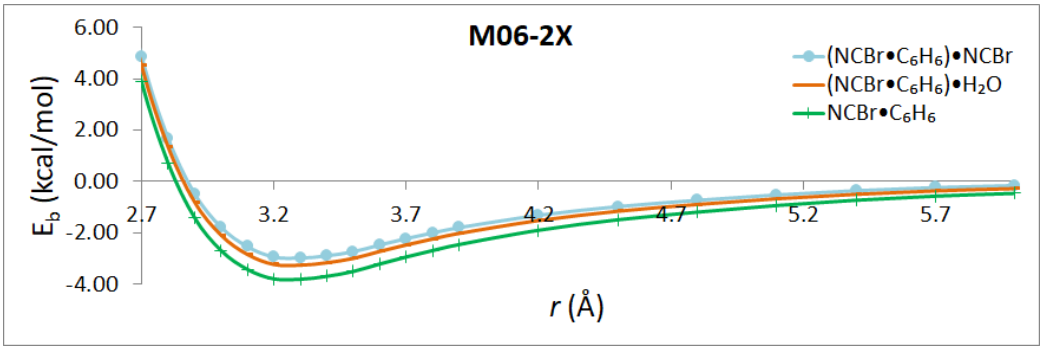

Figure 4.18: Comparison between the binding energy curves obtained at M06-2X/aug-cc-pVTZ level for (1), (1') and (1").

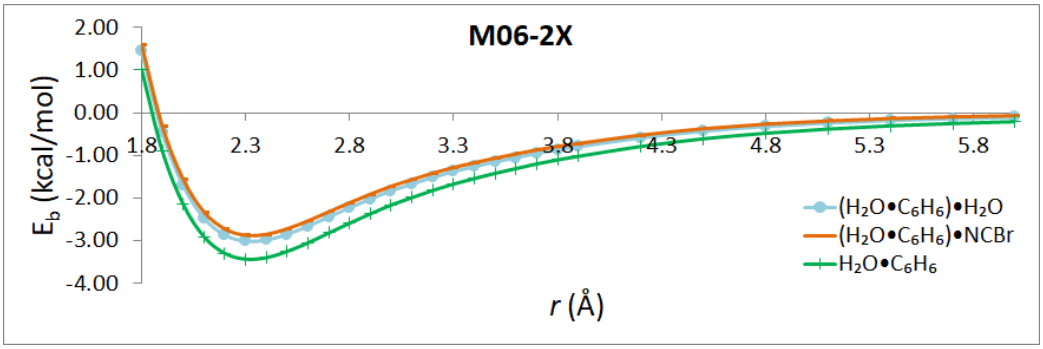

Figure 4.19: Comparison between the binding energy curves obtained at M06-2X/aug-cc-pVTZ level for (2), (2') and (2").

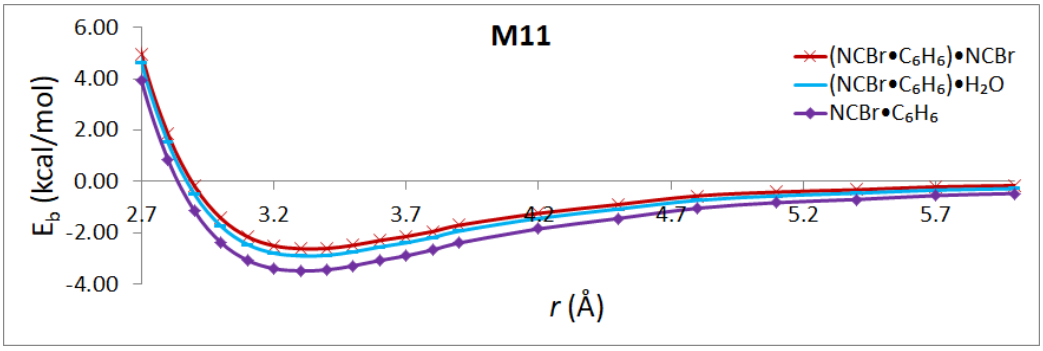

Figure 4.20: Comparison between the binding energy curves obtained at M11/aug-cc-pVTZ level for (1), (1') and (1").

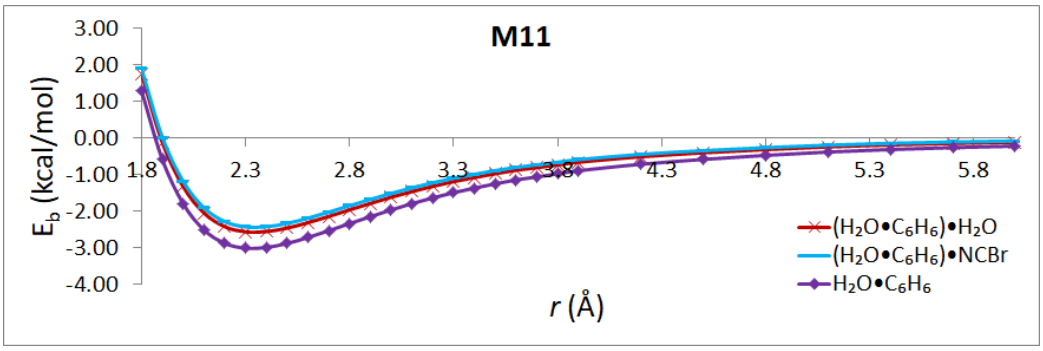

Figure 4.21: Comparison between the binding energy curves obtained at M11/aug-cc-pVTZ level for (2), (2') and (2").

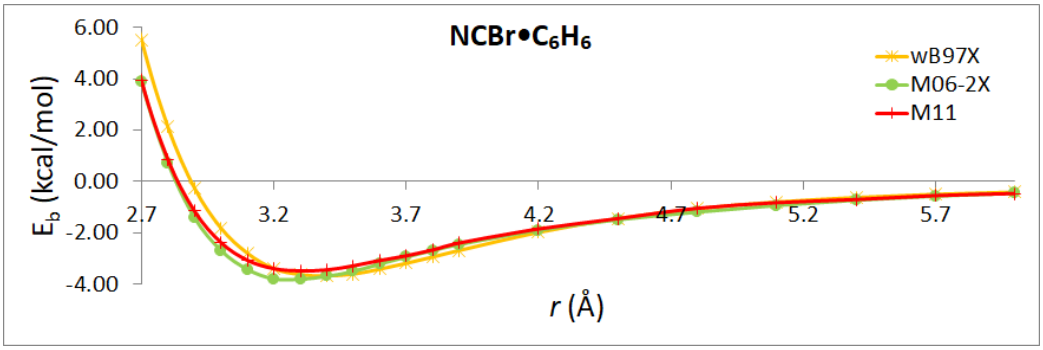

Figure 4.22: Comparison between the binding energy curves obtained for (1) by the wB97X, M06-2X and M11 DFT functionals with the aug-cc-pVTZ basis set.

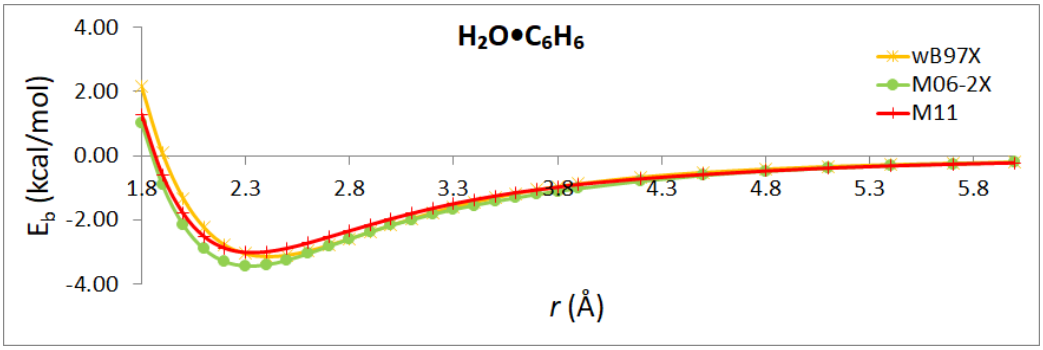

Figure 4.23: Comparison between the binding energy curves obtained for (2) by the wB97X, M06-2X and M11 DFT functionals with the aug-cc-pVTZ basis set.

in the repulsive region and on the tail, while in the minimum the wB97X binding energy is intermediate between the M06-2X and the M11 values. Moreover, there is a small shift among the equilibrium distances (Table 4.5). The same description can be found comparing the behavior of the three functionals for (2) (Fig. 4.23).

Finally we determined the ESP maps of benzene and dimers on the 0.001 a.u. isosurfaces of electron density. The pictures obtained (Fig. 4.24) are indistinguishable for the three functionals. In particular, comparing the values of ESP at the center of the ring, the differences between the functionals are of the order of 10^{-3} a.u. (Table 4.7).

The results provided by the DFT functionals can be summarized as follow:

- the interaction between the monomers of the systems (1) and (2) be-

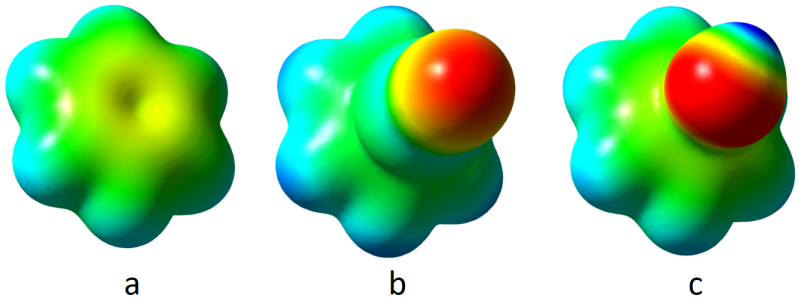

Figure 4.24: The ESP maps of (a) benzene, (b) NCBr·C_6H_6, (c) H_2O·C_6H_6.

	NCBr	NCBr·C_6H_6	H_2O·C_6H_6
wB97X	-0.023	-0.008	-0.014
M06-2X	-0.021	-0.007	-0.012
M11	-0.025	-0.011	-0.016

Table 4.7: The ESP values (a.u.) at the center of the ring for benzene, NCBr·C_6H_6 and H_2O·C_6H_6.

comes less favored when another water or NCBr molecule is present on the other side of the benzene ring (see Tables 4.8 and 4.9);

- the energy destabilization is greater when a NCBr molecule, rather than a water molecule, is added to the C_6H_6·BrCN dimer;
- in the same way, the energy destabilization is greater when a NCBr molecule, rather than a water molecule, is added to the C_6H_6·H_2O dimer.

	r_c	$E_b(1')-E_b(1)$	$E_b(1'')-E_b(1)$
wB97X	3.40	0.50	0.75
M06-2X	3.30	0.55	0.82
M11	3.30	0.59	0.87

Table 4.8: The distances r_c (Å) and the corresponding energy differences (kcal/mol) for the trimers (1') and (1") with respect to (1); for wB97X it is $r_c = r_{eq}$.

We have therefore demonstrated that, for the investigated systems, halogen and hydrogen bonding can not be considered independent interactions.

	r_c	$E_b(2')-E_b(2)$	$E_b(2'')-E_b(2)$
wB97X	2.40	0.50	0.39
M06-2X	2.30	0.56	0.42
M11	2.30	0.59	0.45

Table 4.9: The distances r_c (Å) and the corresponding energy differences (kcal/mol) for the trimers (2') and (2") with respect to (2).

4.3 Use of extremely localized molecular orbitals in the Spin-Coupled method

The ELMOs have been here used for the description of some simple molecules at Spin-Coupled (SC) level. This study is aimed at evaluating the performance of this approach in view of its application for some systems for which the convergence at SC level, using the standard MO approach, cannot be reached. The molecules studied at Spin-Coupled Restricted Hartree-Fock (SC RHF) and SC ELMO levels are (Fig. 4.25):

1. Salicylic acid
2. Pyrazinamide
3. Naphthalene

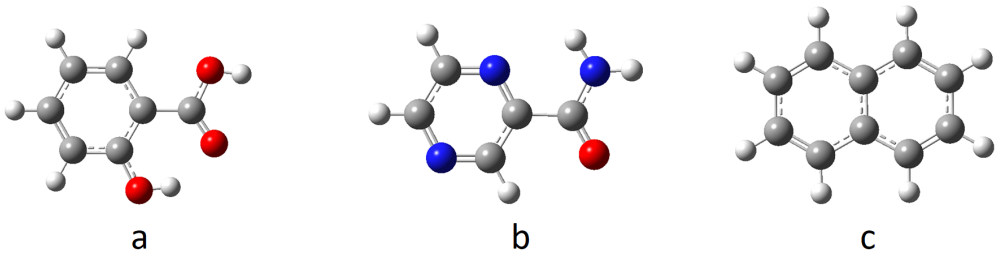

Figure 4.25: The molecules analyzed (a) salicylic acid; (b) pyrazinamide; (c) naphthalene.

	Number of atoms	Number of basis functions				
		6-31G	6-311G	6-31G*	6-31G(d,p)	6-311G*
Salicylic acid	16	102	148	162	180	208
Pyrazinamide	14	91	132	145	160	186
Naphthalene	18	106	154	166	190	214

Table 4.10: The basis sets used and the corresponding number of basis functions for each molecule studied.

4.3.1 Computational details

The geometry optimizations and the SCF calculations were performed using Gaussian09 software [46] *in vacuo* with the use of explicit basis sets [47]. The SC calculations were computed with the Spin-Coupling software employing both the orbitals provided by the RHF calculation and those obtained with the ELMO method. The latter orbitals have been obtained by a code developed by Dr. A. Genoni [39, 48]. MP2 calculations have been also performed as reference. The 6-31G, 6-311G, 6-31G*, 6-31G(d,p) and 6-311G* basis sets have been used. The electronic energy has been evaluated at each level of theory mentioned for every molecule included in the investigation. The number of the basis functions is shown in Table 4.10.

The electrons chosen to represent the valence electrons were selected from the rings of the molecular systems analyzed: we selected 6 valence electrons for salicylic acid and pyrazinamide and 10 valance electrons for naphthalene.

4.3.2 Results

The SC RHF calculations failed to reach convergence only for the pyrazinamide. The SC ELMO approach for the same molecule, instead, converged correctly. Plotting the electronic energies E as a function of the basis set, we obtain the energy trends shown in Figures 4.26, 4.27 and 4.28 (data are in hartrees).

By increasing the number of basis functions, the energy values are obviously gradually lower. Moroever, for a given basis set the trend in energy is the following:

$$E(\mathrm{MP2}) < E(\mathrm{SC\,RHF}) < E(\mathrm{RHF}) < E(\mathrm{SC\,ELMO}) < E(\mathrm{ELMO}). \quad (4.5)$$

Due to the way the ELMOs are defined, the corresponding energy is higher than that obtained by using the MOs. The MP2 provides the lowest energy

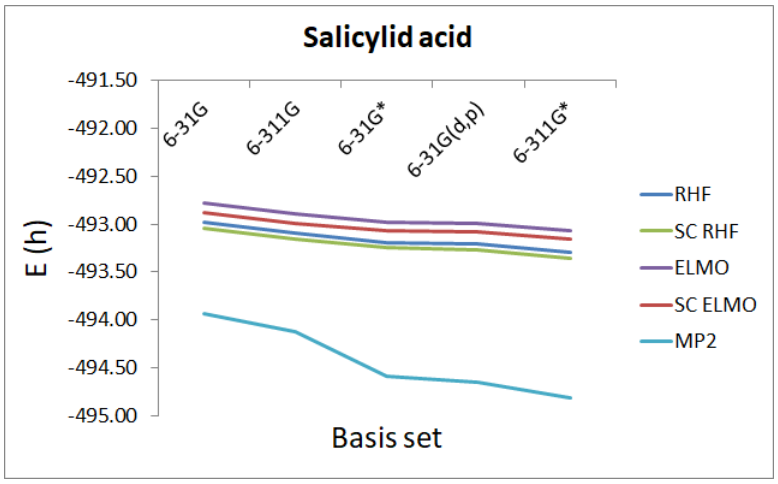

Figure 4.26: The energy trends for salicylic acid.

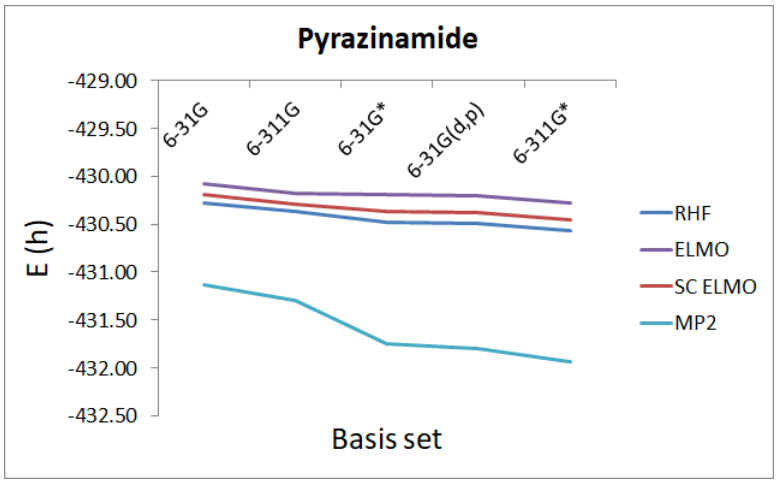

Figure 4.27: The energy trends for pyrazinamide. There is no curve for SC RHF because the convergence failed at that level.

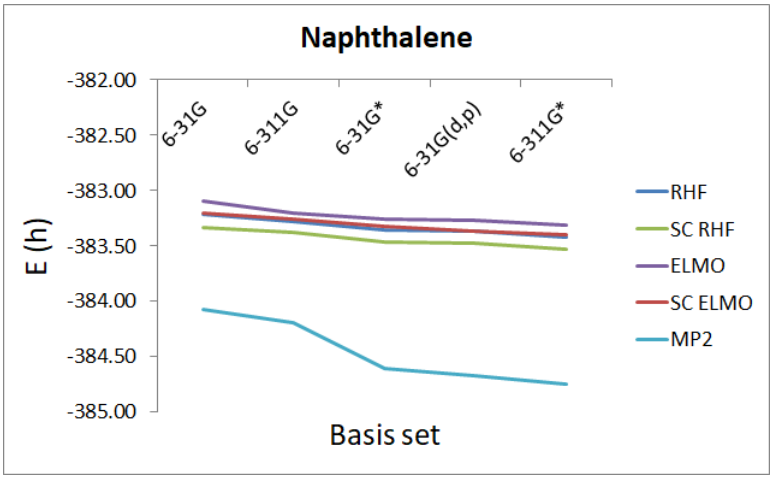

Figure 4.28: The energy trends for naphthalene.

because it takes into account the correlation energy through the perturbative theory. The SC method includes the static correlation, that is not computed at Hartree-Fock level, so that $E(\text{SC RHF}) < E(\text{RHF})$ and $E(\text{SC ELMO}) < E(\text{ELMO})$.

Thus, calculating the differences $E(\text{RHF}) - E(\text{SC RHF})$ and $E(\text{ELMO}) - E(\text{SC ELMO})$, we can determine the correlation contribution introduced at SC level with respect to the Hartree-Fock method. As done with the electronic energies, we can evaluate the differences ΔE mentioned as a function of the basis set (Fig. 4.29).

The two kind of ΔE are almost constant for the salicylic acid, but not for the other two molecules. There is no evaluation of $E(\text{RHF}) - E(\text{SC RHF})$ for the pyrazinamide because the convergence failed at SC RHF level.

As described in the chapter 3, ELMOs are obtained by applying the variational principle limiting the combinations to subsets of the total basis set corresponding to given fragments of the analyzed molecule. Through this procedure, the orthogonalization 'tails' are explicitly excluded. Knowing therefore that the use of ELMOs has a not negligible impact on energy, we define the quantity $\frac{E(\text{SC ELMO})-E(\text{ELMO})}{E(\text{RHF})-E(\text{ELMO})}$ in order to determine the 'correction' to the ELMO energy made at SC ELMO level with respect to the error implied by using ELMOs instead of MOs in the RHF calculation. By expressing the fraction mentioned as a percentage value, we define the recovery rate as

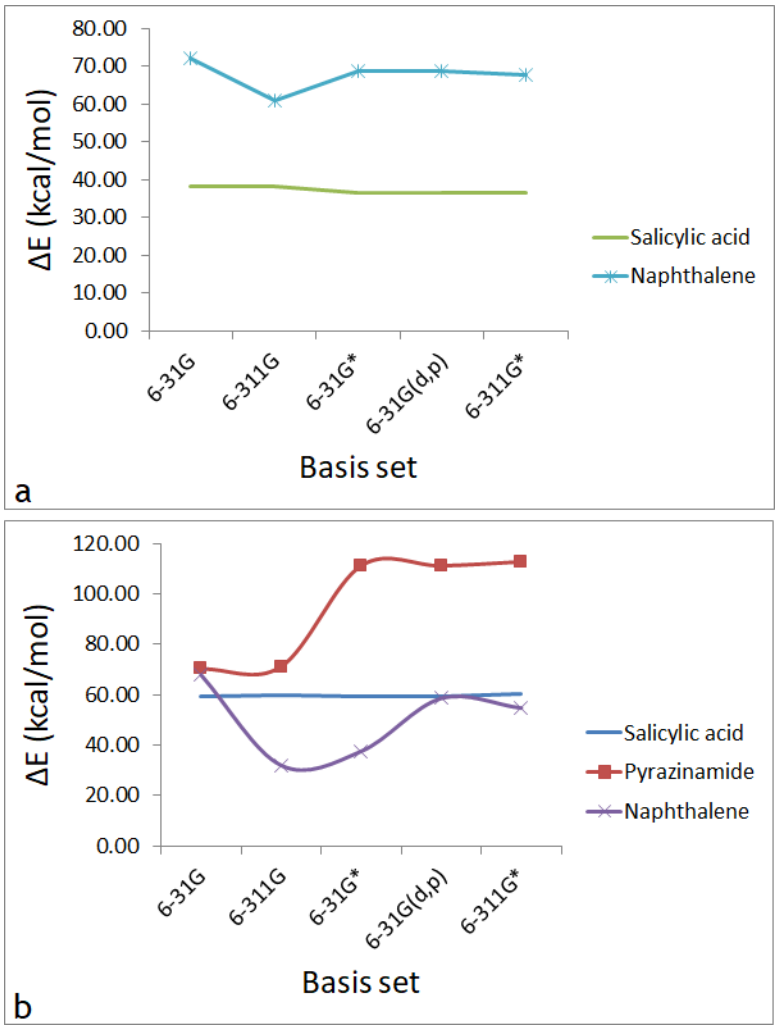

Figure 4.29: The behavior of the correlation energy at level (a) SC RHF, (b) SC ELMO.

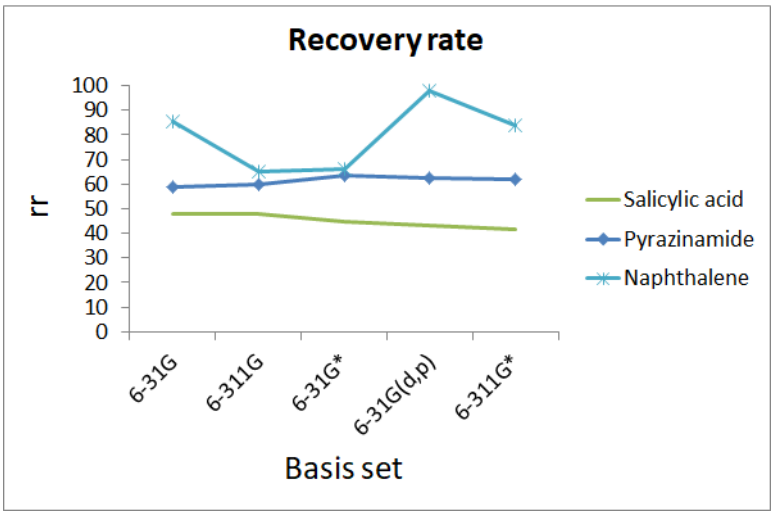

Figure 4.30: The behavior of the recovery rate as a function of the basis set for the molecules investigated.

$$rr = \frac{E(\text{SC ELMO}) - E(\text{ELMO})}{E(\text{RHF}) - E(\text{ELMO})} \cdot 100. \tag{4.6}$$

By plotting the behavior of the recovery rate as a function of the basis set (see Figure 4.30) it results that, while rr is about constant for salicylic acid and pyrazinamide, it changes a lot for the naphthalene and, for the 6-31G(d,p) basis set, it is $rr \approx 100\%$, that is, $E(\text{SC ELMO}) \approx E(\text{RHF})$. In other words, the SC ELMO method allows to fully recover the RHF energy while using ELMOs instead of MOs.

In conclusion, we showed that the adoption of the ELMOs can lead to useful improvements helping to reach the convergence when it cannot be reached using the standard spin-coupled orbitals (i.e. developed on the basis of traditional MOs). Moreover, unlike the latter ones, the convergence is much faster using ELMOs. The price to pay is a little overestimation of the energy due to the way these orbitals are obtained. This detail has to be taken into account depending on the system to analyze and the required accuracy.

Chapter 5

Conclusions

The work reported clearly describes how the halogen bond is not a topic that can be studied through the adoption of a single computational protocol. The choice to use one method rather than another must be made taking into account features and properties of the molecular system to investigate. It can be necessary to use different methods in order to understand which one provides the best description of the system. Within each individual method, a computational hierarchy can be created. The aim is to refine the model obtained when the method with which the analysis was started is not able to provide results of the required accuracy. Furthermore, for a description as complete and coherent as possible, it can be successful to combine different methods when necessary. The work presented tried to provide a satisfactory picture of the molecular systems studied with the tools available. But the picture does not claim to be exhaustive and the way is obviously open to new goals.

Bibliography

[1] M. Adler, M. J. Kochanny, B. Ye, G. Rumennik, D. R. Light, S. Biancalana, and M. Whitlow. Crystal structures of two potent nonamidine inhibitors bound to factor Xa. *Biochemistry*, 41:15514–15523, 2002.

[2] P. Auffinger, F. A. Hays, E. Westhof, and P. S. Ho. Halogen bonds in biological molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 101:16789–16794, 2004.

[3] E. Howard, R. Sanishvili, R. E. Cachau, B. Mitschler, P. Bart, V. Lamour, M. V. Zandt, E. Sibley, C. Bon, D. Moras, T. R. Schneider, A. Joachimiak, and A. Podjarny. Ultrahigh resolution drug design I: Details of interactions in human aldose reductase-inhibitor complex at 0.66 Å. *Proteins: Structure, Function, and Bioinformatics*, 55:792–804, 2004.

[4] Y. Jiang, A. A. Alcaraz, J. M. Chen, H. Kobayashi, Y. J. Lu, and J. P. Snyder. Diastereomers of dibromo-7-*epi*-10-deacetylcephalomannine: Crowded and cytotoxic taxanes exhibit halogen bonds. *Journal of Medicinal Chemistry*, 49:1891–1899, 2006.

[5] E. Cariati, A. Forni, S. Biella, P. Metrangolo, F. Meyer, G. Resnati, S. Righetto, E. Tordin, and R. Ugo. Tuning second-order NLO responses through halogen bonding. *Chemical Communications*, pages 2590–2592, 2007.

[6] J. Hulliger and P. J. Langley. On intrinsic and extrinsic defect-forming mechanisms determining the disordered structure of 4-iodo-4'-nitrobiphenyl crystals. *Chemical Communications*, pages 2557–2558, 1998.

[7] J. A. R. P. Sarma, F. H. Allen, V. J. Hoy, J. A. K. Howard, R. Thaimattam, K. Biradha, and G. R. Desiraju. Design of an SHG-active crystal, 4-iodo-4'-nitrobiphenyl: the role of supramolecular synthons. *Chemical Communications*, pages 101–102, 1997.

[8] P. K. Thallapally, G. R. Desiraju, M. Bagieu-Beucher, R. Masse, C. Bourgogne, and J. F. Nicoud. 1,3-Dibromo-2,4,6-trinitrobenzene (DBTNB). crystal engineering and perfect polar alignment of two-dimensional hyperpolarizable chromophores. *Chemical Communications*, pages 1052–1053, 2002.

[9] H. L. Nguyen, P. N. Horton, M. B. Hursthouse, A. C. Legon, and D. W. Bruce. Halogen Bonding: A New Interaction for Liquid Crystal Formation. *Journal of the American Chemical Society*, 126:16–17, 2004.

[10] J. Xu, X. Liu, and C. He. Synthesis and Self-Assembly of Difunctional Halogen-Bonding Molecules: A New Family of Supramolecular Liquid-Crystalline Polymers. *Macromolecules*, 38:3554–3557, 2005.

[11] J. Xu, X. Liu, J. K. P. Ng, T. Lin, and C. He. Trimeric supramolecular liquid crystals induced by halogen bonds. *Journal of Materials Chemistry*, 16:3540–3543, 2006.

[12] D. W. Bruce, P. Metrangolo, F. Meyer, T. Pilati, C. Präang, G. Resnati, G. Terraneo, S. Wainwright, and A. C. Whitwood. Structure-function relationships in liquid-crystalline halogen-bonded complexes. *Chemistry - A European Journal*, 16:9511–9524, 2010.

[13] P. Metrangolo, Y. Carcenac, M. Lahtinen, T. Pilati, K. Rissanen, A. Vij, and G. Resnati. Nonporous organic solids capable of dynamically resolving mixtures of diiodoperfluoroalkanes. *Science*, 323:1461–1464, 2009.

[14] A. Abate, M. Brischetto, G. Cavallo, M. Lahtinen, P. Metrangolo, T. Pilati, S. Radice, G. Resnati, K. Rissanen, and G. Terraneo. Dimensional encapsulation of $I^- \cdots I_2 \cdots I^-$ in an organic salt crystal matrix. *Chemical Communications*, 46:2724–2726, 2010.

[15] M. Fourmigué and P. Batail. Activation of hydrogen- and halogen-bonding interactions in tetrathiafulvalene-based crystalline molecular conductors. *Chemical Reviews*, 104:5379–5418, 2004.

[16] O. Bolton, K. Lee, H. J. Kim, K. Y. Lin, and J. Kim. Activating efficient phosphorescence from purely organic materials by crystal design. *Nature Chemistry*, 3:205–210, 2011.

[17] R. Kato, T. Imakubo, H. Yamamoto, R. Maeda, M. Fujiwara, and H. Sawa. An application of supramolecular chemistry to molecular conductors. *Molecular Crystals and Liquid Crystals*, 380:61–68, 2003.

[18] D. L. Cooper, J. Gerratt, M. Raimondi, M. Sironi, and T. Thorsteinsson. Expansion of the spin-coupled wavefunction in Slater determinants. *Theoretica Chimica Acta*, 85:261–270, 1993.

[19] F. Liu, D. Zhang L. Du, and J. Gao. *Quantum Chemistry*, 116:110–117, 2016.

[20] P. Politzer, P. Lane, M. C. Concha, and S. Murray Y. Ma. An overview of halogen bonding. *Journal of Molecular Modeling*, 13:305–311, 2007.

[21] S. Kozuch and J. M. L. Martin. Halogen bonds: Benchmarks and theoretical analysis. *Journal of Chemical Theory and Computation*, 9:1918–1931, 2013.

[22] M. Yamada, R. Kanazawa, and F. Hamada. Halogen-halogen interactions and halogen bonding in thiacalixarene systems. *CrystEngComm*, 16:2605, 2014.

[23] C. R. Groom, I. J. Bruno, M. P. Lightfoot, and S. C. Ward. The Cambridge Structural Database. *Acta Crystallographica Section B*, 72:171–179, 2016.

[24] P. Metrangolo and G. Resnati. Type II halogen$\cdots$halogen contacts are halogen bonds. *IUCrJ*, 1:5, 2014.

[25] M. S. Pavan, K. Durga Prasad, and T. N. Guru Row. Halogen bonding in fluorine: experimental chargedensity study on intermolecular F$\cdots$F and F$\cdots$S donor-acceptor contacts. *Chemical Communications*, 49:7558–7560, 2013.

[26] M. T. Messina, P. Metrangolo, W. Panzeri, T. Pilati, and G. Resnati. Intermolecular recognition between hydrocarbon oxygen-donors and perfluorocarbon iodine-acceptors: The shortest O$\cdots$I non-covalent bond. *Tetrahedron*, 57:8543–8550, 2001.

[27] P. Metrangolo, G. Resnati, T. Pilati, R. Liantonio, and F. Meyer. Engineering functional materials by halogen bonding. *Journal of Polymer Science Part A: Polymer Chemistry*, 45:1–15, 2007.

[28] P. Metrangolo, G. Resnati, T. Pilati, and S. Biella. Halogen bonding in crystal engineering. *Structure and bonding (Berlin)*, 126:105–136, 2008.

[29] A. V. Vasilyev, S. V. Lindeman, and J. K. Kochi. Noncovalent binding of the halogens to aromatic donors. Discrete structures of labile Br2 complexes with benzene and toluene. *Chemical Communications*, pages 909–910, 2001.

[30] M. Amati, F. Lelj, R. Liantonio, P. Metrangolo, S. Luzzati, T. Pilati, and G. Resnati. Hybrid iodoperfluoroalkane-ferrocene supramolecular arrays: The shortest contacts iodine forms with nitrogen atoms and unsaturated moieties. *Journal of Fluorine Chemistry*, 125:629–640, 2004.

[31] R. D. Bailey, M. L. Buchanan, and W. T. Pennington. Molecular complexes of 1,4-diazines with iodine. *Acta Crystallographica Section C: Crystal Structure Communications*, 48:2259–2262, 1992.

[32] M. C. Aragoni, M. Arca, F. A. Devillanova, M. B. Hursthouse, S. L. Huth, F. Isaia, V. Lippolis, A. Mancini, H. R. Ogilvie, and G. Verani. Reactions of pyridyl donors with halogens and interhalogens: An X-ray diffraction and FT-Raman investigation. *Journal of Organometallic Chemistry*, 690:1923–1934, 2005.

[33] A. S. Batsanov, J. A. K. Howard, A. P. Lightfoot, S. J. R. Twiddle, and A. Whiting. Stereoselective chloro-deboronation reactions induced by substituted pyridine-iodine chloride complexes. *European Journal of Organic Chemistry*, pages 1876–1883, 2005.

[34] M. Mascal, J. L. Richardson, A. J. Blake, and W. Li. Molecular structure of the S-Triazine-Br2 complex. *Tetrahedron Letters*, 37:3505–3506, 1996.

[35] R. Boese, A. D. Boese, D. Bláser, M. Y. Antipin, A. Ellern, and K. Seppelt. The surprising crystal packing of chlorinefluoride. *Angewandte Chemie International Edition in English*, 36:1489–1492, 1997.

[36] O. Hassel and K. O. Stromme. Crystal structure of the addition compound 1,4-dioxan-chlorine. *Acta Chemica Scandinavica*, 13:1775–1780, 1959.

[37] J. Gerratt and M. Raimondi. The spin-coupled valence bond theory of molecular electronic structure. I. basic theory and application to the $^2\Sigma^+$ states of BeH. *Proceedings of the Royal Society*, 371:525–552, 1980.

[38] G.F. Smits and C. Altona. Calculation and properties of non-orthogonal, strictly local molecular orbitals. *Theoretica Chimica Acta*, 67:461–475, 1985.

[39] A. Fornilli, M. Sironi, and M. Raimondi. Determination of extremely localized molecular orbitals and their application to quantum mechanics/molecular mechanics methods and to the study of intramolecular hydrogen bonding. *Journal of Molecular Structure*, 632:157–172, 2003.

[40] H. Stoll, G. Wagenblast, and H. Preuß. On the Use of Local Basis Sets for Localized Molecular Orbitals. *Theoretica Chimica Acta*, 57:169–178, 1980.

[41] R. Sedlak, K. E. Riley, J. Rezac, M. Pitonak, and P. Hobza. MP2.5 and MP2.X: Approaching CCSD(T) quality description of noncovalent interaction at the cost of a single CCSD iteration. *ChemPhysChem*, 14:698–707, 2013.

[42] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi,

M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. Gaussian 16 Revision C.01, 2016. Gaussian Inc. Wallingford CT.

[43] S. F. Boys and F. Bernardi. Calculation of small molecular interactions by differences of separate total energies - some procedures with reduced errors. *Mol. Phys.*, 19:553, 1970.

[44] S. Tsuzuki, M. Mikami, and S. Yamada. Origin of attraction, magnitude, and directionality of interactions in benzene complexes with pyridinium cations. *J. Am. Chem. Soc.*, 129:8656–8662, 2007.

[45] A. R. Voth, P. Khu, K. Oishi, and P. S. Ho. Halogen bonds as orthogonal molecular interactions to hydrogen bonds. *Nature Chemistry*, 1:74–79, 2009.

[46] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox. Gaussian 09 Revision D.01. Gaussian Inc. Wallingford CT 2009.

[47] Basis set exchange (https://www.basissetexchange.org/).

[48] M. Sironi, A. Genoni, M. Civera, S. Pieraccini, and M. Ghitti. Extremely localized molecular orbitals: Theory and applications. *Theoretical Chemistry Accounts*, 117:685–698, 2007.

Acknowledgements

I would like to thank my supervisor Prof. Maurizio Sironi and my co-tutors Dr. Alessandra Forni, Dr. Stefano Pieraccini and Dr. Alessandro Genoni. I also would like to thank my coworkers Francesco Oliva, Davide Franchini, Enrico Gandini, Giovanni Macetti, Erna Katarina Wieduwilt and Federico Dapiaggi.