UNIVERSITÀ DEGLI STUDI DI MILANO

DIPARTIMENTO DI CHIMICA

PhD Course in Chemistry – XXXI Cycle

Topological Descriptors Enabling Novel Dissections of Electron Position and Spin Properties in Complex Molecular Systems

Tutors:

Dr. Leonardo LO PRESTI

Dr. Carlo GATTI

PhD Candidate:

Giovanni MACETTI

Id. N. R11414

Academic Year 2017-2018

Referees:

- Prof. Evelio Francisco Miguélez, University of Oviedo, Spain
- Prof. Jacob Overgaard, University of Aarhus, Denmark

Committee Members:

- Prof. Piero Ugliengo, University of Turin, Italy
- Prof. Enrique Espinosa, University of Lorraine, France
- Prof. Alberto Vertova, University of Milan, Italy

Discussion:

Tuesday 29$^{\text{th}}$ January 2019, University of Milan

Abstract

Macroscopic and microscopic properties of molecular and solid-state systems are intimately related to the their electronic structure. The electron position and spin densities, which represent the probability distributions to find all or unpaired electrons in the space, contain information concerning several chemical-relevant properties, such as the chemical bonding and the magnetic behaviour. Understanding the fine atomic-level mechanism behind these properties is a key step to design chemical modifications to properly tune and develop materials or molecules with specific features. Topological descriptors can be used to extract information from these electron distributions.

In this work, novel applications of the source function descriptor have been developed to gain further insights on the electron and spin density-related properties. These developments, together with other topological descriptors, were used to get further insights on relevant chemical systems.

Firstly, the source function reconstruction was enlarged to a multi-dimensional grid of points with a particular focus on the two-dimensional maps. This analysis allows to see the ability of chosen subsets of atoms to reconstruct the density in the selected area within a *cause-effect* relationship and to rationalise the chemical or magnetic behaviours. The source function partial reconstructed maps depict if in a molecular region the atomic contributions are important, modest or negligible. Besides, they may also be useful for a proper selection of the reference points and for a full understanding of the source function percentages analysis. In fact, the choice of the reference point where to reconstruct the studied density is neither easy nor objective for non-standard situations, such as for the spin density. This novel application was applied to the study of the spin density on a couple of azido Cu complexes. The source function partial reconstructed maps allow to unravel the different role played by the paramagnetic centre Cu and the ligand atoms and to explain the spin transmission mechanism at a molecular level. Moreover, they enable to highlight the nature of the spin density differences between the two complexes and among adopted computational approaches. DFT functionals tend to over-delocalise the spin density towards the ligand atoms introducing a biased spin-polarization mechanism between the Cu and the ligand atoms.

The same descriptor was then applied to the study of the hydrogen bonds in the DNA base pairs. The source function reveals the delocalised nature of these interactions, highlighting that distant groups and rings have non-negligible effects on the reconstruction of

the electron density in the intermolecular region. Besides, the analysis demonstrates that the purine and pyrimidine bases equally contribute to the reconstruction of the electron density at the hydrogen bond critical points. The source function also reveals that subtle variations of the atomic source contributions occur when the pairs are ionized, revealing that sources and sinks effects redistribution plays an important role in the stabilization of the DNA base pairs.

The source function was also used to develop a method to extract full population matrices purely based on the electron density distribution and then amenable to experimental determination. The peculiar features of this descriptor, in particular the *cause-effect* relationship, assign a profound chemical meaning to the matrix elements in contrast with other population analyses such as the Mulliken's one, where the matrix elements are associated to orbital overlaps. The latest breakthroughs on the development of this method are shown together with some numerical examples on very simple compounds. The full population matrices obtained using the source function descriptor are able to retrieve the major chemical features.

A detailed analysis on the intermolecular interactions involved in the *in vivo* molecular recognition of the antimalarial drug chloroquine with the heme moiety has been carried out using a combined topological-energetic analysis. This work reveals that charged-assisted hydrogen bonds set up between the lateral chains of the chloroquine and the propionate group of the heme are the most important interactions in the drug:substrate recognition process.

Table of Contents

Chapter 1

Introduction

In this first chapter, I recall briefly some theoretical backgrounds concerning the main fields and tools used throughout the thesis.

The aim of this part is not to provide a complete and exhaustive review on all the concepts, since many papers and book chapters have already been written about. Instead, the general idea of this part is to give to the readers a brief and fresh overview of the plethora of tools used and to provide a list of references where interested readers can find more details.

The chapter starts from the general concept of quantum chemical topology and it moves towards different fields belonging to this ensemble, such as the Bader's quantum theory of atoms in molecules and the topological features and descriptors employed. A brief introduction on the scalar electron spin density function is sketched to prepare the ground for the deeper discussion that is held in a following chapter.

In the last part of this first chapter, the source function descriptor is discussed in more detail since it has a large importance in this work. Together with its theoretical background, new developments of this descriptor are introduced, trying to explain why they were needed and what problems were tackled. The detailed application of the source function descriptor for the study of the electron and spin density is shown in the following chapters.

1.1 Quantum Chemical Topology

Chemists and physicists are scientists that study matter and natural phenomena. Although they both use the same scientific approach, they very often disagree about the models used. Molecules are seen by physicist as a sum of particles (*e.g.* protons, neutrons and electrons), while chemists use to see a molecule as an aggregate of interacting atoms or group of atoms. Both models are able to grasp important aspects of reality but none of them is exhaustive. The general idea "the more complex it is, the better it is" is not always true.

Every chemist uses every day, often unconsciously, models. The probably most used

model one can think of is the Lewis' molecular model, developed in the last century by G. N. Lewis.[1] Although it is a very simple and old model, it is still used nowadays. Is this so surprising? Definitely not, because despite its simplicity, it gives useful and reliable information to those who only looks for something able to predict and rationalize chemical connectivity and bonding in most of the ordinary compounds. Of course, its simplicity is also its main drawback and it can not be used if one wants to gain deeper insights in all situations of bonding and of other molecular properties.

The XX and XXI centuries have been the "wonder age" for the blossoming of new methods and models, in chemistry and in all the other scientific disciplines. Among all of these, several methods have been proposed to explore and study the topology of quantum mechanical scalar and vector fields of relevance for chemistry.[2] The set of these methods is part of a branch of the theoretical chemistry called *quantum chemical topology* (QCT).

To confirm the importance of this field for the study of chemical and physical properties in molecules, in 2018 Prof. Paul L. A. Popelier from Manchester university (UK) has been awarded with the *Richard F. W. Bader international prize for excellence in electron density research* at the Sagamore 2018 conference held in Canada. The motivation for this choice was "for his exceptional contributions in the development of pure and applied quantum chemical topology".

Prof. Popelier defines the QCT as:

> a branch of theoretical chemistry that uses the language of dynamical systems (*e.g.* attractor, basin, homeomorphism, gradient path/phase curve, separatrix, critical points) to partition chemical systems and characterise them via associated quantitative properties.[3]

This statement is quite general and it does not refer to any specific field or system. It underlines three important features of QCT:

1. It is based on the *language of dynamical systems*. Technically speaking, a *dynamical system* is defined as a field of bound vectors V on a manifold M, and for all the points the derivative of their coordinates over the time ($d\boldsymbol{r}/dt$) determines a unique trajectory. These trajectories term where $d\boldsymbol{r}/dt$ is equal to zero. Apart from technicalities, we will see later in this chapter what this means.

2. It *partitions chemical systems*. QCT methods are able to decompose chemical systems into smaller pieces, such as a chemist usually partitions molecules into atoms or functional groups.

3. It *characterises chemical systems via associated quantitative properties*. It means that the properties of large systems can be seen as due to the sum of smaller terms, for instance the atomic contributions.

From these three concepts, one can immediately figure out why QCT is so important. It gives to chemists what they want. It allows to "see" the atoms into molecules and it gives

insight on how atoms interact with the environment (the other atoms in the molecule or crystal).[3] Several different tools have been developed in the last decades. They are able to gain details and explain different density-related properties, such as for instance chemical bonding, non-covalent interactions or reactivity of molecules.

A detailed discussion on the main methods belonging to the QCT field and their main outcomes is beyond the focus of this chapter and plenty of papers have been published about (see for example References [3],[4] and [5]). Here I just want to briefly recall some theoretical aspects of the most historically important QCT tool: the quantum theory of atom in molecules (QTAIM)[6–8]. This will be useful in the next chapters, where results obtained from the QTAIM and related descriptors are discussed.

Before moving on, it is essential to underline the fact that the terms (QT)AIM and QCT are sometimes misused and confused. As properly explained by Popelier, the term AIM is correctly used when the studied field is the electron density, because this is the only case where the atoms in the molecules are found.[9] The topological analyses of the Laplacian of the electron density or other fields like the electron localization function (ELF) or the reduced density gradient (RDG) do not yield atoms in molecules and can only be considered as QCT methods.

1.2 The Importance of the Electron Density

The electron density (ED), denoted in this thesis with the symbol $\rho(\boldsymbol{r})$, can be defined as the probability of finding an electron in a volume element $d\boldsymbol{r}$, independently from its spin and regardless of the position and spin of all other electrons of the system.[10] From the quantum mechanical point of view, the $\rho(\boldsymbol{r})$ is seen as the 3(N-1)-dimensional integral over the coordinates of all but one electron of the expectation value of the spinless wavefunction Ψ:

$$\rho(\boldsymbol{r}) = N \int \Psi^*(\boldsymbol{r}, \boldsymbol{r_2}, ..., \boldsymbol{r_N}) \Psi(\boldsymbol{r}, \boldsymbol{r_2}, ..., \boldsymbol{r_N}) \; d\boldsymbol{r_2} d\boldsymbol{r_3} ... d\boldsymbol{r_N} \tag{1.1}$$

where N is the total number of electrons in the system. The ED can also be expressed in term of basis functions Φ and first-order density matrix elements $P_{\mu\nu}$:

$$\rho(\boldsymbol{r}) = \sum_{\mu} \sum_{\nu} P_{\mu\nu} \Phi_{\mu}(\boldsymbol{r}) \Phi_{\nu}(\boldsymbol{r}) \tag{1.2}$$

The ED can be obtained from standard *ab initio* calculations accordingly to equation 1.2 once the wavefunction of the system is evaluated. It is interesting to note that given an ED, its properties are independent from the choice of the orbital representation. Molecular and natural orbitals, although completely different from a conceptual viewpoint, reconstruct the same electron density and retrieve the same properties.

Another advantage in using the electron density instead of the electronic wavefunction is the number of variables needed. While the wavefunction depends on $3N$ (spinless) or $4N$

(considering also the spin) variables, the ED is described only by three spatial variables (x, y and z). These aspects are at the basis of the density functional theory (DFT), where the two Hohenberg-Kohn (HK) theorems state that the external potential and the total energy of a system is a unique functional of the electron density and that the ground state energy can be obtained variationally: the density that minimises the total energy is the exact ground state density.[11] The HK theorems are important because they claim that both the electron density $\rho(\boldsymbol{r})$ and the wavefunction Ψ can describe the ground state of systems. That said, the ground-breaking point here is that the electron density is a quantum mechanics observable while the wavefunction is not.

The ED can experimentally be retrieved through elastic X-ray diffraction experiments. What one actually measures during a crystallographic experiment are the structure factors $F(\boldsymbol{H})$, where H indicates the Miller's indices of the plane *h, k, l*. The mathematical relation linking the electron density in a point **r** with the observed structure factors $F(\boldsymbol{H})$ is the following:

$$\rho(\boldsymbol{r}) = \frac{1}{V}\sum_{\boldsymbol{H}} F(\boldsymbol{H})e^{-2\pi i \boldsymbol{H}\,\cdot\,\boldsymbol{r}} \tag{1.3}$$

where V is the unit-cell volume. Using this relation, it is possible to build a model able to reproduce the experimental structure factors and the electron density of the crystal. In other words, to obtain a reasonable model it is necessary to minimize the difference density $\Delta\rho$ function:

$$\Delta\rho(\boldsymbol{r}) = |\rho_{obs}(\boldsymbol{r}) - \rho_{calc}(\boldsymbol{r})| = \frac{1}{V}\sum_{\boldsymbol{H}} \left|\left(F_{obs}(\boldsymbol{H})e^{-2\pi i \boldsymbol{H}\,\cdot\,\boldsymbol{r}} - F_{calc}(\boldsymbol{H})e^{-2\pi i \boldsymbol{H}\,\cdot\,\boldsymbol{r}}\right)\right| \tag{1.4}$$

For pure geometrical determination, standard X-ray diffraction experiments are sufficient. ρ_{calc} is obtained as a sum of non-interacting spherically densities centred on the atoms: this model is called independent atom model (IAM).

To take into account the effect of the chemical bonds and then to obtain a more accurate density, high-resolution data (usually around 0.5 Å) are necessary. These data allow to properly deconvolve the thermal motion of atoms and to separate it from the deformation of the electron distribution.[12] The reason of this has to be looked for in the behaviour of atomic scattering factors f, which is the ability of an atom to scatter X-ray light. At high Bragg's angle, the scattering factors bear only information concerning the core electrons, while at low angle the scattering intensity is due both to valence and core electrons. The high-angle reflections are then fundamental to properly evaluate atomic positions and thermal motions and consequently to obtain an accurate model.

Assumed to have sufficiently accurate data, the aspherical features of the ED can be reconstructed using a nuclei-centred finite multipole expansion.[13] Hansen and Coppens proposed in 1978 a revised version of this formalism, which is still used in some multipole refinement software.[14,15]

Using this formalism, the ED can be split into three components:

$$\rho(\boldsymbol{r}) = \rho_c(\boldsymbol{r}) + P_v\rho_v(\kappa\boldsymbol{r}) + \rho_d(\kappa'\boldsymbol{r}) \tag{1.5}$$

where $\rho_c(\boldsymbol{r})$ and $\rho_v(\boldsymbol{r})$ are the spherical part of the core and valence densities, respectively, while $\rho_d(\boldsymbol{r})$ is the aspherical valence part expressed as multipolar expansion

$$\rho_d(\kappa'\boldsymbol{r}) = \sum_l R_l(\kappa'\boldsymbol{r}) \sum_{m=-l}^{l} P_{lm}Y_{lm}(\frac{\boldsymbol{r}}{r}) \tag{1.6}$$

where R_l are radial functions and Y_{lm} are real spherical harmonics. κ and κ' are variables that take into account the radial expansion (or contraction) of the valence shell. P_v and P_{lm} are the population coefficients of the spherical and multipole part of the model, respectively. The electron density retrieved from accurate structure factors through the use of the multipole model is absolutely comparable to that obtained by good-quality *ab initio* calculations.

Thanks to the ED, theory and experiment have an intersection point and their results can be expressed on the same ground, making their comparison easier and more reliable. This is reason why analysis of the electron density is so important and a large number of topological descriptors based on this scalar field have been proposed and used.

1.3 The Electron Density Topology: Quantum Theory of Atoms in Molecules

The quantum theory of atoms in molecules (QTAIM) is an ED-based theory developed by the Canadian professor Richard F. W. Bader and co-workers.[6–8] Through the $\rho(\boldsymbol{r})$ scalar field, its gradient fields ($\nabla\rho(\boldsymbol{r})$ and the Laplacian $\nabla^2\rho(\boldsymbol{r})$), this theory is able to reformulate the classical chemical concepts (atom, bond, *etc.*) in a more clear and rigorous way.

The QTAIM is based on the electron density topology. The $\rho(\boldsymbol{r})$ exhibits four kinds of non-degenerated critical points (CPs), defined as the points where the gradient $\nabla\rho(\boldsymbol{r}) = 0$. Critical points can be classified according to two properties of their Hessian matrix (H): the rank R, *i.e.* the number of non-zero eigenvalues, and the signature S, *i.e.* the sum of the signs of the eigenvalues. The CPs are then classified by the pair (R, S).[6]

It is worth to recall that the Hessian is a matrix of second derivatives ($\nabla\nabla f$) of a function with respect to the coordinates. In the case of the electron density $\rho(\boldsymbol{r})$, the Hessian is a 3x3 matrix expressed as:

$$H(\boldsymbol{r}) = \nabla\nabla\rho(\boldsymbol{r}) = \begin{pmatrix} \frac{\partial^2\rho(\boldsymbol{r})}{\partial x^2} & \frac{\partial^2\rho(\boldsymbol{r})}{\partial x\partial y} & \frac{\partial^2\rho(\boldsymbol{r})}{\partial x\partial z} \\ \frac{\partial^2\rho(\boldsymbol{r})}{\partial y\partial x} & \frac{\partial^2\rho(\boldsymbol{r})}{\partial y^2} & \frac{\partial^2\rho(\boldsymbol{r})}{\partial y\partial z} \\ \frac{\partial^2\rho(\boldsymbol{r})}{\partial z\partial x} & \frac{\partial^2\rho(\boldsymbol{r})}{\partial z\partial y} & \frac{\partial^2\rho(\boldsymbol{r})}{\partial z^2} \end{pmatrix} \tag{1.7}$$

When a critical point is found, the Hessian matrix can be diagonalized and the corresponding *eigenvectors*, v_i, and *eigenvalues*, λ_i, can be obtained. The *eigenvectors* are three orthogonal vectors indicating the three principal axes of curvature of the $\rho(\boldsymbol{r})$, while the eigenvalues indicate the nature of the curvature in the studied point.[7]

The electron density Laplacian, $\nabla^2\rho(\boldsymbol{r})$, is a function which equals the sum of the eigenvalues of the Hessian, that is:

$$\nabla^2\rho(\boldsymbol{r}) = \lambda_1(\boldsymbol{r}) + \lambda_2(\boldsymbol{r}) + \lambda_3(\boldsymbol{r}), \quad \text{with } \lambda_1(\boldsymbol{r}) \leq \lambda_2(\boldsymbol{r}) \leq \lambda_3(\boldsymbol{r}). \tag{1.8}$$

It can be demonstrated that the Laplacian is an intrinsic property of the ED because it is invariant with respect to coordinate transformation.

Four different critical points of rank 3 can be found for the electron density and they can be classified as follows:

- (3,-3) CP: this CP has three negative eigenvalues (three negative curvatures) and correspond to the maximum of the ED. This kind of CP is found in correspondence of the position of the atomic nuclei. Since the (3,-3) CP is by topological definition an *attractor*, *i.e.* it is a point on which all the gradient lines converge, this point is called *nuclear attractor* (NA). This 1:1 relationship among nuclei and topological maxima is the demonstration that the most dominant force for the electron distribution is the electrostatic attraction of the nuclei. A (3,-3) CP can rarely appear far from the nuclei position and in this case it is called *non-nuclear attractor* (NNA).

- (3,-1) CP: the signature equal to -1 means that in such a point two eigenvalues are negative and one is positive. The CP is then a minimum in one direction and a maximum in the other two directions. Topologically speaking, this CP is a saddle point of index 1. In the case of the ED, it is called *bond critical point* (BCP) and it is found between two bonded atoms. The two gradient paths that start from the BCP and initially follow the maximum curvature along the *eigenvector* v_3 until the NAs are called *bond path* (BP).

- (3,+1) CP: this CP is a maximum in two directions and minimum in one. It is a saddle point of index 2 and correspond to a *ring critical point* (RCP) in the ED topology. It is found within a ring of atoms, like for instance in the middle of the benzene ring.

- (3,+3) CP: is a local minimum and it is a topological repellor. This CP is found where a cage is present and it is called *cage critical point* (CCP).

Table 1.1 summarises the four rank 3 non degenerate CPs detectable in the electron density distribution.

The number and the type of the CPs depend on the studied molecule. The only existing relationship among them is the Poincarè-Hopf relation:

Table 1.1: Summary of the four types CPs of rank 3 in the electron density topology.

Name		λ_1	λ_2	λ_3	(R,S)
(Non)-nuclear attractor	(N)NA	-	-	-	(3,-3)
Bond critical point	BCP	-	-	+	(3,-1)
Ring critical point	RCP	-	+	+	(3,+1)
Cage critical point	CCP	+	+	+	(3,+3)

$$n_{NA} - n_{BCP} + n_{RCP} - n_{CCP} = m \tag{1.9}$$

where n_{NA}, n_{BCP}, n_{RCP} and n_{CCP} are the number of (N)NAs, BCPs, RCPs and CCPs, respectively, and m is a constant equal to 1 for a molecule and 0 for a crystal. This mathematical formula is valid for any well-behaved three dimensional scalar function.[2]

1.3.1 QTAIM Atom

QTAIM defines the concept of atom as the union of a nucleus, *i.e.* the (3,-3) CP attractor, and its basin. A basin can be described in two different ways, depending on which characteristic one wants to highlight.

The first definition refers to the basin as the portion of the molecular space that is delimited by a surface S through which the gradient of the ED has zero flux in any point.[6–8,16] This condition can be written as:

$$\nabla\rho(\boldsymbol{r}) \cdot \boldsymbol{n}(\boldsymbol{r}) = 0 \qquad \forall \boldsymbol{r} \in S(\boldsymbol{r}) \tag{1.10}$$

where $\boldsymbol{n}(\boldsymbol{r})$ is the unit vector normal to the surface at $\boldsymbol{r}$.[16] The fulfilment of this condition highlights that the atom in a molecule defined according to QTAIM is a *proper quantum open system*.[6] The observables of QTAIM atom, such as the kinetic energy or the ED, are correctly described by the quantum mechanics equations of motion. Moreover, the *proper quantum open systems* meet the validity criteria for the virial theorem and all the other quantum mechanics theorems. Using the Gauss' divergence theorem[17], the surface-integral of Equation 1.10 can be rewritten as

$$\oint_{S_\Omega} \nabla\rho(\boldsymbol{r}) \cdot \boldsymbol{n}(\boldsymbol{r})\, dS = \int_\Omega \nabla^2\rho(\boldsymbol{r})\, d\tau = 0 \tag{1.11}$$

where now the zero-flux surface condition yields that the Laplacian of the electron density, $\nabla^2\rho(\boldsymbol{r})$, vanishes when integrated within such a basin. Equation 1.11 has an important practical consequence. In fact, since the integrated Laplacian has to be equal to zero, it can be used to check the quality of a numerical atomic integration. In principle, any deviation from zero by this integral can be considered as an integration error. Naturally, since analytical integrations are usually not possible and numerical methods are required, the Laplacian value can not be exactly null. In general, the lower the value of $\nabla^2\rho(\boldsymbol{r})$, the

more accurate the integration is.

On the other hand, a basin can be seen as the space traversed by all the ascending gradient paths of the electron density which terminate in a (3,-3) CP, an attractor (a nucleus). In simplest words, a basin is the ensemble of points that, when the electron density gradient is "climbed", reach the same maximum. The nucleus can be therefore defined as a sink of steepest ascent paths. Referring to Section 1.1, this definition coincides with that provided for the dynamical systems. Critical points different from the (3,-3) one can be both source and sink of ascending gradient paths according to their classification. For example, a BCP is a source when the direction toward its linked nuclei is considered, while it behaves as a sink in the other directions. Common points are passed through by only one gradient path.

Both of these classifications are equally valid and define the atoms using two complementary approaches. Nevertheless, it is clear that the full set of QTAIM atoms (including any non-nuclear basin, if present) provide an exhaustive partitioning of the space. Each atom has its own basin Ω, which is unique and does not overlap to any other. Two atoms are separated by a zero-flux surface defined as in Equation 1.10 and the points enclosed by a bounding interatomic surface uniquely belongs to only one domain.

Moreover, this exhaustive partitioning yields that the expectation value of any quantum mechanical observable (<A>) for the whole system is given by the sum of all the single atomic contributions:

$$< A >= \sum_{\Omega} A(\Omega) \tag{1.12}$$

where the sum extends over all the atoms of the molecule or crystal.

The two definitions of an atom are valid both in the gas phase for an isolated molecule and in the crystal. In the molecule, an interatomic surface defines the boundary between two atoms. However, since the isolated molecule does not have a finite volume, some atoms have a large "open" part that extend to infinity.[6–8] The molecule and the atoms are then open domains with infinite volumes. This is true for the majority of the atoms, while it can happen that some internal attractors are fully surrounded by interatomic surfaces. On the contrary, in a three-dimensional lattice each atom is always surrounded by other atoms so that no infinity boundaries are present. In this case, each basin has a proper volume and shape. Obviously, this last situation is only valid for an infinite repetition of atoms, while in the real cases this is true in the bulk but not onto the surfaces of the crystal.

1.3.2 Bond Path

The point onto the interatomic surface that reaches the maximum value of electron density is the bond critical point (BCP). As already shown before, the BCP is the (3,-1) saddle point that indicates the presence of a bond between two atoms. This point has a great importance in the context of QTAIM. The ensamble of the maximum density paths that

link the BCP with two nuclei is called *atomic interaction line* (AIL). If the system is at the equilibrium, that is the forces acting on the nuclei are null, the presence of an AIL is the necessary condition for two atoms to be considered bonded.[6,16,18] In this case, the AIL takes the name of bond path (BP).

It has been demonstrated that the BP between two nuclei is determined by quantum-mechanical exchange-correlation energies that always contribute to stabilize the interaction between the involved atoms. The bond path can be seen as a privileged exchange channel.[19]

The BP can also be defined as the line of maximum density, with respect to any lateral deviation, that links the nuclei of two atoms.[18] Along this line, the BCP is the point where the density attains its minimum value.[18] The network of nuclei and bond paths defines the so-called *molecular graph* that fully defines the chemical connectivity and the interactions between atoms.

Properties evaluated at the BCP, considered the most significant point along a bond path and on the interatomic surface, are able to disclose insights on the studied bond. Unfortunately, there is not a unique recipe to classify and characterise bond properties. Moreover, although the BCP is considered the most characteristic point able to describe a chemical interaction, it is still a local descriptor and can not take into account all the possible features influencing the bond.[16] That said, one must be aware that conclusions based only on the analysis of the BCP could not be complete or fully reliable. In fact, some recent works have criticised the ability of BCP and BP to describe the interactions between atoms.[20]

A detailed analysis of the major chemical bond descriptors is beyond the aim of this section. The interested readers can find detailed and clear highlights in References [5] and [16]. However, a brief discussion on some topological descriptors used in this work are reported in following sections.

1.4 Topology of the Electron Density Laplacian

As already anticipated in Section 1.3, the Laplacian of a scalar field in a specific point $\boldsymbol{r}$ is related to the sum of the Hessian eigenvalues (see Equation 1.8). The topology of the electron density Laplacian is more complicated than that of ρ. However, it carries a large amount of information that can be useful from a chemical point of view and that are not visible studying only the electron density distribution. For instance, through the local virial theorem $\nabla^2\rho(\boldsymbol{r})$ can be used to classify the nature of the chemical bond when evaluated at the BCP, or it provides a mapping of the shared electron pairs postulated in the Lewis model[1] and of the localized lone pairs assumed in the VSEPR model[21,22].

$\nabla^2\rho(\boldsymbol{r})$ measures the local curvature of the electron density distribution. Referring to Equation 1.8, Laplacian values greater than zero indicate that the positive λ autovalue dominate over the negative ones, *vice versa* if $\nabla^2\rho(\boldsymbol{r})$ is lower than zero. According to the divergence theorem (Equation 1.10), the sign of the electron density Laplacian indicates the accumulation/depletion of $\rho(\boldsymbol{r})$. $\nabla^2\rho(\boldsymbol{r})>0$ (positive sign) is physically associated with

an outward net flux of $\nabla\rho$ in an infinitesimal volume centred at $\mathbf{r}$: the electron density is *locally depleted* from that point. On the contrary, a negative value of $\nabla^2\rho(\boldsymbol{r})$ indicates that on average the inward flux is greater than the outward one: the electron density is *locally concentrated*.

The electron density in an isolated atom decreases monotonically from regions very close to the nucleus position, where it is maximum, to the infinite. Conversely, the electron density Laplacian oscillates more, showing alternating shells of charge concentration ($\nabla^2\rho(\boldsymbol{r}) < 0$) and charge depletion ($\nabla^2\rho(\boldsymbol{r}) > 0$). This alternation of minima and maxima is in general in correspondence with the quantum shells (K, L, *etc.*). The outer shell, called *valence shell*, shows an inner region where the Laplacian is negative and an outer one where it becomes positive and it goes to infinity. The former region is called *valence shell charge concentration* (VSCC, $\nabla^2\rho(\boldsymbol{r}) < 0$), while the latter is called *valence shell charge depletion* (VSCD, $\nabla^2\rho(\boldsymbol{r}) > 0$). When the atom forms chemical bonds, the spherical VSCC and VSCD distort forming maxima, minima and saddle points. The position, the type and the number of these points depend on the chemical environment. It is then clear that the study of the topology of the electron density Laplacian is extremely useful to gain further insights on the chemical bonding or other properties.

The study of the $\nabla^2\rho(\boldsymbol{r})$ field is generally performed in terms of the function $L(\boldsymbol{r}) = -\nabla^2\rho(\boldsymbol{r})$, so that charge localization is associated with positive $L(\boldsymbol{r})$ and *vice versa* for the charge delocalization.

(3,-3) $L(\boldsymbol{r})$ maxima in the VSCC region indicate, as already anticipated above, the location of the bonded and non-bonded pairs of the Lewis and VSEPR models. Local electron density Laplacian maxima and minima in the valence shell can be associated with chemical reactivity.[6–8] From the above association between $L(\boldsymbol{r})$ maxima and electron lone pairs, local charge concentration behaves as a *Lewis base* or *nucleophile*. On the other hand, local charge depletion acts as a *Lewis acid* or an *electrophile*. That said, the formation of a chemical bond is associated with the presence of (3,-3) $L(\boldsymbol{r})$ maximum (or (3,+3) $\nabla^2\rho(\boldsymbol{r})$ minimum) on the nucleophile molecule and of a facing (3,+3) $L(\boldsymbol{r})$ minimum (or (3,-3) $\nabla^2\rho(\boldsymbol{r})$ maximum) on the electrophile. The three-dimensional disposition of the bonds, predicted by the simple VSEPR theory, is defined by the alignment of maxima and minima of the Laplacian distribution. Several papers in literature have used this predictive properties to rationalize peculiar bonds both in molecules and in crystals.[4,6,16,23,24]

Another feature that the electron density Laplacian is able to unravel is the asphericity of the metal electrons distribution as a consequence of the incomplete filling of d-orbitals.[16,24] In this case the Laplacian distribution is able to give a physical proof of the ligand field theory. (3,+3) minima of $L(\boldsymbol{r})$, VSCD points, are usually located on the metallic centres with the geometry predicted by the ligand field theory. These points are aligned to the VSCC regions located on the ligand atoms. The electron density Laplacian distribution is then able to show the so-called *key-lock* mechanism between a nucleophile and an electrophile.

A last important feature is worth to be mentioned. The local expression of the virial

theorem[6,25] states the following mathematical relation (in atomic units):

$$\frac{1}{4}\nabla^2\rho(\boldsymbol{r}) = 2G(\boldsymbol{r}) + V(\boldsymbol{r}). \tag{1.13}$$

$G(\boldsymbol{r})$ is the positive kinetic energy density evaluated in the point $\boldsymbol{r}$, while $V(\boldsymbol{r})$ is the potential energy density. The former physical quantity is defined positive everywhere, while on the contrary the latter is always negative. The electron density Laplacian is then the results of two opposing quantities and it is the expression of which energy density, kinetic or potential, locally prevails. Positive values of $\nabla^2\rho(\boldsymbol{r})$ indicates that the positive kinetic energy density is higher, on an absolute scale, than half the potential one. On the contrary, a negative $\nabla^2\rho(\boldsymbol{r})$ value indicates that the potential term is higher in absolute value than twice the kinetic energy term. This property is exploited in the classification of the bonding, as it is briefly shown in the following sections.

Recalling the definition of $\nabla^2\rho(\boldsymbol{r})$ expressed in Equation 1.8, one can associate the prevalence of $G(\boldsymbol{r})$ or $V(\boldsymbol{r})$ thinking about the curvature of the electron density in the space. If $\nabla^2\rho(\boldsymbol{r}) > 0$, the positive curvature at the BCP along the bond path prevails and the kinetic energy density term prevails. When $\nabla^2\rho(\boldsymbol{r}) < 0$, the two negative curvatures perpendicular to the bond path prevail and the potential energy density term dominates. As it is shown further, this fact has a large implication on the classification of chemical bonds.

1.5 Non-Covalent Interactions

The term *non-covalent interactions* (NCIs) is a very wide concept that collects a large number of types of intermolecular interactions where a negligible electron sharing occurs between the strongly bonded moieties.[5] The interest on this kind of bonding is extremely high because NCIs are fundamental in several fields, from material science to *in vivo* study of the drug activity. However, their elusive nature makes them difficult to detect and study. Several descriptors were proposed in the last decades for the study of NCIs, many of them based on the electron density scalar field.[4,5,16] In the next sections, some topological and energetic descriptors used in the next chapters are introduced and briefly discussed. A detailed overview is outside the scope of this introductory chapter. The interested readers can found more information in References [4, 5, 16].

1.5.1 Bonds Classification Through BCP Properties

As already said in the previous sections, BCP is usually taken as the most representative point for the study of chemical bonds and its properties can be used to classify the nature of the chemical interactions.

Bader and Essen[26] proposed a criterion to classify chemical interactions as *shared* or *not shared* based only on the value of the electron density Laplacian at the BCP. As already seen in Section 1.4, the sign of $\nabla^2\rho(\boldsymbol{r})$ (or $L(\boldsymbol{r})$) indicates whether the $\rho(\boldsymbol{r})$ accumulates in

or depletes from the BCP. Moreover, the same value indicate which energy density, namely positive kinetic (G) or potential (V), dominates in that point. $\nabla^2\rho(\boldsymbol{r}) < 0$ values describe *shared* interactions (covalent or polar) because it is associated with an accumulation of density in the interatomic region and with a dominating potential energy density. On the contrary, $\nabla^2\rho(\boldsymbol{r}) > 0$ refers to *closed-shell* interactions (ionic and hydrogen bonds, van der Waals contacts, *etc.*) because the ED is preferentially depleted from the BCP towards the atoms and the interaction is dominated by the kinetic energy density G. The difference of these two types of bonds can also be seen looking at the $\rho(\boldsymbol{r})$ values and at the components of the kinetic energy density G. In particular, *closed-shell* interactions are associated with a low value of ED at the BCP and with a parallel component of G, $G_{\parallel}$, greater than the perpendicular ones. To compare the two kinds of interaction on the same ground, it is useful to consider the quantity G_{BCP}/ρ_{BCP}. *Shared* interactions are associated with $G_{BCP}/\rho_{BCP} < 1$, while the *non-shared* ones exhibit a value greater than one. Table 1.2 summarises the features that distinguish the two different interactions.

This simple classification only distinguishes bonds in two separate types and it is not able to treat intermediate situations. Espinosa and co-workers suggested to use the ratio $|V_{BCP}|/G_{BCP}$ as a criterion for the bond classification.[27] Pure *shared* and *closed-shell* interactions are identified by $|V_{BCP}|/G_{BCP}$ values greater than 2 and lower than 1, respectively, accordingly to Equation 1.13. The *transition* region is characterised by intermediate values, together with positive electron density Laplacians. The introduction of the electron energy density (H)[28] (Equation 1.14) and in particular of the so called *bond degree parameter* (BD)[27] (Equation 1.15), which is the ratio between H and the electron density value at the bond critical point (ρ_{BCP}), has allowed a more flexible bond classification.

$$H_{BCP} = G_{BCP} + V_{BCP} \tag{1.14}$$

$$BD = \frac{H_{BCP}}{\rho_{BCP}} \tag{1.15}$$

The BD is negative both in the pure *shared-shell* and in the *transition* region, while it is positive for *closed-shell* interactions. The difference between the *shared-shell* and the *transition* region is associated to the value of the electron density Laplacian: the former shows negative values, while the latter is characterized by positive values. The degree of covalency of a bond depends on the value of the BD parameter: the more negative the BD is at the BCP, the more covalent is the interaction.

Table 1.2 summarises the criteria useful to classify the chemical bond according to what presented in this section.

1.5.2 Reduced Density Gradient

Johnson and co-workers proposed the reduced density gradient (RDG) as a descriptor to evaluate the presence and the relative strength of non-covalent interactions.[29,30] The RDG

Table 1.2: summary scheme of the bonds classification according to Bader & Essén[26] and Espinosa *et al.*[27] criteria. All the quantities are evaluated at the bond critical point of the studied interaction.

Criterion	Descriptor	Shared-shell (SS)	Transition	Closed-shell (CS)
Bader & Essén	$\nabla^2\rho_{BCP}$	< 0	//	> 0
	$G_{BCP,\parallel}$	$\ll G_{BCP,\perp}$	//	$\gg G_{BCP,\perp}$
	ρ_{BCP}	Large	//	Small
Espinosa	$\lvert V_{BCP}\rvert / G_{BCP}$	> 2	$1 < \mathrm{x} < 2$	< 1
	$\nabla^2\rho_{BCP}$	< 0	> 0	> 0
	BD	$\ll 0$	< 0	> 0
	Covalence degree	High	Any Value	//
	ρ_{BCP}	Large	Large	Small

is mathematically defined as:

$$RDG(\boldsymbol{r}) = \frac{1}{2\,(3\pi^2)^{1/3}} \frac{|\nabla\rho(\boldsymbol{r})|}{\rho(\boldsymbol{r})^{4/3}} \tag{1.16}$$

The RDG quantifies the local departure of the electron density from an homogeneous distribution (an ideal electron gas). In this ideal case, $\nabla\rho(\boldsymbol{r})$ is everywhere null and the RDG descriptor is consequently everywhere equal to zero. In real systems, RDG assumes values which span from zero, where the gradient is null, to large values, far from the nuclei where the numerator ($\nabla\rho(\boldsymbol{r})$) drops to zero faster than the denominator ($\rho(\boldsymbol{r})^{4/3}$) in Equation 1.16.

Being based only on the electron density and its gradient, the RDG is amenable to theoretical and experimental determination. It has been shown that ED obtained from multipole refinements and from *ab initio* calculations present the same RDG features.[31]

This quantity was introduced for the first time in the context of DFT as an independent variable of the generalized gradient approximation (GGA) for the definition of exchange-correlation functionals.[32–34] In 2010, Johnson *et al.* realised that RDG low-values isosurfaces in low-density regions are able to reveal the presence of non-covalent interactions.[29] The shape of the RDG isosurfaces depends on the type of the interaction and it can provide chemical insights on its nature. For instance, strong hydrogen bonds usually show *lenticular-shaped* RDG isosurfaces between the hydrogen and the acceptor atom, while C-H$\cdots$ring interactions very often present *conical-shaped* RDG regions. In general, strong interactions present small and well-defined domains while weak contacts are characterized by broader and multiform RDG regions.[5,31,35]

Low RDG regions are by definition found around the electron density critical points of the systems and they share, at least in part, the same features. However, RDG sometimes is able to highlight NCIs where no atomic interaction lines are found.

Qualitative information on the nature and the strength of the NCIs can be obtained plotting the value of the function $\rho(\boldsymbol{r}) \cdot sign(\lambda_2)$ on RDG isosurfaces. $\rho(\boldsymbol{r})$ is the value of the electron density evaluated at the surface point, while $sign(\lambda_2)$ is the sign of the second

curvature of the Hessian matrix (see Equation 1.7 in Section 1.3) at the same point.

The sign of λ_2 eigenvalue is assumed to indicate the nature of the NCI: if $\lambda_2 > 0$ the interaction is repulsive, while negative λ_2 values are associated with attractive interactions.[29] The value of $\rho(\boldsymbol{r})$ is a measure of the strength of the interaction: the higher is the ρ value, the stronger is the NCI.

Through this descriptors it is possible to compare different interactions and disclose which one is the most attractive/repulsive one.

1.5.3 Zhikol's Functional

As it was shown in the previous sections, non-covalent intermolecular interactions are usually studied evaluating properties at the bond critical point or using some non-local descriptors, such as the reduced density gradient. While some interactions, like the hydrogen-bonds, can be easily studied using these descriptors, the same is not true for the ring-ring stacking interactions, where the π electronic system is involved.

The dispersion/repulsion energy is often the main important term in the stabilization of complexes, like in the case of DNA double strand. However, the evaluation of its energy is not straightforward and it is extremely sensitive to the level of theory and basis set used. Commonly used DFT functionals usually fail to properly evaluate dispersion contributions and empirical or semi-empirical corrections are required to obtain reasonable interaction energies.[36] More accurate quantum mechanical methods, such as the Møller-Plesset MP2 perturbation approach, are necessary to obtain a more accurate description of the dispersion/repulsion interactions. However, for large systems these methods are too computational demanding and they can not be used.

In 2005, Zhikol and co-workers proposed a functional based on topological features fitting the interaction energies evaluated at the MP2/6-31G+(2p,2d) level of theory for a series of stacked benzene dimers. Through this functional, classical DFT functionals can be used to evaluate the $\pi \cdots \pi$ interaction energy.

The proposed functional exploits the electron density properties evaluated at the ring (RCP) and cage critical points (CCP) found in the intermolecular region between the stacked rings:

$$E_{\pi\cdots\pi} = \sum_i C_i\{1.214(14) \cdot \nabla^2\rho_i - 131.6(3.9) \cdot (\nabla^2\rho_i)^2 + 635(84) \cdot \rho_i^2\}. \qquad (1.17)$$

In Equation 1.17, $\nabla^2\rho_i$ is the Laplacian of the electron density ρ_i evaluated at the i-th cage critical point and C_i is the cosine of the angle between the CCP and the two RCPs of its closest aromatic rings. The summation includes all the CCPs related to the ring stacking.

In principle, no physical reasons are known to produce a numerical correlation of ED properties in the CCPs with pairing energy, but for the intermolecular region between

stacked rings such correlation has been found with a satisfactory level of accuracy.[37]

1.5.4 NBO Analysis

Another way to evaluate the energy of intermolecular interactions, in particular hydrogen bonds, is to analyse the natural bond orbitals (NBOs).

In a nutshell, NBOs are an orthonormal set of localized maximum-occupancy orbitals describing the Lewis-like bonding pattern of the electron pairs. The localization of these orbitals occurs on one, two or rarely three atoms. NBOs are obtained through linear combination of natural atomic orbitals (NAOs), which are natural orbitals (NOs) centred on the atoms.

NBOs can be classified according to the number of centres (atoms) on which they are localized (1- or 2-centres) and to their nature (*Lewis-type* and *non-Lewis-type*). *Core* (C_A) and *nonbonded* (n_a) NBOs are 1-centre *Lewis-type* orbitals that clearly describe the core and the valence lone pairs electrons on the atoms. *Non-Lewis-type* 1-centre orbitals are the so-called *Rydeberg* and *unfilled nonbonded* NBOs which describes other features, usually not important from a chemical point of view.

Within the family of the 2-centres NBOs we can found the *bond* (*Lewis-type*) and the *antibond* (*non-Lewis-type*) orbitals. The former are typically found between covalent- or ionic-bonded atoms, while the latter are usually found to be important non-Lewis acceptor orbitals, contributing to the intermolecular hydrogen bonds or other donor-acceptor interactions.

The second-order perturbation theory analysis of Fock matrix in NBOs basis allows to disclose the presence of important interactions among the orbitals. Usually, hydrogen bonds are represented by the interaction between the *lone pair* NBO localised on the H-bond acceptor atom and the *antibond non-Lewis-type* NBO localized on the bond between the H-bond donor and the hydrogen atom. The estimation of the energy associated with this interaction can be evaluated according to the formula:

$$E(2) = \frac{q_i \cdot F(i,j)^2}{(\varepsilon_i - \varepsilon_j)} \tag{1.18}$$

where q_i is the donor orbital occupancy, ε_i and ε_j are the eigenvalues (energies) of the donor and acceptor orbitals, respectively, and $F(i,j)$ is the the off-diagonal NBO Fock matrix element.

1.5.5 Experimental Charge Density Approach

The total interaction energy between two molecules can be decomposed in terms of different contributions. Classically, when using an experimental charge density approach, at least three terms are defined: *electrostatic* (E_{es}), *dispersive* (E_{disp}) and *repulsive* (E_{rep}).

$$E_{int} = E_{es} + E_{disp} + E_{rep} \tag{1.19}$$

The dispersive and repulsive terms are the most difficult contributions to evaluate because an analytical form of these energies does not exist. Usually they are treated using *ad hoc* empirical or semi-empirical functionals calibrated to correctly reproduce experimental properties values.[36,38–41]

Instead, the electrostatic term can be expressed in term of Coulomb interaction between two charge distributions ρ_A and ρ_B:

$$E_{es} = \int_A \int_B \frac{\rho_A(\boldsymbol{r}_A)\rho_B(\boldsymbol{r}_B)}{|\boldsymbol{r}_A - \boldsymbol{r}_B|} d\boldsymbol{r}_A d\boldsymbol{r}_B. \tag{1.20}$$

Although exact, this function is not very often used because of the six-dimensional integration required. Usually, the electrostatic interaction is evaluated using multipole expansion models centred on the atomic nuclei in the molecules. Volkov and co-workers have demonstrated that the interaction energy obtainable from the multipole model approach is usually lower than that obtainable from an analytical integration.[42,43] This discrepancy arises from the interpenetration of the electron clouds of different molecules, contribution that is not taken into consideration in the multipole expansion.

Recently, Spackman has proposed an approach for the evaluation of the electrostatic energy taking into account the penetration correction.[38,44–47] In his model, Spackman suggests to decompose the electron density distribution in two terms: a spherical not-interacting term, defined *promolecular term* ρ_A^{pro}, and a deformation term $\Delta\rho_A$:

$$\rho_A(\boldsymbol{r}_A) = \rho_A^{pro}(\boldsymbol{r}_A) + \Delta\rho_A(\boldsymbol{r}_A) \tag{1.21}$$

The promolecular term is obtainable as the sum of the spherical atomic densities and it accounts for the larger part of the penetration contribution,[48] while the deformation term accounts for the anisotropic distribution of the electron cloud due to chemical bonding and NCI interactions.

Substituting Equation 1.21 into 1.20, one obtains:

$$E_{es} = \int_A \int_B \frac{\rho_A^{pro}\rho_B^{pro}}{|\boldsymbol{r}_A - \boldsymbol{r}_B|} d\boldsymbol{r}_A d\boldsymbol{r}_B + \int_A \int_B \frac{\rho_A^{pro}\Delta\rho_B + \rho_B^{pro}\Delta\rho_A}{|\boldsymbol{r}_A - \boldsymbol{r}_B|} d\boldsymbol{r}_A d\boldsymbol{r}_B + \int_A \int_B \frac{\Delta\rho_A\Delta\rho_B}{|\boldsymbol{r}_A - \boldsymbol{r}_B|} d\boldsymbol{r}_A d\boldsymbol{r}_B = E_{pro-pro} + E_{pro-def} + E_{def-def} \tag{1.22}$$

The electrostatic energy is now seen as the sum of three contributions. The $E_{pro-pro}$ is calculated as the Coulomb contribution between the promolecular non-interacting spherical distributions. The $E_{pro-def}$ takes into account the interaction between the deformation density of one molecule and the promolecular term of the other and *vice versa*. The last term, $E_{def-def}$, accounts for the interaction between the atom-centred multipoles that describes the deviation from the spherical symmetry of the electron density. Multipoles are usually evaluated up to l = 4 in terms of Cartesian tensor formulation of Buckingham.[49]

The total interaction energy within this model is obtained summing all the atom-atom

contributions of the terms reported in the previous equations. This decomposition in atomic contributions allows one to evaluate the strength of specific intermolecular interactions. For instance, in the case of the hydrogen bond, the energy of interaction can be evaluated considering the atom-atom contributions between the H and the acceptor atoms.

1.6 Spin Density

The *electron spin density* (SD) $s(\boldsymbol{r})$ (hereinafter *spin density*) is defined as the probability of finding an unpaired electron in a volume element $d\boldsymbol{r}$.[10] While in the case of the electron density $\rho(\boldsymbol{r})$ this probability is independent on the spin, $s(\boldsymbol{r})$ indicates the probability to find those electrons that are not formally paired in the classical double-occupied orbitals approach. The majority of the molecules are singlet state in their ground state and they do not show unpaired electrons, which means that their spin densities are null. However, metal atoms or some peculiar molecules, such as the oxygen O_2 molecule, can show ground-states of higher spin multiplicity, such as *doublet* or *triplet*. In these cases, the spin density distributions are not zero.

Using a spin-resolved formalism, the electron and spin densities can be written as a linear combination of the two spin-resolved densities, indicated as $\rho_\alpha(\boldsymbol{r})$ and $\rho_\beta(\boldsymbol{r})$. As it is clear, these two densities are related to the probability to find an electron with spin α $(+1/2)$ or β $(-1/2)$ in the space, respectively. In this context, the electron density $\rho(\boldsymbol{r})$ is the sum of the spin-resolved densities $\rho_\alpha(\boldsymbol{r})$ and $\rho_\beta(\boldsymbol{r})$ (Equation 1.23). On the contrary, the spin density $s(\boldsymbol{r})$ is their difference (Equation 1.24).

$$\rho(\boldsymbol{r}) = \rho_\alpha(\boldsymbol{r}) + \rho_\beta(\boldsymbol{r}) \tag{1.23}$$

$$s(\boldsymbol{r}) = \rho_\alpha(\boldsymbol{r}) - \rho_\beta(\boldsymbol{r}) \tag{1.24}$$

Although $\rho(\boldsymbol{r})$ and $s(\boldsymbol{r})$ are defined by the same densities, they carry different information. The electron density contains chemical information on the bonds, the reactivity and other related properties. On the contrary, the spin density gives information on the magnetic properties of the systems. By studying them it is the possible to disclose complementary features.

Electron and spin density are also largely different from a topological point of view. While the former has been largely studied since Bader's theory came out for the first time and it has been largely used to tackle chemical issues,[16] the latter has a more complicated topology and, to the best of my knowledge, no work has thus far been published on such topology in the literature. However, in a recent MSc thesis work, Bruno and others have tried to disclose similarities/differences between $\rho(\boldsymbol{r})$ and $s(\boldsymbol{r})$ studying the simple, but not trivial, triplet state of the water molecule.[50]

The spin density has by definition lower values than the electron density. This is associated with the number or electrons considered in the two cases. The integration of

$\rho(\boldsymbol{r})$ in the whole three-dimensional space retrieves the total number of electrons. On the contrary, the integraton of $s(\boldsymbol{r})$ yields the number of unpaired electrons which for a common organic radical or molecular magnet compound is usually equal to one or two. For instance, integrating the two densities for the water molecule in its triplet state, one retrieves 10 and 2 e for the $\rho(\boldsymbol{r})$ and $s(\boldsymbol{r})$, respectively. This divergence increases dramatically if heavier elements, such as transition metal atoms, are present in the molecule or if the number of atoms is very high. Even if this is not a big deal from a theoretical point of view, it implies some drawbacks in numerical treatment of the spin density.

Spin density, such as the electron density, can be obtained both from theoretical calculations and experiments. For closed-shell systems, the quantum chemical methods are able to predict properties in a fairly accurate way.[51–53] Unfortunately, the same is not true for open-shell systems, particularly for metal complexes and clusters.[54] Some of the most used theoretical approaches, such as Hartree-Fock (HF) and the largely used density functional theory (DFT) approach very often fail in these situations. In particular, the latter have shortcomings related to energy, geometry and more important spin density accuracy, which is qualitatively incorrect in some cases.[53]

According to Hohenberg-Kohn theorem, only the total electron density is required to obtain the exact ground-state of a system.[11] The spin state and the spin-resolved ρ_α and ρ_β densities are not necessary in the minimization process. The standard DFT exchange-correlation functionals are usually calibrated only on the properties related to the electron density and not on the spin density. This is the main reason why DFT fails to reproduce the spin density distribution.[53,55] Multireference post-HF methods, such as CASSCF, CASPT2 or DMRG[56,57], are usually required to obtained accurate enough spin density distributions. However, these methods can not be always applied to large systems because they are computationally demanding.

From an experimental point of view the situation is even worse. The spin density can be reconstructed using the reflections obtained from the polarized neutron diffraction technique through the use of a modified multipolar model.[58–61] The complex experimental set-up together with the intrinsic limitation of this technique (small number of reflections, neutron source required, *etc.*) exponentially increase the difficulty to retrieve experimental spin densities from single crystals.

A detailed analysis on the difficulty to obtain accurate spin densities and the disagreement obtained when different levels of theory are used are discussed in more details in Chapter 2.

1.7 Source Function Descriptor

As seen in the previous sections, the electron and the spin density (plus their associated fields) are two important features to study properties of molecules or crystals. Within the QTAIM and more generally the QCT, in 1998 Richard F. W. Bader and Carlo Gatti proposed an interesting topological descriptor: the *source function* (SF).[62]

In their seminal paper, they showed that the electron density $\rho(\boldsymbol{r})$ at any point $\boldsymbol{r}$ can be described as determined by the contributions from a function called *local source* (LS), which operates in any point $\boldsymbol{r'} \neq \boldsymbol{r}$ in the whole space R^3:

$$\rho(\boldsymbol{r}) = \int_{R^3} LS(\boldsymbol{r}, \boldsymbol{r'}) d\boldsymbol{r'} = \int_{R^3} -\frac{1}{4\pi} \frac{\nabla^2 \rho(\boldsymbol{r'})}{|\boldsymbol{r} - \boldsymbol{r'}|} d\boldsymbol{r'}. \tag{1.25}$$

This formula puts the ED Laplacian and the ED within a *cause-effect* relationship. The *cause* here is the Laplacian $\nabla^2\rho(\boldsymbol{r'})$, which indicates how the electrons concentrate in or dilute from the point $\boldsymbol{r'}$, and the *effect* is the reconstructed $\rho(\boldsymbol{r})$. The factor $(4\pi\,|\boldsymbol{r} - \boldsymbol{r'}|)^{-1}$ is a Green function that works as a weighting factor and it measures the effectiveness of the *cause-effect* relationship.[17] Although interesting, Equation 1.25 alone is not exploitable to improve the chemical sense in the analysis of the electron density field. When the total space is partitioned according to the QTAIM criteria (Equation 1.10), the $\rho(\boldsymbol{r})$ is reconstructed as a sum of atomic (Ω) LS contributions:

$$\rho(\boldsymbol{r}) = \sum_{\Omega} \int_{\Omega} -\frac{1}{4\pi} \frac{\nabla^2 \rho(\boldsymbol{r'})}{|\boldsymbol{r} - \boldsymbol{r'}|} d\boldsymbol{r'} = \sum_{\Omega} S(\boldsymbol{r}, \Omega) \tag{1.26}$$

$$S(\boldsymbol{r}, \Omega) = \int_{\Omega} -\frac{1}{4\pi} \frac{\nabla^2 \rho(\boldsymbol{r'})}{|\boldsymbol{r} - \boldsymbol{r'}|} d\boldsymbol{r'}. \tag{1.27}$$

The single basin contribution, $S(\boldsymbol{r}, \Omega)$ is called *source function of* Ω *to* $\boldsymbol{r}$. It defines the contribution that an atom gives to the reconstruction of the electron density in the selected reference point (RP) $\boldsymbol{r}$. Although any mutual exclusive or fuzzy space partitioning scheme could be used to define the SF contributions, the use of the QTAIM partitioning recipe ensures a quantum mechanical rigorous association (see Section 1.3.1).[63]

Equation 1.26 can be further splitted into inner $S_{in}(\boldsymbol{r}, \Omega)$ and outer $S_{out}(\boldsymbol{r}, \Omega)$ contributions, where the adjective inner refers to the basin Ω hosting the RP where reconstruction is analysed:

$$\rho(\boldsymbol{r}) = S_{in}(\boldsymbol{r}, \Omega) + \sum_{\Omega' \neq \Omega} S_{out}(\boldsymbol{r}, \Omega') \tag{1.28}$$

By requiring only the Laplacian of the electron density, the SF descriptor is amenable to experimental determination from X-ray diffraction experiment and multipole model refinement.[4,14,15,64] Then, the SF analysis represents a natural choice for comparing on the same ground results coming from theoretical and experimental electron density distributions, provided that the quality of the electron density Laplacian is accurate enough.[21,63,65]

The source function is in principle a powerful topological descriptor because it allows to see the $\rho(\boldsymbol{r})$ under a different light: not as a local property how usually it is referred to, but instead as a more delocalised phenomenon. As already discussed previously, the Hohenberg and Kohn theorems states that the electron density is uniquely mapped to the external potential, which is defined by the position of the nuclei and by their charges. The SF descriptor has the ability to bring this expression on a chemical field, showing if an

atom or a group atoms are relevant in the definition of the electron density at $\boldsymbol{r}$ and to what extent.

From a qualitative point of view, the SF is able to discriminate a basin acting as a *source* of electron density ($S(\boldsymbol{r},\Omega) > 0$, the basin injects electrons in $\boldsymbol{r}$) or as a *sink* ($S(\boldsymbol{r},\Omega) < 0$, the basin withdraws electrons from $\boldsymbol{r}$). Quantitatively, the larger in absolute value the SF contribution, the stronger the effect on the reference point $\boldsymbol{r}$. As clear from equation 1.25, the *source/sink* behaviour of a basin is strictly related to its $\nabla^2\rho(\boldsymbol{r'})$ distribution. The points of the basin where the ED concentrates ($\nabla^2\rho(\boldsymbol{r'}) < 0$) act as *sources*, *vice versa* if the ED is depleted ($\nabla^2\rho(\boldsymbol{r'}) > 0$) the points act as *sinks* of electrons. The final source function effect depends on the distribution of the *source* and *sink* points within the basin. Table 1.3 summarises the effect of the Laplacian in a point $\boldsymbol{r'}$ in the reconstruction of the electron density in $\boldsymbol{r}$.

Table 1.3: Effect of $\nabla^2\rho(\boldsymbol{r'})$ and $L(\boldsymbol{r'}) = -\nabla^2\rho(\boldsymbol{r'})$ on the reconstruction of the electron density $\rho(\boldsymbol{r})$.

$\nabla^2\rho(\boldsymbol{r'})$	$L(\boldsymbol{r'})$	$LS(\boldsymbol{r},\boldsymbol{r'})$	Effect on $\rho(\boldsymbol{r})$
-	+	+	*source* ($\rho(\boldsymbol{r})$ increases)
+	-	-	*sink* ($\rho(\boldsymbol{r})$ decreases)

1.8 Source Function for the Electron Spin Density

Although all the equations shown in the previous section refer to the ED, one can in principle use the source function descriptor to study any scalar field. Three years ago, Gatti an co-workers applied for the first time the SF descriptor to the study of the spin density on the triplet state of the water molecule.[66] The extension to the SD case is mathematically straightforward. One just needs to replace $\rho(\boldsymbol{r})$ with $s(\boldsymbol{r})$ and $\nabla^2\rho(\boldsymbol{r'})$ with $\nabla^2 s(\boldsymbol{r'})$, while the Green's function $4\pi\left|\boldsymbol{r}-\boldsymbol{r'}\right|^{-1}$ is independent of the studied scalar field:

$$s(\boldsymbol{r}) = \int_{R^3} LS_s(\boldsymbol{r},\boldsymbol{r'})d\boldsymbol{r'} = \sum_{\Omega}\int_{\Omega} LS_s(\boldsymbol{r},\boldsymbol{r'})d\boldsymbol{r'} = \sum_{\Omega} S_s(\boldsymbol{r},\Omega). \tag{1.29}$$

$$S_s(\boldsymbol{r},\Omega) = \int_{\Omega} -\frac{1}{4\pi}\frac{\nabla^2 s(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|}d\boldsymbol{r'} \tag{1.30}$$

$S_s(\boldsymbol{r},\Omega)$ is the *source function spin density* (SFSD), which is the contribution from the atom Ω to the reconstruction of the spin density in the point $\boldsymbol{r}$. The local source becomes:

$$LS_s(\boldsymbol{r},\boldsymbol{r'}) = -\frac{1}{4\pi}\frac{\nabla^2 s(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|} = -\frac{1}{4\pi}\frac{\nabla^2\left[\rho_\alpha(\boldsymbol{r'})-\rho_\beta(\boldsymbol{r'})\right]}{|\boldsymbol{r}-\boldsymbol{r'}|} = \frac{1}{4\pi}\frac{\nabla^2\rho_\beta(\boldsymbol{r'})-\nabla^2\rho_\alpha(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|} \tag{1.31}$$

In the case of the spin density, the local source effect is strictly related to difference

between the concentration/dilution degree of the Laplacian of the spin-resolved densities $\rho_\alpha(\boldsymbol{r'})$ and $\rho_\beta(\boldsymbol{r'})$.

The space partitioning applied in equation 1.29 is the same as equation 1.26. Also in the case of the SD, the partition of the space is done using the QTAIM zero-flux surface condition expressed in equation 1.10. This allows to keep the rigorous association of source contributions with the quantum atoms or group of atoms definition.

The substitution of $\nabla^2\rho(\boldsymbol{r'})$ by $\nabla^2 s(\boldsymbol{r'})$ has as a consequence a different pattern of atomic sources.[66] The two Laplacian functions have noticeably different portraits and the two scalar fields concentrate ($\nabla^2 f < 0$) and dilute ($\nabla^2 f > 0$) in two completely different ways, implying that ED and SD information are transmitted with different mechanisms.[66,67] Especially the spin density, which is defined as the difference between the α and β components of the electron density, has a complicate behaviour. Concentration or dilution of both α and β density components is not a sufficient condition to define the positive or negative source contributions. What is more important is the relative concentration/dilution of one density component compared to the other.[66]

Points where $LS_s(\boldsymbol{r}, \boldsymbol{r'})$ is positive will preferentially increase the α component of the $\rho(\boldsymbol{r})$ (and then the $s(\boldsymbol{r})$), while points where the local source contribution is negative will preferentially increase the $\rho_\beta(\boldsymbol{r'})$ component and then they will decrease the $s(\boldsymbol{r})$ value. Through the whole text and accordingly with the definition given in Reference [66], the former effect is called α *effect* ($s(\boldsymbol{r})$ is increased), while the second one β *effect* ($s(\boldsymbol{r})$ is decreased).

Table 1.4 summarises the relation between the $\nabla^2\rho_\alpha(\boldsymbol{r'})$ and $\nabla^2\rho_\beta(\boldsymbol{r'})$ concentration/depletion and the effect on the reconstruction of the SD in $\boldsymbol{r}$.[66] For a comparison, a similar table for the electron density is reported above (see Table 1.3).

Table 1.4: Effect of $\nabla^2\rho_\alpha(\boldsymbol{r'})$ and $\nabla^2\rho_\beta(\boldsymbol{r'})$ on the reconstruction of the electron spin density $s(\boldsymbol{r})$. α effect means that the SD is increased, while β means the opposite.

$\nabla^2\rho_\alpha(\boldsymbol{r'})$	$\nabla^2\rho_\beta(\boldsymbol{r'})$	Relative Magnitude	$\nabla^2 s(\boldsymbol{r'})$	$LS_s(\boldsymbol{r}, \boldsymbol{r'})$	Effect on $s(\boldsymbol{r})$
+	+	$\nabla^2\rho_\alpha(\boldsymbol{r'})>\nabla^2\rho_\beta(\boldsymbol{r'})$	+	-	β
+	+	$\nabla^2\rho_\alpha(\boldsymbol{r'})<\nabla^2\rho_\beta(\boldsymbol{r'})$	-	+	α
+	-	Any	+	-	β
-	+	Any	-	+	α
-	-	$\nabla^2\rho_\alpha(\boldsymbol{r'})>\nabla^2\rho_\beta(\boldsymbol{r'})$	-	+	α
-	-	$\nabla^2\rho_\alpha(\boldsymbol{r'})<\nabla^2\rho_\beta(\boldsymbol{r'})$	+	-	β

1.8.1 Magnetic/Relaxation Decomposition

Compared with the source function applied to the electron density case, interpretation of the source function spin density S_s results is not straightforward. To ease the chemical and physical interpretation of these results, usually it is helpful to split the source function contribution into a *magnetic* ($S_{s,mag}$) and *relaxation* (or *reaction*, $S_{s,relax}$) term.[66] The former contribution is defined as due to the unpaired α electrons distribution, while the

latter is related to the distribution of the remaining α and β electrons.[66,68,69]

$$S_s(\boldsymbol{r},\Omega) = S_{s,mag}(\boldsymbol{r},\Omega) + S_{s,relax}(\boldsymbol{r},\Omega) \tag{1.32}$$

By definition, the *magnetic* density integrates in the whole space to the number of unpaired electrons and it is positively-defined everywhere. On the contrary, the *relaxation* contribution integrates to zero in the whole space, but locally can assume positive or negative values. This statement is not true for restricted open approach, where the formally paired electrons are forced to share the same orbital. In general, the strict relationship $\rho_\alpha(\boldsymbol{r}) \equiv \rho_\beta(\boldsymbol{r}), \forall \boldsymbol{r} \in R^3$ is relaxed because of the presence of the unpaired electrons and then different values for $\rho_\alpha(\boldsymbol{r})$ and $\rho_\beta(\boldsymbol{r})$ are possible. The *magnetic* density is easily obtainable by diagonalizing the first-order density matrix and selecting only the natural orbitals (NOs) with occupation equal or close to one.

It is worth noting that, although the magnetic density is positive-defined everywhere, the $LS_{s,mag}$ and consequently the $S_{s,mag}(\boldsymbol{r},\Omega)$ values can assume both positive (α effect) and negative (β effect) values. What has to be positive or at limit equal to zero is the sum of all the $S_{s,mag}$ contributions. $S_{s,mag}$ and $S_{s,relax}$ can cooperate with or counteract each other in the definition of the total S_s contribution.

1.9 Representation of the Source Function Descriptor

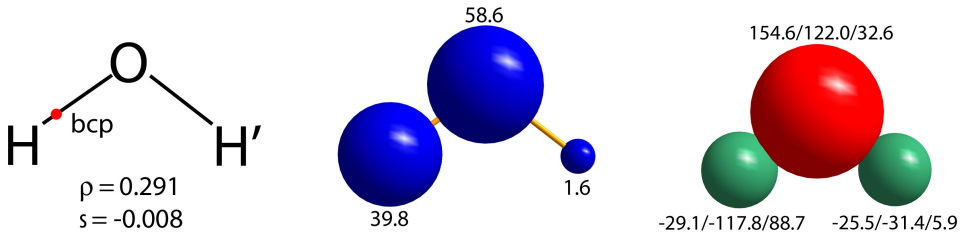
1.9.1 Source Function Percentage

The SF values obtained through equations 1.27 and 1.30 are expressed as absolute values in term of *atomic units* (*a.u.*, $e/bohr^3$). To have a clearer picture on the capability of an atom or a group of atoms Ω to reconstruct the density in a system, it is usually better to scale the absolute source function value with respect to the total reconstructed density in the selected reference point, representing it as a source function percentage (SF%) contribution:[21,63,70]

$$S\%(\boldsymbol{r},\Omega) = \frac{S(\boldsymbol{r},\Omega)}{\rho(\boldsymbol{r})} \cdot 100; \qquad S_s\%(\boldsymbol{r},\Omega) = \frac{S_s(\boldsymbol{r},\Omega)}{s(\boldsymbol{r})} \cdot 100 \tag{1.33}$$

The atomic percentage contributions are shown in a ball-and-stick representation for all the atoms in a system. The volume of the sphere representing the atom is proportional to the value of its source function percentage. The colour of the sphere refers to the sign of the contribution: blue and yellow refer to the electron density and they define positive (source) and negative (sink) contributions, respectively; green and red refer instead to the spin density and they correspond to α and β effects, respectively.

An example of a typical SF percentage analysis is depicted in Figure 1.1, where the atomic source functions are evaluated at the bond critical point of the O-H bond for a triplet water molecule. The selected case was studied at the CASSCF(8,8)/6-311++G(2d,2p) level of theory.



(a) BCP reference point (b) $S\%(BCP, \Omega)$ (c) $S_s\%(BCP, \Omega)$

Figure 1.1: source function percentage contributions to the reconstruction of the electron and spin density in the O-H bond critical point of the triplet state water. (a) position of the reference point in the molecule (red dot); (b) source (blue) and sink (yellow, not shown) atomic effect for the reconstruction of the electron density $\rho(\boldsymbol{r})$ at the BCP. Numerical values indicate the percentage atomic contributions; (c) α (green) and β (red) effect for the reconstruction of the spin density $s(\boldsymbol{r})$ at the BCP. Percentage values indicate the total/magnetic/relaxation atomic contributions. The volumes of the spheres are proportional to the total atomic contributions. Full results on this study case are reported in Reference [66].

The figure shows two completely different source scenarios if one reconstruct the $\rho(\boldsymbol{r})$ or the $s(\boldsymbol{r})$ in the same RP.

The $S\%(\boldsymbol{r})$ picture is coherent with the classical covalent bond description, where the two atoms involved in the interaction account for the 98.4 % of the total ED at the BCP, while the distant H' atomic contribution is negligible.[66] On the other hand, the information given by the source function for the reconstruction of the spin density in the same reference point, $S_s\%(\boldsymbol{r})$, shows a complete different co-operation between the atoms. The O atom yield a β effect as high as ≈156%. The positive value associate with the β effect is related to the fact that the spin density evaluated in the BCP is negative (see Equation 1.33). On the contrary, the two H atoms oppose to the oxygen atom yielding α effects equal to -29.1% and -25.5% for H and H', respectively. It is worth noting that in the reconstruction of the spin density in the BCP, both H atoms equally contribute. The same evidence is not found in the reconstruction of the electron density, where there is a clear difference in the source contributions. However, the decomposition of the total SF% contribution into the magnetic and relaxation terms highlights the different nature of these similar effects. H atom shows large magnetic (-117.8%) and relaxation (-88.7%) terms while the same values for H' are -31.4% and 5.9%, respectively. For both atoms the two contributions counteract each other, but for the former the absolute values are definitely larger than for the latter. This fact can be explained considering that the BCP reference point is closer to H than H' and then the scaling factor $|\boldsymbol{r} - \boldsymbol{r'}|$ (see Equation 1.30) is less effective.

This analysis expectedly confirms that spin transmission occurs through a different channel with respect to the electron density information.

The choice of the reference point $\boldsymbol{r}$ depends on the studied property. The BCP is usually taken as reference when the electron density reconstruction is studied to discuss

about chemical bonding.[21,63,70] Other choices, such as (3,-3) and (3,+3) CPs of the Laplacian distribution, can be considered and they sometimes result more useful than the BCP itself.[21,65,68,69]

This kind of analysis has large potentiality to highlight subtle atomic features in the reconstruction of the density, but it also shows some limitations. The most problematic limitation is related to the choice of the reference point, which is sometimes not so obvious. While it has been well established that the BCP is a good (but not exhaustive) point for the analysis of the chemical bonding, this is not the case for the study of other phenomena or other scalar fields, such as the spin density. For instance, it has been found that the SD at the BCP is usually close to zero because in that region the electron pairing is maximised (particularly in the case of covalent bonds).[66–69]

That said, which criteria one has to take into account in the choice of the reference point? And can one be sure that the chosen point is representative of the studied molecular region?[68,69]

1.9.2 n-Dimensional Source Function Partial Reconstruction

To overcome the issue related to the choice of the reference points in not-trivial situations, a new representation of the SF descriptor has been developed as a part of this three year PhD work and it has been reported for the first time in Reference [68].

Instead of choosing a single reference point, the SF contributions for a given subset of atoms are calculated at an N-dimensional grid of points (N = 0-3). In this way, one can visualize in a more clear and immediate way the contribution from different sources along a line (N = 1), on a surface (N = 2) or in a volume (N=3). The case N = 0 coincides with the SF classical evaluation in a single reference point.

Partial reconstructed densities, $\rho\{\Omega, subset\}$ and $s\{\Omega, subset\}$, can be expressed as:

$$\rho\{\Omega, subset\}(\boldsymbol{r}) = \sum_{\Omega=1}^{subset} S(\boldsymbol{r}, \Omega) = \sum_{\Omega=1}^{subset} \int_{\Omega} LS(\boldsymbol{r}, \boldsymbol{r'}) d\boldsymbol{r'} \tag{1.34}$$

$$s\{\Omega, subset\}(\boldsymbol{r}) = \sum_{\Omega=1}^{subset} S_s(\boldsymbol{r}, \Omega) = \sum_{\Omega=1}^{subset} \int_{\Omega} LS_s(\boldsymbol{r}, \boldsymbol{r'}) d\boldsymbol{r'} \tag{1.35}$$

where $\boldsymbol{r}$ is a point belonging to the ensemble $R^N \subseteq R^3$ and $\Omega, subset$ refers to the atoms used in the *partial* reconstruction.

When all the atoms are included within equations 1.34 and 1.35, the reconstructed densities are no longer partial and they are formally equal to the real ones.[68] The total reconstruction is useful to check the quality of the numerical integration adopted but it does not show any chemical relevance.

When only the SF contributions of a subset of atoms are included in Equations 1.34 and 1.35, the role these atoms have in the reconstruction of the analysed density in the selected region becomes manifest. These maps depict if in a molecular region the atomic

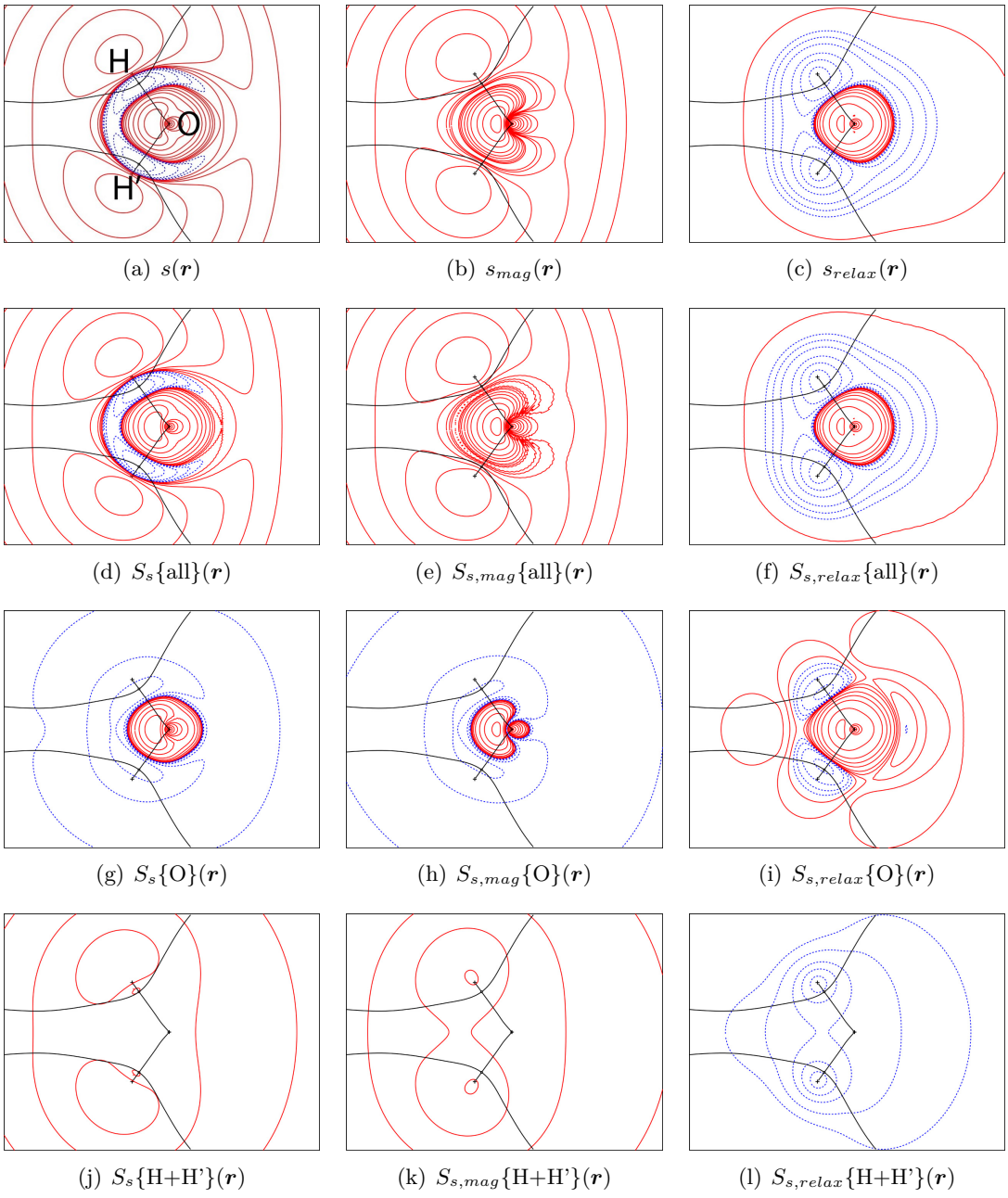
contributions are important, modest or negligible. They could also be used for a proper selection of reference points for the source function percentage analyses.

Figure 1.2 shows an example of the source function partial reconstruction analysis applied to the study of the spin density. The panels illustrate the spin density distribution and the spin density two-dimensional reconstructed maps using different subsets of atoms for the triplet state of the water molecule.[68]

In the first row, the spin density evaluated in the molecular plane is shown together with its magnetic and relaxation contributions obtained as explained in Section 1.8.1. The total spin density (Figure 1.2(a)) shows positive regions around the O and H atoms. In the interatomic region, a negative half-crown is present, indicating a possible polarization mechanism along the O-H bonds. The decomposition of this density into magnetic and relaxation contributions helps to understand the nature of these features. The magnetic term is the dominant one and it is able to reconstruct the major part of the total density, except for the negative region. On the other hand, the relaxation contribution clearly presents a negative spin density region around the almost spherical positive area centred on the O nucleus. The relaxation term is able to counteract the magnetic contribution in the interatomic region, producing the negative area in the total map, while it cooperates to construct the spin density on the O atom.

The second row depicts partial reconstructed maps obtained taking into account all the atomic source function contributions ($S_s\{\text{all}\}(\boldsymbol{r})$). As it is clearly evident, the reconstructed maps are almost undistinguishable from the real ones (first row) except for very small features in the low density regions, as it is evident in the horn-like parts around the O atom in the magnetic map or in the contraction of the most external isovalue in the relaxation term. These small differences can be imputed to numerical approximations due to the integration required to evaluate the source function contributions (see Equation 1.30). However, these small errors are located in poorly interesting regions of the maps and they can be considered negligible for our purposes. The fact that the reconstructed density and its components are barely distinguishable from the primary ones allow us to trust the reconstruction obtained through the source function descriptor and then to consider valid all the relate discussion.

The third and fourth rows show the source function partial spin density maps reconstructed using as subset the oxygen atom, $S_s\{\text{O}\}(\boldsymbol{r})$, and the two Hydrogens, $S_s\{\text{H+H'}\}(\boldsymbol{r})$. This atomic-level decomposition allows one to intimately understand the role played by atoms in the reconstruction of the density in the selected plane. The positive area around the O atom in the total map is mainly due, from a qualitative point of view, to the $S_s\{O\}$ contribution. The O atom yields an α effect in the region around its nucleus which has the same shape to that observed into the spin density map (Figure 1.2(a)) but more contracted. Around this positive core, the oxygen atom generates a wide β effect responsible for the formation of the negative region in Figure 1.2(a). In fact, looking at the contribution from the two hydrogen atoms (Figure 1.2(j)), they produce a broad α effect in the whole space and they can not be responsible for the negative spin density. However, their effect is



(a) $s(\boldsymbol{r})$ (b) $s_{mag}(\boldsymbol{r})$ (c) $s_{relax}(\boldsymbol{r})$

(d) $S_s\{\text{all}\}(\boldsymbol{r})$ (e) $S_{s,mag}\{\text{all}\}(\boldsymbol{r})$ (f) $S_{s,relax}\{\text{all}\}(\boldsymbol{r})$

(g) $S_s\{\text{O}\}(\boldsymbol{r})$ (h) $S_{s,mag}\{\text{O}\}(\boldsymbol{r})$ (i) $S_{s,relax}\{\text{O}\}(\boldsymbol{r})$

(j) $S_s\{\text{H+H'}\}(\boldsymbol{r})$ (k) $S_{s,mag}\{\text{H+H'}\}(\boldsymbol{r})$ (l) $S_{s,relax}\{\text{H+H'}\}(\boldsymbol{r})$

Figure 1.2: spin density (s) and two-dimensional source function partial spin density maps (S_s) in the molecular plane for the triplet H_2O at the CASSCF(8,8)/6-311++G(2d,2p) level of theory (first column) and their magnetic (second column) and relaxation (third column) terms. Solid red lines indicate positive value (α effect) of spin density, while dotted blu denote the negative values (β effect). Contours are drawn at interval of $\pm(2,4,8) \cdot 10^{-n}$, $-3 \leq n \leq 0$ *a.u.*. Full black lines mark boundaries of atomic basins.

evident in the definition of the spin density profile on themselves and of the enlargement of the positive density around the O nucleus (compare Figures 1.2(g) and 1.2(j)) with 1.2(d)). Both magnetic and relaxation terms of the oxygen atom subset co-operate in the definition of the positive density of the total map while the negative spin density region is mainly dominated by the effect of the latter contribution. The magnetic and relaxation sources from the hydrogen atoms behave in an opposite way, with the former yielding α effect and the latter producing a counteracting β effect in the whole space. It is worth noting once more here that although the magnetic spin density is everywhere positively defined (Figure 1.2(b)), the single atomic magnetic source function can be negative as in the case of the $S_{s,mag}\{\mathrm{O}\}(\boldsymbol{r})$. As shown in Figure 1.2(k) the contribution from H and H' is positive and overcomes the negative part of the $S_{s,mag}\{\mathrm{O}\}(\boldsymbol{r})$ all over the space.

Concluding this section, it is important to underline the profound difference between the partial density reconstruction given by the source function descriptor and the one obtainable using other approaches, like that using subsets of pseudoatom densities in the multipole models[15] or that using the contribution from the basis functions of a subset of atoms in a theoretical calculation.[68] The information the SF partial reconstructions provide is only a function of the total density and depicts the atomic contributions within a *cause-effect* picture. The SF partial reconstructed densities are then independent from any model. On the contrary, the standard partial reconstructed densities depend on the multipole model or atomic basis set used and are then biased by this choice. For absurd, a change in the origin of a complete atomic basis set will not change the total density of the system but it will change the atomic standard partially reconstructed densities.[68]

In the following chapters, the combination of these two source function analyses is discussed both for the spin density and the electron density distributions. In particular, it is worth noting that the former scalar field has not been studied in detail so far and the here presented works have to be considered as a starting point for a wider and deeper analysis of the spin density distribution.

1.10 Full Population Matrix through Source Function Integration

As already shown in the previous sections, deep chemical insights can be extracted from the study of the electron and spin density through the use of the source function descriptor. Both the classical percentage contributions in a selected reference point and the novel multi-dimensional representation give information on how the basins in the system contribute to the reconstruction of the selected density, highlighting which atoms or group of atoms play a major role.

The SF tool may also be used to build a *full population matrix*, *i.e.* to define to what extent one basin contributes to determine the electron population of the other basins and its own population.[63,71]

By integrating the source contribution from all the basins Ω' over the basin Ω, Equation 1.26 becomes:

$$N(\Omega) = \int_{\Omega} \rho(\boldsymbol{r})\ d\boldsymbol{r} = \sum_{\Omega'} \int_{\Omega} S(\boldsymbol{r}, \Omega') d\boldsymbol{r} = \sum_{\Omega'} \int_{\Omega} \int_{\Omega'} -\frac{1}{4\pi} \frac{\nabla^2 \rho(\boldsymbol{r'})}{|\boldsymbol{r} - \boldsymbol{r'}|} d\boldsymbol{r'} d\boldsymbol{r} \tag{1.36}$$

where $N(\Omega)$ is the total population of the basin Ω. It has been said in Equation 1.28 that the source function contribution can be divided into internal and external terms. The same concept can be applied here, decomposing $N(\Omega)$ into an *inner* self-population (N_{in}) and an *outer* one (N_{out}).

$$N(\Omega) = N_{in}(\Omega) + N_{out}(\Omega) = M(\Omega, \Omega) + \sum_{\Omega'} M(\Omega, \Omega') \tag{1.37}$$

N_{in} indicates the total population that the basin Ω gives to itself, while N_{out} is the sum of all the contributions that the other basins Ω' give to Ω.

The total contribution of a basin Ω_Y to another basin Ω_X, that is the element (Ω_X, Ω_Y) in the population matrix, is indicated as $M(\Omega_X, \Omega_Y)$, where $\Omega_X = \Omega_Y$ for the intra-atomic contributions and $\Omega_X \neq \Omega_Y$ for the inter-atomic ones.

The *full population matrix* assumes the form:

Table 1.5: scheme of a generic full population matrix.

Ω_X	Ω_1	Ω_2	$\cdots$	Ω_N	$N(\Omega_X)$
Ω_1	$M(\Omega_1, \Omega_1)$	$M(\Omega_1, \Omega_2)$	$\cdots$	$M(\Omega_1, \Omega_N)$	$N(\Omega_1)$
Ω_2	$M(\Omega_2, \Omega_1)$	$M(\Omega_2, \Omega_2)$	$\cdots$	$M(\Omega_2, \Omega_N)$	$N(\Omega_2)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
Ω_N	$M(\Omega_N, \Omega_1)$	$M(\Omega_N, \Omega_2)$	$\cdots$	$M(\Omega_N, \Omega_N)$	$N(\Omega_N)$
$M(\Omega_X)$	$M(\Omega_1)$	$M(\Omega_2)$	$\cdots$	$M(\Omega_N)$	$N(R^3)$

The elements in the last row and in the rightmost column of the Table 1.5 report the sum of the full population matrix elements along the corresponding column and along the corresponding row, respectively.

$M(\Omega_X)$ indicates the number of electron "distributed" to the other basins in the system by the basin Ω_X (included itself), $N(\Omega_X)$ is the total electronic population of the basin as already stated before and $N(R^3)$ is the total number of electrons in the molecule.

The diagonal elements are the *self-* or *inner contributions*, while the *off-diagonal* terms are the interatomic contributions from one basin to the population of the other. One peculiarity of the off-diagonal elements of the matrix is that usually $M(\Omega_X, \Omega_Y) \neq M(\Omega_Y, \Omega_X)$ if no molecular symmetry constraints are present. This implies that the source function can discriminate the different behaviour of two atoms in term of their reciprocal contributions. This feature is quite important for the preservation of the chemical difference between different elements. Many population matrices, such as the Mulliken's[72], do not preserve this asymmetry and any two off-diagonal elements are always equal by construction.

The matrix reported above defines a population analysis based only on the electron density Laplacian $\nabla^2\rho$ and it is not based on any model or discrete representation of the electron density or of the one-electron density matrix. As a consequence, this population analysis can in principle be applied to experimental electron density distributions. This fact represents a unique chance to compare theoretical and experimental results on the same ground and on an unbiased basis. It is worth noting that there is a profound difference between the elements of the matrix obtained through the use of the SF descriptor together with the QTAIM space partitioning and the elements of such matrix obtained in term of a sum of contributions given by products of atomic basis set functions overlaps weighted by one-electron density matrix elements.[63] The use of the source function for determining such elements largely increases the chemical meaning of the population matrix. While in Mulliken's analysis the interatomic terms are related to the overlap between the set of orbitals centred on the atomic nuclei, the SF sees these contributions within a *cause-effect* relationship: the cause is the Laplacian distribution within a basin, while the effect is the reconstructed density inside another (or the same) basin. In other words, the elements of the population matrix obtained using the source function tool can be read exactly as the number of electrons determined by basin Ω_X into basin Ω_Y. This fact is particularly interesting since it is almost unique.

For instance the well-known localization/delocalization indices (LI/DI, respectively) enable to obtain localization/delocalization matrices that partition the electrons in a system between those that are fully localized on the atoms of the system and those that are instead equally shared between the various pair of atoms in the system.[73–77]. This partition also yields a population matrix,[76,77] but the meaning of its elements is clearly quite different from that inherent to the SF reconstruction. Besides, the LI/DI population matrix is symmetric by definition.

Chapter 2

Spin Density Investigation through the Source Function Descriptor[1]

Electron distributions, such as the electron position density (ED), $\rho(\boldsymbol{r})$, and the electron spin density (SD), $s(\boldsymbol{r})$, are quantum mechanical observables and scattering techniques enable their reconstruction in crystals.[78] The electron density can be retrieved through X-ray and electron diffraction of crystalline materials, while the spin density can be derived from polarized neutron diffraction (PND) experiments on magnetic crystals.[58]

In the last half-century the charge density field has strongly developed thanks to a large community of practitioners.[4,12,78–80] New methods to derive accurate X-ray electron densities and the development of several topological descriptors have allowed to fully characterise and study the electron distribution $\rho(\boldsymbol{r})$ and its properties.

If we can consider the study of the ED as a mature (but not fully saturated) field, this is not true when we consider other densities such as the momentum and the spin densities. From an experimental point of view, the sophisticated set-up, the very small number of reflections obtainable from PND experiment ($\approx 10\%$ of the reflections in the reciprocal space can be collected) and the restricted access to neutron beamtime are the bottle-neck and the major issues in the reconstruction of good experimental spin density.[58,81,82] However, the study of the latter density distribution is fundamental to get insight on the molecular details of spin information transmission as a way to design specific magnetic networks and to tune their properties.[68,69,82–84] A large scope for important developments appears to be still at hand in these fields and some steps toward this direction have been moved in the last years. In a 2015 paper, Macchi *et al.* wrote about the possibility to "extract more information from experimental data, which necessarily means challenging well established models and testing extensions, corrections or even alternative routes".[85] Gillet *et al.* have recently achieved a relevant intermediate step in the long-term quest

[1] This chapter fully reports the results already published in References [68] and [69]

for a simultaneous refinement of charge, spin and momentum densities.[59–61] Deusch *et al.* successfully introduced for the first time a spin-split (ρ_α and ρ_β) version of the Hansen and Coppens multipolar model[14] for a joined X-ray and PND data fitting.[82,84] Spin-up and spin-down distributions were found to be in quite good agreement with those obtained from theoretical computations.[82,84] However, this agreement is not a sufficient condition to prove the quality of the spin density distribution, as it is shown in the next sections.[69]

Theoretical SDs can be easily calculated using, for instance, unrestricted methods but the results are usually far from being reliable. Systematic studies using density functional theory (DFT) exchange-correlation (XC) functionals, which are largely used in literature, have demonstrated to be unable to treat open-shell systems properly and to retrieve accurate spin density distributions.[55,57,68,69,86] *Ab-initio* post-HF methods are needed to predict more reliable spin densities since they are able to properly treat the correlation among electrons. However, these methods are sometimes too computational demanding and cannot be used for large systems.

In future, this spin-split model expansion should lead to increasingly accurate spin density distributions in crystal, which represent the most important and valuable quantity to visualize the magnetic interactions in complex solid-state networks.[82–84] Deciphering the mechanisms through which spin information propagates from paramagnetic to non-magnetic centres and the reasons for possible spin polarization effects requires interpretative models which are inherently unavailable from experiment and are somehow arbitrary. A more rigorous alternative to these models is to resort to descriptors directly based on observables, that can be applied both on theoretical and experimental spin densities.

As already shown in the Chapter 1, the source function descriptor (SF) has been extended to the study of the spin density scalar field. The spin density source function provides quantitative insights into the relative capability of different atoms or group of atoms to reconstruct the spin density in specific reference points (RPs) within a *cause-effect* relationship. Differently from other methods, the source function is able not only to show how spin information propagates from paramagnetic to non-magnetic centres, but also *vice versa* how the latter may in turn influence the spin density distribution of the paramagnetic centres themselves. Moreover, the high sensitivity of the SF descriptor allows to bring to light small differences and to discuss the accuracy of theoretical or experimental SD distributions, or to compare SD obtained from different levels of theory.[68,69]

In this chapter, the results of the analysis of the theoretical spin density distribution of two azido dinuclear Cu(II) complexes, an End-On (EO)[87] and an End-to-End (EE) Cu(II) azide[88], are shown.

This work has several main purposes. Firstly, the application of the SF to the study of the spin density has to be tested on more complex systems than those carried out thus far.[66,67] Secondly, the kind of information that may be retrieved from a spin density distribution using a descriptor that potentially also applies to experimentally derived spin densities requires further exploration. Thirdly, the SF analysis distribution of these two azide complexes should help in understanding why they have different magnetic behaviour.

In both systems the Cu centres are ferromagnetically coupled, but with quite different coupling constant J, large for the EO and almost negligible for the EE system.[87,88] As a further purpose, the source function is used here to highlight the origins of the differences among the spin densities retrieved using DFT functionals or complete active space self consistent field (CASSCF) methods.[55,68,69]

2.1 Structural and magneto-structural information

Azido-bridged compounds have been largely synthesised and studied since 1980s for their interesting magneto-structural properties.[88] The azido group N_3^- is one of the most used magnetic coupler and it is classified as a non-innocent ligand since it influences the oxidation state of the coordinated metal(s).[87] The N_3^- can couple metallic centres in two different ways: (i) End-On coordination (μ-1,1 EO), where only one terminal N atom bonds the two metallic centres; (ii) End-to-End coordination (μ-1,3 EE) where both the two terminal N atoms coordinate the metals.

In both cases, the coordination can be symmetric or asymmetric. In the former case, the two M-N bonds are equivalent, while in the latter case they have different lengths.

The geometries of the two studied dinuclear Cu azides were taken from their molecular crystal structures, using the 18 K (EO, Reference [87]) and 150 K (EE, Reference [88]) X-ray structures. The three-dimensional arrangement of the atoms, together with the most important atomic labels and bond lengths are shown in Figure 2.1.

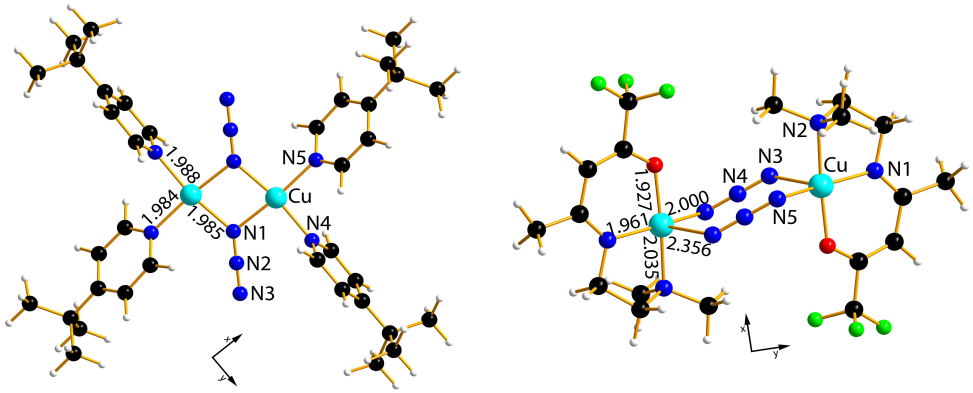


Figure 2.1: Molecular structure of the End-On (left) and End-to-End (right) centrosymmetric azido-bridged dicopper molecular complexes. Bond lengths are expressed in Å. Atoms color code: Cu (light blue), F (light green), O (red), N (blue), C (black), H (white).

A summary of the main structural and magnetic features is reported in the following list:

- Symmetric End-On (EO) complex[87]: $[Cu_2(t\text{–Bupy})_4(N_3)_2](ClO_4)_2$, where t-Bupy = *p-tert*-butylpyridine. The complex is centrosymmetric with Cu-Cu distance ≈ 3 Å and two short equivalent $Cu–N_{azide}$ bonds (d = 1.985 Å). The coordination of the

Cu centres is square-planar, with the two *p-tert*-butylpyridine ligands coordinating through the nitrogen atoms. The EO complex exhibits a triplet ground state, with a ferromagnetic coupling constant J = 300 cm^{-1}.

- Asymmetric End-to-End (EE) complex[88]: $Cu_2L_2(N_3)_2$, where L = 7-dimethylamino-1,1,1-trifluoro-4-methyl-5-azahept-3-en-2-onato. The complex is centrosymmetric with Cu-Cu distance ≈ 5 Å and one long and one short $Cu-N_{azide}$ bonds (d = 2.356 Å (Cu-N5) and 2.000 Å (Cu-N3), respectively). The coordination of the Cu centres is square pyramidal. The terminal azide N5 atom together with N1, N2 and O1 of the L ligand form the square pyramidal base, while the other terminal N of the other azido group, N3, coordinates from the top. The EE complex exhibits a very weak ferromagnetic behaviour (J = 17 cm^{-1}).

2.2 Computational Details

Theoretical spin density evaluations for the triplet ground states of the End-On and End-to-End azido dicopper complexes were performed *in vacuo* using the Gaussian09/16 program packages.[89,90] For the EO system, both the double positively charged (without the chlorate anions) and the neutral complex were calculated, but results are discussed only for the former as both of them provide almost identical spin density distributions.

Unrestricted Hartree-Fock (UHF or HF), unrestricted density functional theory (UDFT or DFT) and complete active space self consistent field (CASSCF) levels of theory were employed for both compounds, but not all the methods are discussed here. Table 2.1 summarises which methods have been used and here reported for both azido complexes.

Table 2.1: Summary of the theoretical methods employed. The meaning of the signs is: calculated/shown, ✓= yes, X = no

Level of Theory	EO complex	EE complex
HF	✓/X	✓/✓
DFT - B3LYP	✓/✓	✓/✓
DFT - BLYP	X/X	✓/✓
CASSCF(6,6)[a]	✓/✓	✓/✓
CASSCF(10,10)[a]	✓/X	✓/✓

[a] CASSCF(Y,Z): number of electrons (Y) and orbitals (Z) in the active space.

Triple-ζ pob-TZVP basis set was used for all the methods.[91] From now on, the letter *U* is be omitted from every method name or acronym for simplicity.

All the reported DFT and HF results refer to spin-contamination annihilated wavefunctions, where the correction was necessary to avoid contamination by higher multiplicity states. The only case where spin-annihilation procedure did not work properly was the EO Hartree-Fock calculation (not reported here - see Table 2.1). The hybrid B3LYP and

pure BLYP DFT functionals for the EE complex were chosen analogously to those used by Boguslawski *et al.* on iron nitrosyl to compare the spin density distribution retrieved by hybrid and pure exchange-correlation Hamiltonians.[55]

Complete active space calculations for both azido complexes were performed using 6 electrons in the active space composed by 6 orbitals (hereinafter CASSCF(6,6)), 2 doubly occupied, 2 half-occupied and 2 virtual orbitals. A larger active space with 10 electrons in 10 orbitals (hereinafter CASSCF(10,10)) was considered for the EE system (4 doubly occupied, 2 half-occupied and 4 virtual orbitals). We were not able to achieve this latter active space in the EO Cu azide because of symmetry issues. To take into account both molecular and orbital symmetries, we should have performed a CASSCF(14,14) calculation, which was unfeasible to reach because of intrinsic technical limitations of computing machines and of Gaussian09/16 software[89,90]. The starting orbital guesses for all the CASSCF computations were taken from B3LYP (EO) and HF (EE) natural orbitals (NOs). For the EO system, the 6 orbitals were chosen in an analogous manner to those adopted in the model by Aebersold *et al.*, selecting both A_g *gerade* and A_u *ungerade* orbital symmetries.[87] For the EE system, two half-full NOs mainly localized on the Cu(II) cations with a *d-like* shape and some tails on the coordinated ligand heteroatoms plus four/eight π/π^*-like NOs located on the backbone of the ligand L were chosen as active orbitals in CASSCF(6,6)/CASSCF(10,10) calculations, respectively. Like for the EO system, both A_g and A_u orbital symmetries were taken into account. The number of active configurations for the CASSCF were 225 and 44100 for the (6,6) and (10,10) active spaces, respectively.

The spin densities for HF, DFT and CASSCF wavefunctions were calculated using the option "pop = NO" in Gaussian09/16 input and selecting only those NOs with occupation equal or very close to one.[66,68,69,89,90] In all the computations performed, no relevant deviations from the unity were detected (maximum shift = 0.00029 electrons).

Topological analyses of the ρ, s and $\nabla^2\rho$ fields were performed through a modified version of the AIMPAC program package by Biegler-König *et al.*[92–95] Other not-published home-developed codes were used to evaluate other properties. EXTREMESPIN code was used to evaluate spin density, α and β components of the electron density and the spin density Laplacian at given positions. SPINSF2016 code was used to evaluate the atomic source function contributions to the electron and spin densities reconstruction at selected reference points or on a N-dimensional grid of points. This latter option was developed and tested during these three years of PhD and it is described in Section 1.9 of Chapter 1. PLOTDEN2016 was finally used to plot contour two-dimensional maps of the studied fields and the SF reconstructed partial spin density maps, together with their magnetic and relaxation density counterparts. SPINSF2016 and EXTREMESPIN are heavily modified versions of the original PROAIMV end EXTREME codes of the AIMPAC program package,[92–95] while PLOTDEN2016 is an updated version of the PLOTDEN2013 code (also unpublished but with a brief description in the supporting information (SI) of Reference [66]). All the structures reported in this chapter were drawn using the DIAMOND program[96], while the 3D isosurfaces were plotted using VESTA code[97].

2.3 End-On Cu Azide Complex

2.3.1 Population Analysis

Quantum theory of atom in molecules (QTAIM) and Mulliken's theoretical atomic charges q(Ω) and spin population SP(Ω) for selected atoms are shown in Table 2.2, together with PND spin population estimates $SP_{PND}(\Omega)$ and Mulliken's charges evaluated by Aebersold *et al.*.[87]

Table 2.2: Atomic charges (q) and spin population (SP) for selected atoms or group of atoms Ω in the End-On azido Cu complex. q(Ω) and SP(Ω): QTAIM partitioning; $SP_{Mull}(\Omega)$ and $SP_{Mull,Aeb}(\Omega)$: Mulliken's partitioning; $SP_{PND}(\Omega)$: polarized neutron diffraction estimate.

Ω	q(Ω)[a]	SP(Ω)[a]	$SP_{Mull}(\Omega)$[a]	$SP_{PND}(\Omega)$[b]	$SP_{Mull,Aeb}(\Omega)$[b]
Cu	1.433/1.095	0.899/0.619	0.919/0.628	0.783(7)	0.425
N1	-0.642/-0.488	0.032/0.095	0.024/0.089	0.069(6)	0.167
N2	-0.216/-0.138	-0.009/0.005	-0.013/-0.007	-0.016(6)	-0.005
N3	0.126/0.039	0.028/0.088	0.030/0.100	0.057(7)	0.122
Azide	-0.733/-0.587	0.051/0.188	0.041/0.182	0.110(19)	0.284
N4	-1.464/-1.056	0.022/0.087	0.018/0.088	0.067(8)	0.129
N5	-1.496/-1.083	0.022/0.079	0.018/0.080	0.049(7)	0.120
$\sum C_{py}(N4)$[c]	1.224/0.877	0.002/0.008	0.002/0.004	-0.037(46)	0.009
$\sum C_{py}(N5)$[c]	1.264/0.914	0.002/0.008	0.002/0.004	0.005(52)	0.008

[a] CASSCF(6,6)/B3LYP data.
[b] Data from Table 5 of reference [87]. $SP_{Mull,Aeb}(\Omega)$ results were obtained from the simplified ($[Cu_2(py)_4(N_3)_2]^{2+}$) system.
[c] Sum of the properties of the C atoms belonging to the pyridine rings having N4 or N5 has their N atom.

The theoretical data listed in Table 2.2 underline the strong sensitivity of the spin density and then spin populations on the level of computation used. CASSCF(6,6) predicts ≈ 1.80 electrons over 2 to be resident on the two Cu atoms, while the exchange-correlation DFT functional used predicts instead only ≈ 1.24 electrons. The Mulliken's spin population predicted for the two Cu atoms on the complex model by Aebersold *et al.* is further lower, being equal to 0.85 electrons. This discrepancy can be explained considering that (i) Aebersold *et al.*[87] used a model system instead of the whole Cu azide complex (the *tert-butyl* groups were substituted by H atoms); (ii) they used a TZVP-type basis set, different from ours even if of similar quality; (iii) they used a local Becke-Perdew (BP) exchange-correlation functional instead of an hybrid B3LYP. The Cu spin population estimated by the PND technique, 1.56 electrons for the two Cu atoms, is intermediate between the predicted CASSCF(6,6) and B3LYP values, but closer to the former. QTAIM and Mulliken's spin populations are close to each other when using the same level of theory. However, their small differences are capable of changing the sign of the populations from positive to negative or *vice versa* when the population is close to zero. Also the predicted atomic net charges are strongly different: CASSCF(6,6) method predicts a +1.433 *e* on the Cu atom, while the B3LYP functional finds only +1.092 *e*. This large variation may be in

some way related with the large difference found for the spin populations.

The CASSCF(6,6) method predicts that ≈ 5.1% of the spin population is delocalized on the N_3^- ligand and the remaining ≈ 4.4% is mainly localized onto the pyridinic nitrogen atoms. As can be noted from Table 2.2, the carbon atoms of the pyridines bear an almost negligible population. The same values evaluated by the DFT functional are quite different, being ≈ 18.8% and 8.4% (average value), respectively. The spin delocalization is about four times larger at the B3LYP level than at CASSCF one, both on the azido and on the pyridine ligands. The C atoms have a four time larger spin population in the case of the DFT method, although its value is still close to zero. This is a well known behaviour of the classical (both pure and hybrid) exchange-correlation functionals, which are not able to correctly reproduce the spin density features.[55] The reason of this is the absence of any spin-dependent term in the Hamiltonian of almost all the most used DFT functional. These latter, accordingly to the Hohenberg-Kohn theorem, are accurately tuned to reproduce the electron density of a system, not its spin density.[11] Instead, CASSCF wavefunctions with an adequate active space are definitely much closer to the spin density obtained using other more sophisticated methods, such as the density-matrix renormalization group (DMRG) method, which may be considered as an (almost) exact reference for the spin density.[57] Besides, the spin density distributions are known to also strongly depend on the exchange-correlation functional when a non-innocent ligand is present in transition metal complexes.[55,86] Despite these issues, we chose to analyse the B3LYP spin density for the sake of comparison with previous studies on the End-to-End complex[82,88] and between the two Cu azido complexes here studied.

Assuming, therefore, CASSCF(6,6) as the most accurate theoretical estimate of the electron spin density distribution here discussed for the EO complex, both B3LYP and the local functional adopted by Aebersold *et al.*[87] exaggerate the spin delocalization through the EO azido complex. The CASSCF and DFT methods, although quite different from a quantitative point of view, are qualitatively in agreement in predicting the relative weights of spin delocalization among the ligands. In particular, both approaches predict N2 to have the smallest spin population in the N_3^- group, although they disagree on the sign of this population (-0.009 *e* and +0.005 *e* for CASSCF(6,6) and B3LYP level of theory, respectively).

As already seen analysing the spin population on the Cu atom, CASSCF(6,6) outcomes are closer to the PND results but still far. Concerning this deviation, one should take into account the shortcomings of PND data collection and refinement. Aebersold *et al.* used only 152 magnetic structure factors to define the multipolar model, refining only the spherical part or one angular Slater atomic function on each atom.[87] Another parameter to consider for this discrepancy is the convergence of the CASSCF active space. Increasing the number of orbitals from 6 to 8 and then to 10, that is increasing the active configuration from 225 to 1930 and 9450, does not influence significantly either the energy and the spin population. Energy is lowered by 0.00012 and 0.00017 *a.u.* for CASSCF(6,8) and CASSCF(6,10), while the SD populations remain unaltered within 0.0005 *e*. On the

other hand, increasing the number of correlated electrons is a far more delicate issue. To take into account both the orbital and geometrical symmetries of the system, the number of electrons should be increased from 6 to 14 (two π electrons for each pyridine unit), leading to not easy feasible computations as already discussed above. CASSCF models with an intermediate number of active electrons, like 8 or 10, have nonetheless been tested. CASSCF(10,10) leads to small asymmetries in the population of the two symmetric parts of the complex. For instance, the QTAIM populations for the pyridine N4 and N5 atoms change from 0.022 e for both atoms to 0.045 and 0.021 e, respectively. Moreover, CASSCF(10,10) retrieves an enhanced spin population for the pyridine N atoms whose electrons became active and a decreased population on the Cu atoms (0.850 e). In general, increasing the number of active space seems to lead to a closer agreement among CASSCF and PND results. However, considering the non-physical bias introduced with an intermediate number of electrons and the non-feasible CASSCF(14,14) level of theory, in this work we only showed the CASSCF(6,6) results.

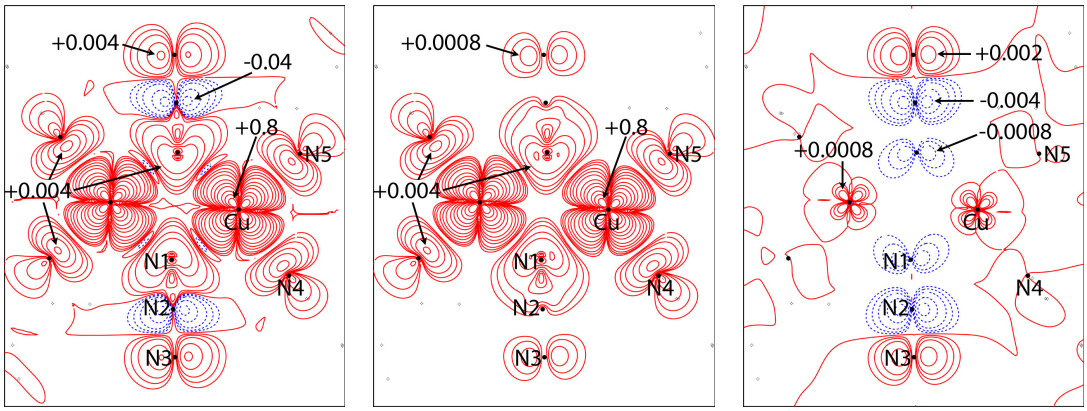
2.3.2 Total, Magnetic and Relaxation Spin Density Maps

This section discuss about the spin density maps in the chosen molecular plane, highlighting differences and analogies between different theoretical Hamiltonians for the End-On azido compound. Moreover, the decomposition of the total spin density into magnetic and relaxation components, as discussed in Section 1.8.1, is applied and the analysis on these partial maps is performed.

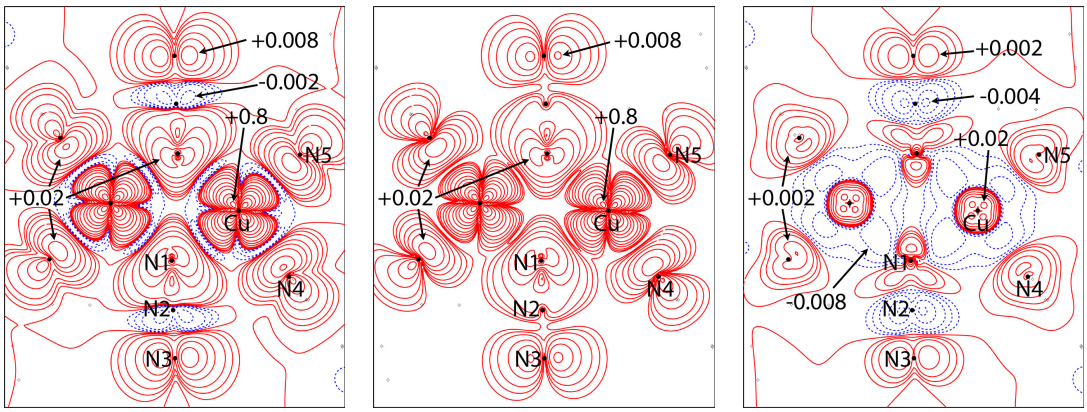
Spin density maps, along with their magnetic and relaxation components, were evaluated in the least-square plane of the four ligand N atoms around each Cu metallic centre, namely N1, N4, N5 and their symmetry equivalent N1', N4' and N5'.

The unpaired electron distributions in the ligand N planes for both employed theoretical level, CASSCF(6,6) and B3LYP, are reported in Figure 2.2, together with their magnetic and relaxation components.

From a first-look comparison of the total spin density retrieved by the two methods, it is further evident the high tendency of DFT functionals to delocalize densities. Although isosurfaces shapes are similar for the two methods, the size of positive and negative regions around the azido and ligand atoms are neatly larger in the B3LYP case than in the CASSCF one. Moreover, the DFT functional predicts maximum isosurfaces levels that are higher on the ligands atoms than the CASSCF(6,6) outcomes (+0.02 and +0.004 *a.u.* for B3LYP and CASSCF, respectively). The opposite behaviour is found around the Cu nucleus, where both methods predict a maximum isovalue line of 0.8 *a.u.* but it is more contracted in the DFT map than in the CASSCF one. Both methods agree that there is a spin polarization mechanism in the azido group, with a positive distribution on the N1 and terminal N3 atoms, while a negative p-like shaped distribution is located around the central N2 atom. The alternation of positive and negative spin densities around N2 may explain its almost zero spin population found in Table 2.2. Around N1, the isosurfaces



(a) s(**r**), CASSCF (b) s_{mag}(**r**),CASSCF (c) s_{relax}(**r**),CASSCF



(d) s(**r**), B3LYP (e) s_{mag}(**r**), B3LYP (f) s_{relax}(**r**), B3LYP

Figure 2.2: CASSCF(6,6) and B3LYP spin density maps (s(**r**), first panel) and their magnetic (s_{mag}(**r**), second panel) and relaxation (s_{relax}(**r**), third panel) contributions for the End-On complex evaluated in the least-square plane of the four ligand N atoms around each Cu. Solid red lines indicate positive value of spin density, while dotted blu denote the negative values. Solid black dots indicate the position of the atoms in the plane. Contour maps are drawn at an interval of $\pm(2,4,8) \cdot 10^n$, $-4 \leq n \leq 0$ *a.u.*.

are squeezed along the two Cu-N bond directions, resulting in triangular-like distribution. On the contrary, N3 has no deformation due to bonding and the isolines assume a p-like distribution perpendicular to the N_3^-, like for the N2 atom. CASSCF(6,6) and B3LYP methods retrieve quite different spin densities around the Cu nuclei. In particular, the latter predicts a negative region between the d-like positive distribution centred on the Cu nuclei and the ligand atoms. This region is completely absent in the multi-reference method (except for a very small area), suggesting that this phenomenon is a pure artifact of the DFT inability to correctly reproduce the spin density of a system.

The decomposition of the total spin density into magnetic and relaxation contributions highlights interesting details (see Figure 2.2, second and third column panels). First of all, it is evident that the magnetic part is responsible for the main positive features of the total spin density maps, such as the $d_{x^2-y^2}$ distribution around the Cu nuclei and the positive isosurfaces close to the ligand atoms. The two employed methods predict

similar magnetic contributions but with different spatial extent, reflecting what already found in the discussion of the total density reported above. Also, it becomes manifest that the negative regions in the total spin maps are due, as expected, to the relaxation components. Differences between relaxation contributions obtained from CASSCF and DFT are evident. The former method shows a very clean and contracted map, with well-shaped distribution around N2 (positive $s_{relax}(\mathbf{r})$) and N3 (negative $s_{relax}(\mathbf{r})$), and a small positive $d_{x^2\text{-}y^2}$ contribution close to the metallic centre. On the contrary, the relaxation component of DFT function is largely delocalized with positive and negative values. B3LYP retrieves a spherical positive distribution around the Cu nuclei and a large delocalized negative contribution between them and the ligands, that is responsible for the negative region around the Cu in the total density map.

The spin density Laplacian maps, depicted in Figure 2.3, provide further details on the behaviour of atoms. CASSCF(6,6) and B3LYP methods predict similar Laplacian distribution in the studied plane, with small differences. For instance, the two methods find a different distribution for the region around N1 and N1' atoms, where B3LYP yields a close negative distribution surrounded by a positive region while for the CASSCF the negative part is more expanded and the boundaries are not so clear. It is here essential to remark that, even if the two Laplacian maps in Figure 2.3 are qualitatively similar, they are responsible for very different spin density distributions.

The analysis of these Laplacian distributions helps the interpretation of the integrated atomic spin density Laplacians (ASDL, $\int_\Omega \nabla^2 s d\boldsymbol{r}$) obtained for some atoms in the molecules. Numerical values of ASDL are reported in Table 2.3, together with the average atomic effect on the spin density (see Section 1.8, Table 1.4)

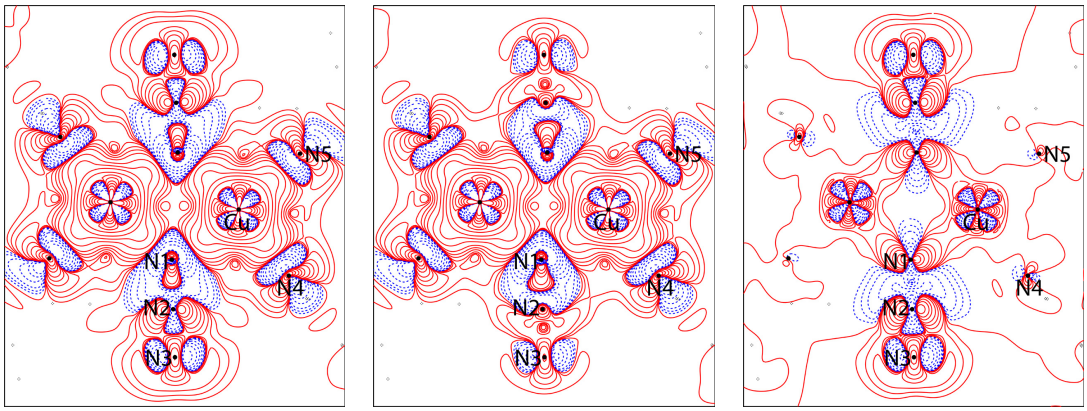
Table 2.3: QTAIM atomic spin density laplacian (ASDL, $\int_\Omega \nabla^2 s \; d\boldsymbol{r}$) for selected atoms or group of atoms (Ω) in the End-On azido complex.

Basin, Ω	ASDL, $\int_\Omega \nabla^2 s \; d\boldsymbol{r}$[a]	Average effect on s[a,b]
Cu	-0.206/-0.085	α
N1	0.098/0.036	β
N2	0.025/0.064	β
N3	-0.018/-0.047	α
Azide	0.105/0.053	β
N4	0.048/0.001	β
N5	0.048/0.002	β

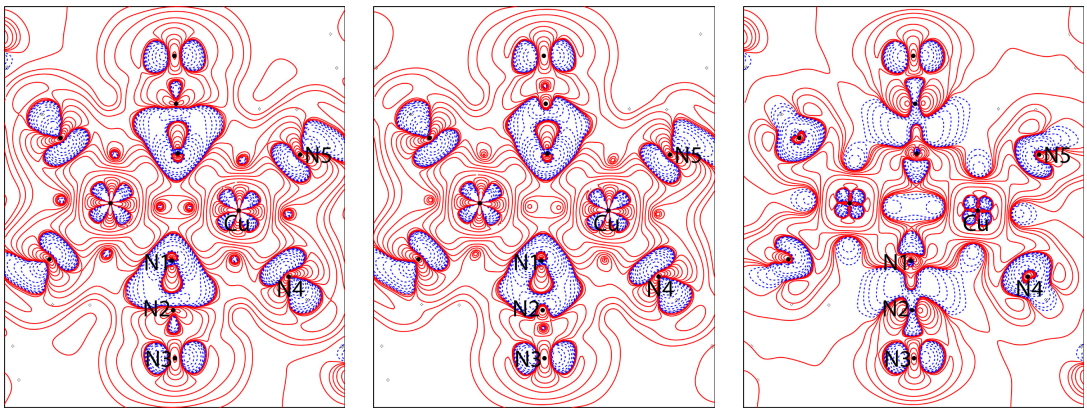
[a] CASSCF(6,6)/B3LYP data. Values are expressed in atomic units.

[b] α/β effect: the atom increases/decreases, on average, the spin density. The effect of a single reference point $\mathbf{r}$ can be different from the average behaviour, depending on its position.

At variance with the atomic-electron density Laplacian ($\int_\Omega \nabla^2 \rho \; d\boldsymbol{r}$), the integrated spin density Laplacian is usually different from zero because the atomic boundaries were



(a) $\nabla^2 s(\mathbf{r})$, CASSCF (b) $\nabla^2 s_{\mathrm{mag}}(\mathbf{r})$, CASSCF (c) $\nabla^2 s_{\mathrm{relax}}(\mathbf{r})$, CASSCF



(d) $\nabla^2 s(\mathbf{r})$, B3LYP (e) $\nabla^2 s_{\mathrm{mag}}(\mathbf{r})$, B3LYP (f) $\nabla^2 s_{\mathrm{relax}}(\mathbf{r})$, B3LYP

Figure 2.3: CASSCF(6,6) and B3LYP spin density Laplacian maps ($\nabla^2 s(\mathbf{r})$, first panel) and their magnetic ($\nabla^2 s_{\mathrm{mag}}(\mathbf{r})$, second panel) and relaxation ($\nabla^2 s_{\mathrm{relax}}(\mathbf{r})$, third panel) contributions for the End-On complex evaluated in the least-square plane of the four ligand N atoms around each Cu. Color code and isosurface levels are the same of Figure 2.2.

defined in terms of the local zero flux of $\nabla\rho$ and not of that of ∇s. The values reported in Table 2.3 provide the average behaviour of one atom to increase the spin density (ASDL < 0, α effect) or to decrease it (ASDL > 0, β effect). However, the specific effect of each atom at one reference point $\mathbf{r}$ depends on its position in the space and can be different from the average behaviour. The Cu atom has a large negative ASDL, indicating a prevailing α effect in the space (-0.206 and -0.085 for the CASSCF(6,6) and the B3LYP method, respectively). All the N atoms, except for the terminal N3 azido atoms, behave in the opposite way yielding, on average, β effects. CASSCF and B3LYP methods qualitatively agree on the sign of the ASDLs.

The analysis of the spin density Laplacian distribution in Figure 2.3 provides further details on these values. The prevalence of the α effect for the Cu atoms is due to the presence of very localized negative $\nabla^2 s$ regions (dotted blu isolines) around the metallic centres and directed towards the ligands atoms. Similarly to what found in the spin density maps, they have a $d_{x^2\text{-}y^2}$ shape, with very large absolute values and even more contracted to the nuclei than those maps. A large positive Laplacian region surrounds this d-like

negative distribution, creating a counteracting β effect from the same basin. The Cu α effect prevails over the β one, leading to the atomic integrated Laplacian values reported in Table 2.3.

The two bridging N1 atoms bear negative $\nabla^2 s$ pointing toward the direction of the Cu-N bonds, yielding as for the Cu an α effect. Although this region is quite wide, it is unable to prevail over the β effect coming from the Cu basin. The α effect regions of the bridging N1 make manifest the ability of these atoms to determine the delocalization of the unpaired density. A spin polarization mechanism is evident along the azido moiety, where N2 is dominated by β effect while the terminal N3 behaves in an opposite way. While the N2 positive $\nabla^2 s$ distribution has no clear shape, the N3 shows a π-like distribution as highlighted in the spin density maps.

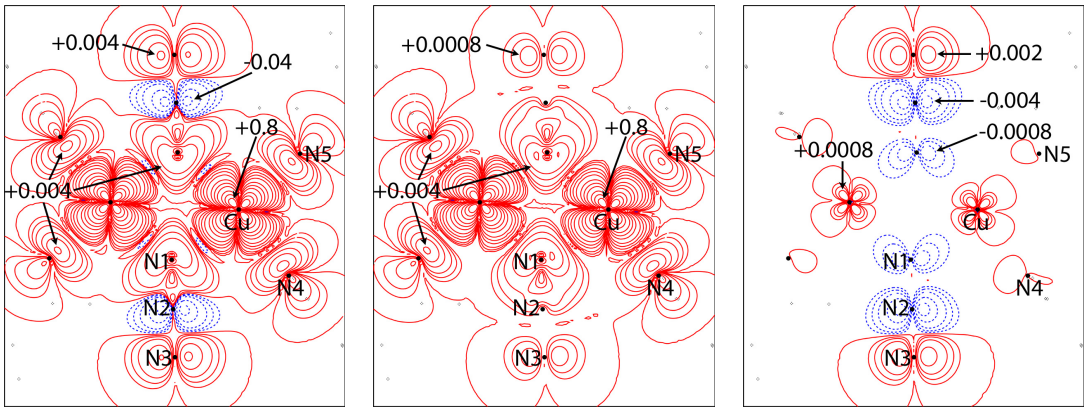
The decomposition of the Laplacian maps into magnetic and relaxation terms (Figure 2.3, second and third panels) shows that the major differences in the two distributions reside in the relaxation contributions. These are significantly different in shape and local scalar magnitude, with CASSCF distribution more contracted around the Cu and N_3^- atoms and DFT more delocalized. On the contrary, the magnetic components are more closer to each other, at least for their shape, and they are the main contributors to the reconstruction of the total Laplacian distribution.

2.3.3 Source Function Partial Densities Reconstructions

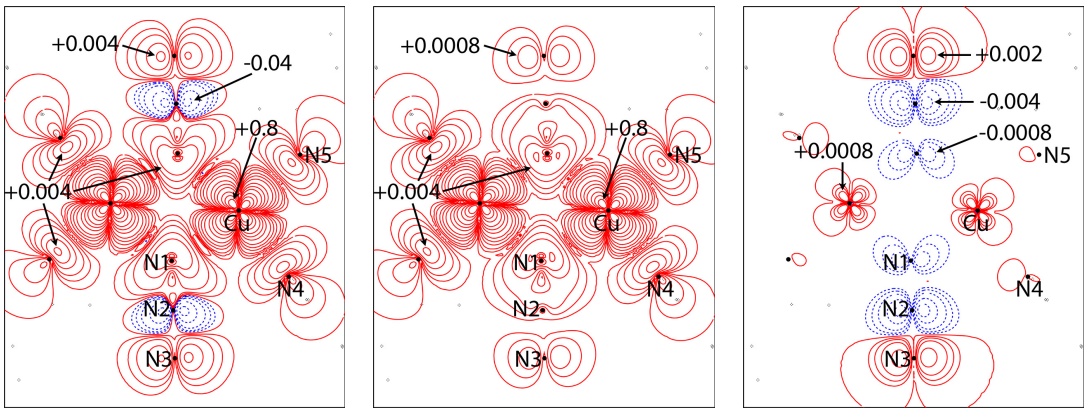
Source Function partial densities provide a *cause-effect* view of the spin density. Figures 2.4 and 2.5 show partial densities obtained only using subsets of atoms at the CASSCF(6,6) and B3LYP level of theory. The notation used to express the subset chosen is the same of equation 1.35, where the labels of the employed atoms are reported in braces. That said, the considered group subsets are {Cu,N,C}, {Cu,N}, {N} and $\{N_{azido}\}$, where {Cu}, {N} and {C} denote the subsets including the two Cu atoms, all the nitrogen atoms and the carbon atoms belonging to pyridine rings, while $\{N_{azido}\}$ refers only to azido N atoms.

When all the atoms but the Hydrogens and the *tert*-butyl groups are considered in the spin density reconstructions ({Cu,N,C} subset), the partial maps obtained are almost indistinguishable from their primitive ones, independently from the method used (compare S_s{Cu,N,C} maps in Figures 2.4 and 2.5 with s($\mathbf{r}$) in Figure 2.2). This fact suggests that the excluded atoms give a null, or very small, spin contribution in the selected plane. Actually, some small differences are present in the peripheral regions of the maps. Besides the negligible contribution due to the excluded atoms, these are the result of numerical errors due to the finite precision of the integration required to calculate the source function contributions. However, since these issues do not concern regions we are interested in, the introduction of these small errors has no influence on the discussion.

When the contribution from the pyridine C atoms is removed (compare S_s{Cu,N} row with S_s{Cu,N,C} one in Figures 2.4 and 2.5), important differences become manifest. For the DFT case, the negative region surrounding the Cu atoms showed in Figures 2.2 and



(a) $S_s\{Cu,N,C\}$ (b) $S_{s,mag}\{Cu,N,C\}$ (c) $S_{s,relax}\{Cu,N,C\}$

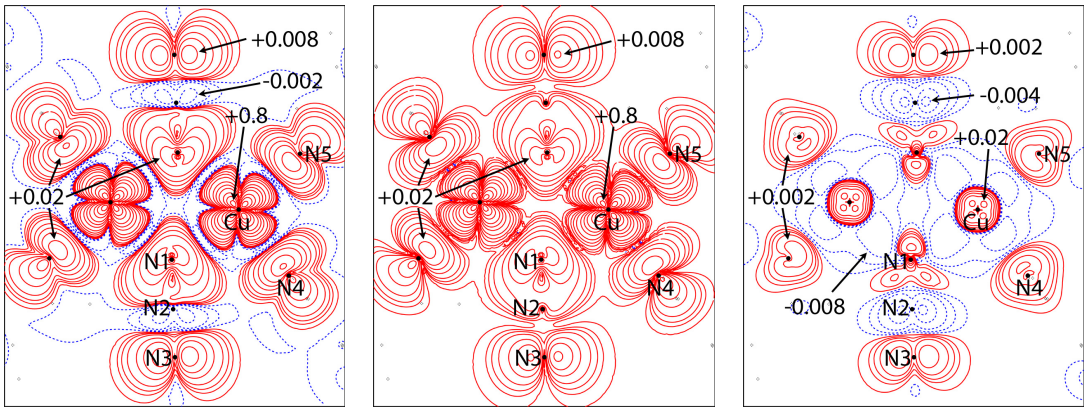


(d) $S_s\{Cu,N\}$ (e) $S_{s,mag}\{Cu,N\}$ (f) $S_{s,relax}\{Cu,N\}$

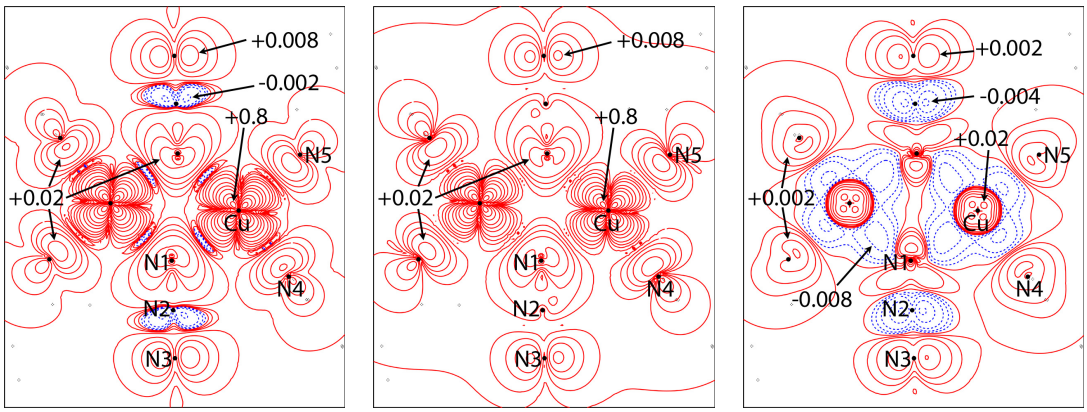
Figure 2.4: CASSCF(6,6) source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-On complex evaluated in the least-square plane of the four ligand N atoms around each Cu. Subsets containing Cu, N and C atoms ({Cu,N,C}) and without the C contributions ({Cu,N}) have been considered. Color code and isosurface levels are the same of Figure 2.2.

2.5 completely disappears, revealing that this negative spin density region is mostly due to the C atoms of the pyridine rings. The same variation occurs in the relaxation partial spin density maps where the negative regions contract, while in the magnetic partial maps there is a contraction of the very high positive density around the metallic centre (Figure 2.5). CASSCF differences among $S_s\{Cu,N,O\}$ and $S_s\{Cu,N\}$ are minimal and only involves the expansion/contraction degree of the positive spin distributions (see Figure 2.4). The fact that only B3LYP densities appear to be significantly affected by the exclusion of the carbon source function contributions further confirms what stated in the previous sections: B3LYP functional exaggerates the spin delocalization from the paramagnetic centre to the non-magnetic ones if compared to CASSCF.

It is worth saying that source function analysis not only permits to analyse the effect of a magnetic atom towards the other basins, but also it allows to see the opposite way, from the non-magnetic centres to the magnetic one. This reverse direction may be only appreciated through the source function descriptor, which by the way describes this in a



(a) S_s{Cu,N,C}, B3LYP (b) $S_{s,mag}${Cu,N,C}, B3LYP (c) $S_{s,relax}${Cu,N,C}, B3LYP



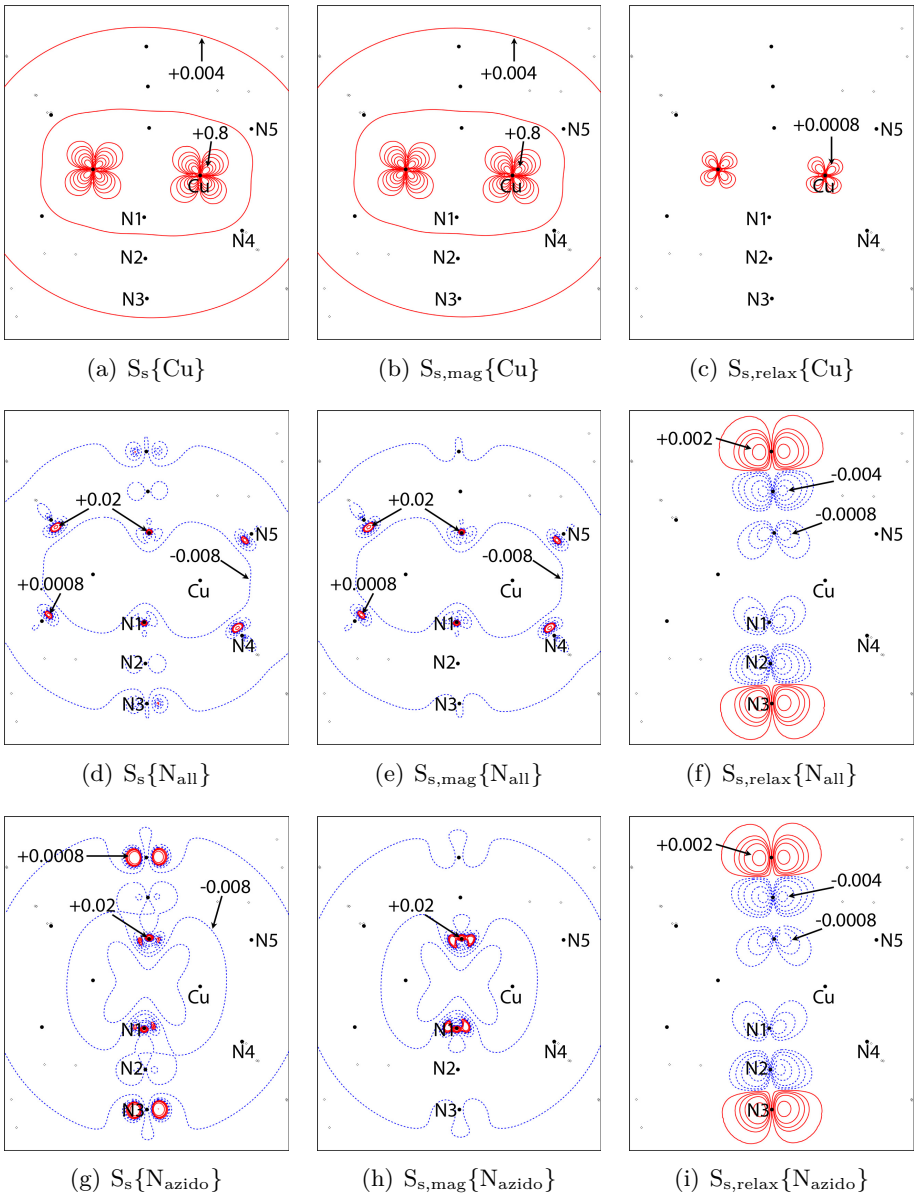
(d) S_s{Cu,N}, B3LYP (e) $S_{s,mag}${Cu,N}, B3LYP (f) $S_{s,relax}${Cu,N}, B3LYP

Figure 2.5: B3LYP source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-On complex evaluated in the least-square plane of the four ligand N atoms around each Cu. Subsets containing Cu, N and C atoms ({Cu,N,C}) and without the C contributions ({Cu,N}) have been considered. Color code and isosurface levels are the same of Figure 2.2.

cause-effect context.

Figures 2.6 and 2.7 show the source function contributions yielded by the Cu, all the N and the azidic N atoms together with their magnetic and relaxation terms.

As expected, the S_s{Cu} contributions are the most relevant sources that impart the $d_{x^2-y^2}$ distribution around the metallic centres. These contributions are everywhere positive and magnetic and relaxation components cooperate to retrieve the total density, with the former being the most dominant term. The maximum total contribution coming from the Cu atoms is slightly higher and more contracted for the CASSCF level of theory than for DFT (compare Figure 2.6(a) and 2.7(a)). The two proposed methods retrieve similar total and magnetics S_s{Cu} from a qualitatively point of view, but they completely disagree on the relaxation term. B3LYP predicts large regions of negative spin density sources around the positive spherical area close to the Cu nuclei (Figure 2.7(c)). Instead, CASSCF retrieves the classical $d_{x^2-y^2}$ shape distribution, but with a very low and contracted isocontours (Figure 2.6(c)).



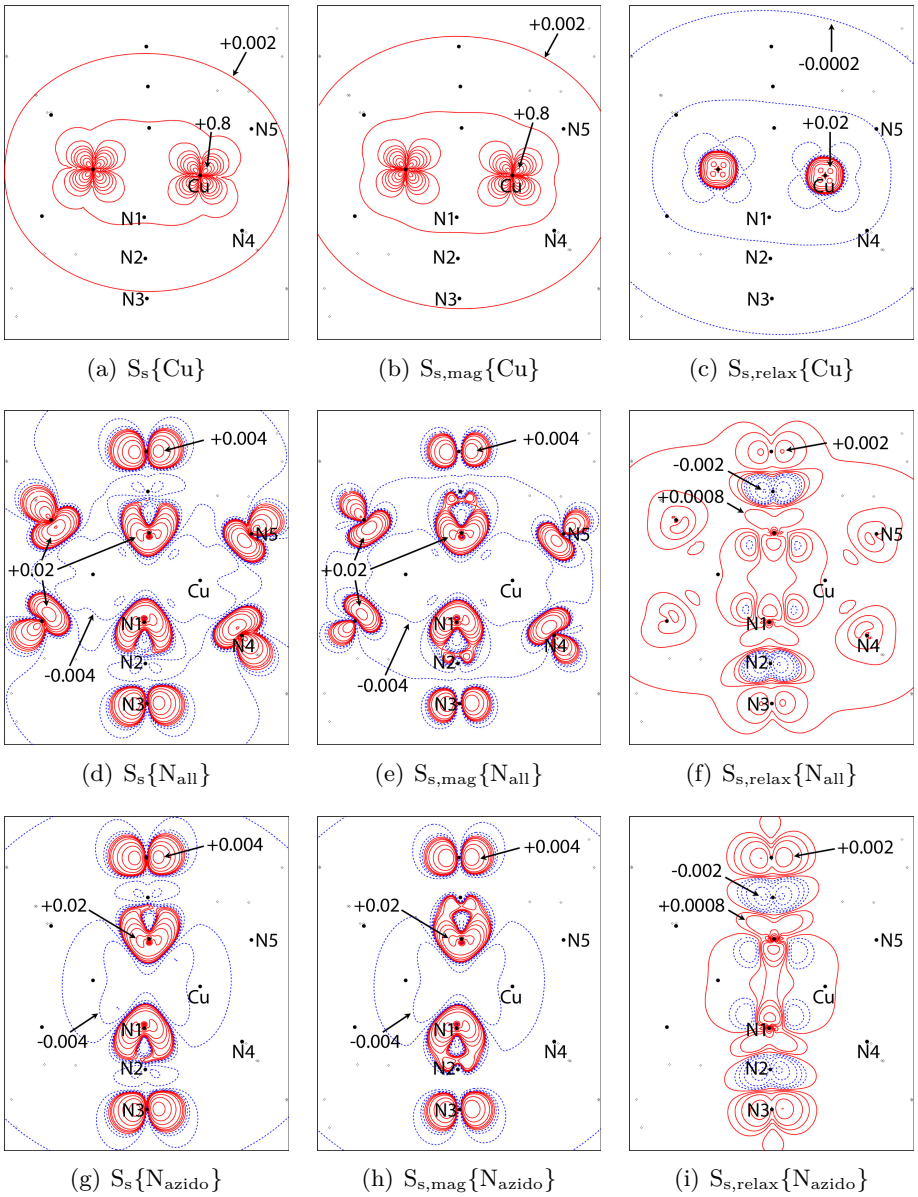
(a) $S_s\{Cu\}$ (b) $S_{s,mag}\{Cu\}$ (c) $S_{s,relax}\{Cu\}$

(d) $S_s\{N_{all}\}$ (e) $S_{s,mag}\{N_{all}\}$ (f) $S_{s,relax}\{N_{all}\}$

(g) $S_s\{N_{azido}\}$ (h) $S_{s,mag}\{N_{azido}\}$ (i) $S_{s,relax}\{N_{azido}\}$

Figure 2.6: CASSCF(6,6) source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-On complex evaluated in the least-square plane of the four ligand N atoms around each Cu. Subsets containing Cu atoms ($\{Cu\}$), all the N atoms ($\{N_{all}\}$) and only the azido N atoms ($\{N_{azido}\}$) have been considered. Color code and isosurface levels are the same of Figure 2.2.

The contributions coming from all the N atoms of the system ($S_s\{N_{all}\}$) and from the N atoms belonging to the azido moieties ($S_s\{N_{azido}\}$) are also reported in Figures 2.6 and 2.7. At the CASSCF(6,6) level (Figure 2.6), nitrogen atoms yield β effect except for very small regions around the nuclei, where the contributions are positive and highly concentrated. Although the β effect from the N atoms is weak, it is locally able to overcome the α effect coming from the Cu atoms in the region close to middle N2 atoms (see Figures 2.2 and 2.4). When only the azido Nitrogens are considered, these positive regions slightly increase



(a) $S_s\{Cu\}$ (b) $S_{s,mag}\{Cu\}$ (c) $S_{s,relax}\{Cu\}$

(d) $S_s\{N_{all}\}$ (e) $S_{s,mag}\{N_{all}\}$ (f) $S_{s,relax}\{N_{all}\}$

(g) $S_s\{N_{azido}\}$ (h) $S_{s,mag}\{N_{azido}\}$ (i) $S_{s,relax}\{N_{azido}\}$

Figure 2.7: B3LYP source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-On complex evaluated in the least-square plane of the four ligand N atoms around each Cu. Subsets containing Cu atoms ($\{Cu\}$), all the N atoms ($\{N_{all}\}$) and only the azido N atoms ($\{N_{azido}\}$) have been considered. Color code and isosurface levels are the same of Figure 2.2.

in size, underlining that the negative contributions from the pyridinic N atoms partially inhibit the azidic positive sources. On the contrary, at the B3LYP level of theory the N atoms yield larger and more delocalised α effects (Figure 2.7) when compared to CASSCF. The shapes of these positive regions reflect the presence of the Cu atoms. N1 and N1' bridging atoms have a split-lobe (triangular-like) shape, with two edges perpendicular to the two Cu-N bonds. Pyridine N atoms instead have a T-like shape, with the head directed toward the metallic centres. Differently from the CASSCF case, the removal of

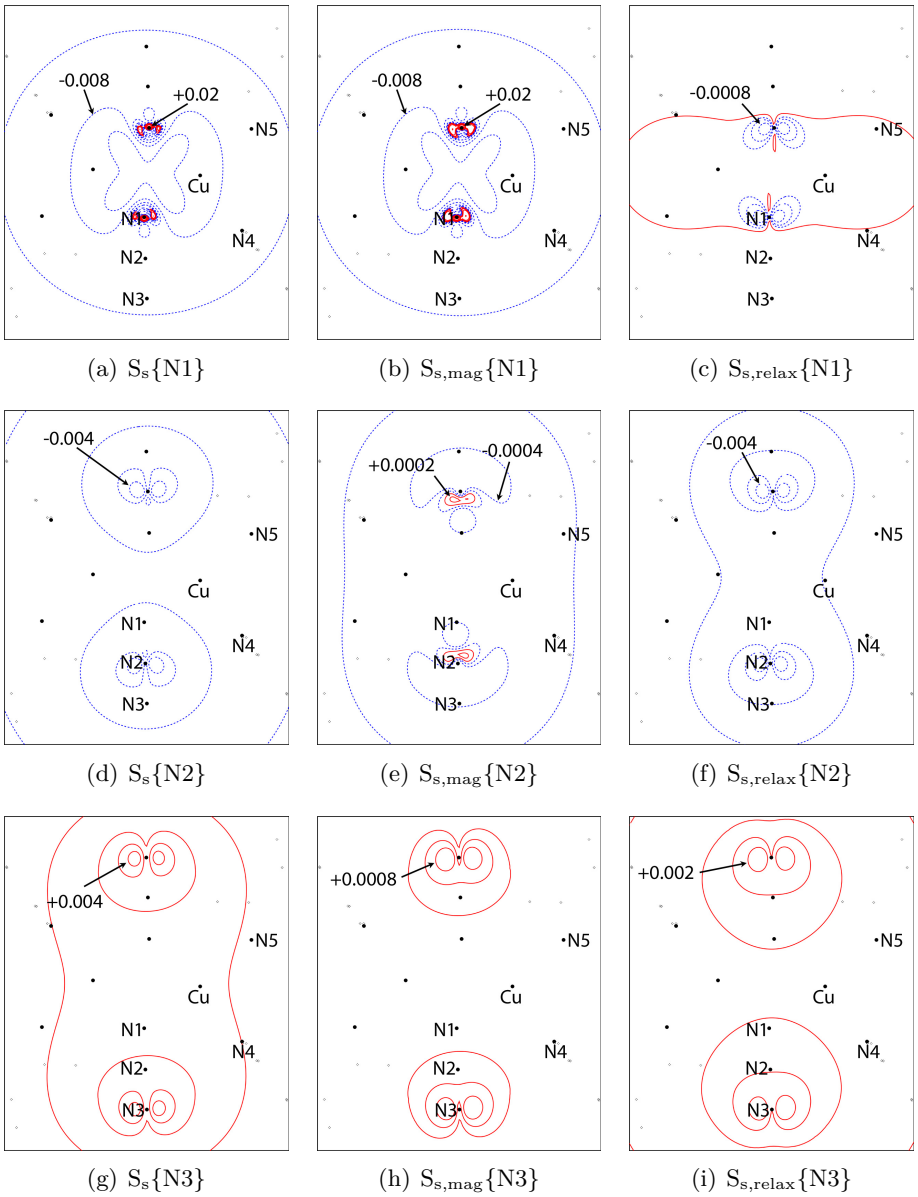
the S_s{N4,N5} contributions seem to not have large effect on the azido terms. These two sets of N atoms appear to be more independent than for the multi-reference calculation, prompting that an higher delocalization of unpaired electrons does not necessarily imply a greater connection between them. The magnetic and relaxation terms from the two applied methods largely differ as well (Figures 2.6 and 2.7). The magnetic term imparts the shape of the spin density source at the B3LYP level, while the relaxation partial density is largely delocalized and only has a minor role (Figure 2.7). On the other hand, at the CASSCF level the relaxation term plays a major role and equally cooperates with the magnetic contribution to the reconstruction of the partial density, especially for the central and terminal nitrogen atoms (N2 and N3) of the azido groups (Figure 2.6).

Deeper details can be obtained analysing the single contributions from the azido N atoms, to distinguish their different roles. The reconstructed spin density for these atoms are depicted in Figures 2.8 and 2.9 for the CASSCF(6,6) and the B3LYP level of theory, respectively.

Both CASSCF and DFT approach retrieve qualitatively similar behaviours. The bridging N1 atom and its symmetry-equivalent atom N1' yield both local positive and diffuse negative sources, with a spin polarization mechanism of sources with their bonded atoms. The central N2 and the terminal N3 instead yield only negative and positive diffused sources, respectively. The sources from the N1 atom are dominated by the magnetic term, both for CASSCF and DFT methods. However, the shapes and magnitudes recovered with the two models are quite different. As seen before, the former predicts more contracted positive sources than DFT around the bridging nitrogen atoms (compare Figures 2.8 and 2.9).

For the N2 and N3 atoms, the relaxation contributions assume a major role, becoming dominant in the case of the central N2 and N2' nitrogen atoms and yielding β effect regions. The source function spin density and both its components have a π-like shape. This is reminiscent of the symmetry of the unoccupied molecular orbitals which result by mixing the $d_{x^2\text{-}y^2}$ metal orbitals with the unoccupied π_u azido orbitals within a fragment orbital approach.[87] The magnetic and relaxation terms are equally important for the terminal N3 atoms, but CASSCF(6,6) and B3LYP differ on the relative weights of these two components in defining the total spin density. The multi-configurational approach predicts the relaxation term to be the most relevant, while B3LYP finds that $S_{s,mag}${N3} prevails over the $S_{s,relax}${N3} contribution.

To conclude this part, it has been shown that the local spin density source functions from a single basin are in general quite different from that due to the sources of the closest atoms on that same basin. These differences provide a local measure of the co-operating effects inherent to the electron spin delocalization and polarization mechanisms. Saying it in other word, apart from the Cu sources, the other atomic contributions are unable to describe the spin density distribution by themselves: the effect of other atoms is, in general, far from being negligible.



(a) $S_s\{N1\}$ (b) $S_{s,mag}\{N1\}$ (c) $S_{s,relax}\{N1\}$

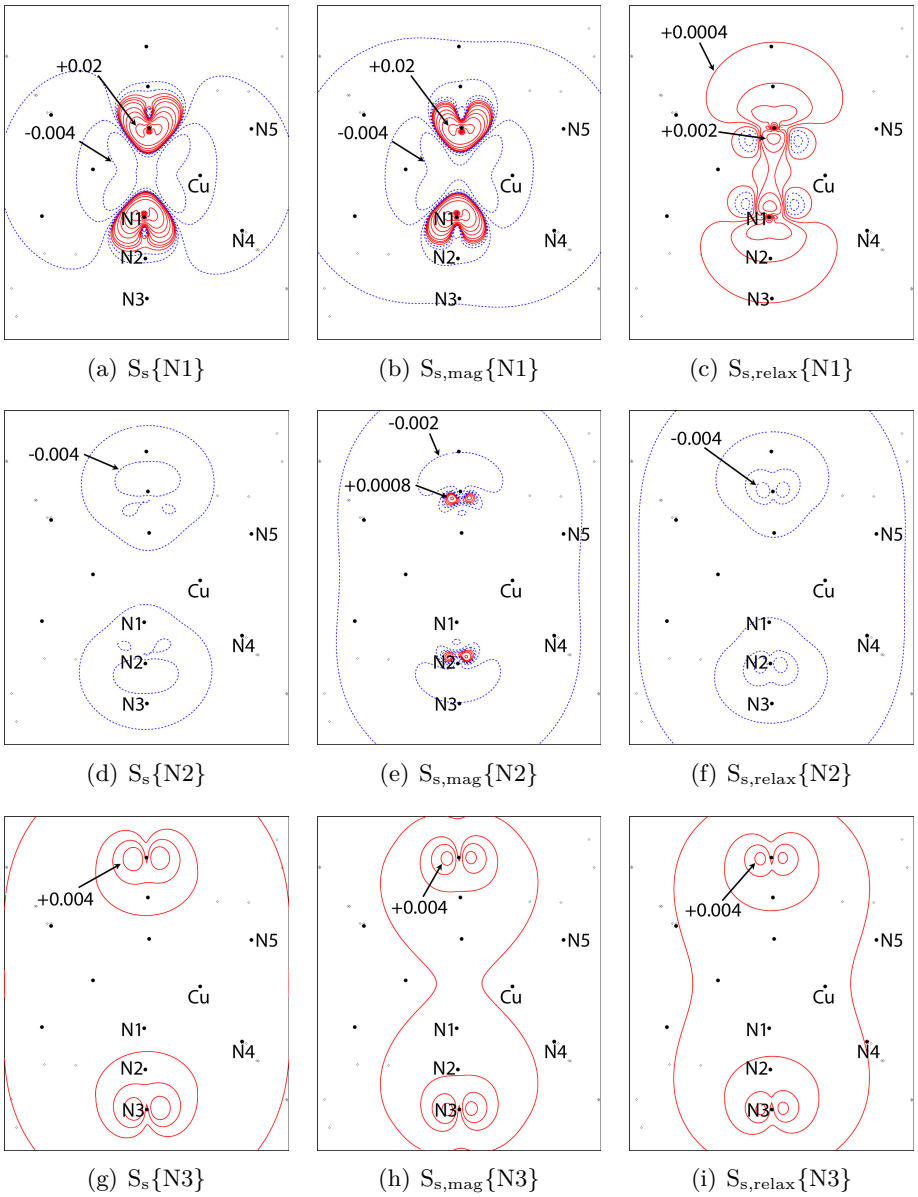
(d) $S_s\{N2\}$ (e) $S_{s,mag}\{N2\}$ (f) $S_{s,relax}\{N2\}$

(g) $S_s\{N3\}$ (h) $S_{s,mag}\{N3\}$ (i) $S_{s,relax}\{N3\}$

Figure 2.8: CASSCF(6,6) source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-On complex evaluated in the least-square plane of the four ligand N atoms around each Cu. Subsets formed by N1+N1' atoms ({N1}), N2+N2' atoms ({N2}) and N3+N3' atoms ({N3}) have been considered. Color code and isosurface levels are the same of Figure 2.2.

2.3.4 Cu(II) 3d Electron Asphericity

In the End-On complex the Cu(II) cation has a square-planar coordination, with four different ligands (two azido and two pyridine) coordinating through their N atoms. According to the crystal field theory, for this geometry and taking into account the electronic structure of the copper (Cu(II) - d^9 configuration), the half-occupied orbital (the magnetic one) should be the $d_{x^2-y^2}$ based on the axis orientation shown in Figure 2.32. As already



(a) $S_s\{N1\}$ (b) $S_{s,mag}\{N1\}$ (c) $S_{s,relax}\{N1\}$

(d) $S_s\{N2\}$ (e) $S_{s,mag}\{N2\}$ (f) $S_{s,relax}\{N2\}$

(g) $S_s\{N3\}$ (h) $S_{s,mag}\{N3\}$ (i) $S_{s,relax}\{N3\}$

Figure 2.9: B3LYP source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-On complex evaluated in the least-square plane of the four ligand N atoms around each Cu. Subsets formed by N1+N1' atoms ({N1}), N2+N2' atoms ({N2}) and N3+N3' atoms ({N3}) have been considered. Color code and isosurface levels are the same of Figure 2.2.

discussed in the theoretical introduction, evidences of this simple model should be visible studying the distribution of the Laplacian of the electron density.[16] In particular, the asphericity of the 3d electron distribution around a metal atom, due to an inhomogeneous occupation of the 3d valence orbitals, should be manifest.[16,24,98,99]

The topological analysis of the $-\nabla^2\rho$ in the valence-shell concentration and depletion regions finds four minima and four maxima around the Cu atoms. The four (3,+3) minima of $-\nabla^2\rho$, *i.e.* charge depletion points (CDs), lye almost on the Cu-N internuclear axes,

while the four (3,-3) maxima, *i.e.* charge concentration points (CCs), are found in between the CDs. A sketch of the critical points found and their positions is represented in Figure 2.10.

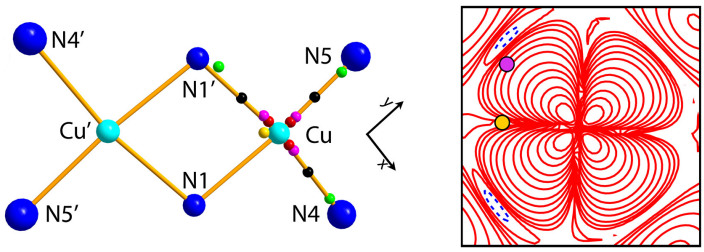


Figure 2.10: Cu atom 3d electron asphericity in the End-On azido Cu complex. Left: Critical points distribution around the Cu atom. Only the charge concentration critical point between the two Cu atoms is shown. Reference points colour code is: pink: $-\nabla^2\rho$ (3,+3) charge depletion (CD); red: $-\nabla^2\rho$ (3,+1) ring-like (R); yellow: $-\nabla^2\rho$ (3,-3) charge concentration (CC) on the Cu atom; green: $-\nabla^2\rho$ (3,-3) charge concentration (CC) on the N atom; black: ρ (3,-1) bond critical point (BCP). Right: location of (3,-3) CC between the two Cu atoms and (3,+3) CD along the Cu-N1' bond on the spin density map evaluated in the least-square plane of the four ligand N atoms around each Cu.

The CDs are located at 0.43 Å from the Cu nucleus and may be associated with the lobes of the magnetic singly occupied $d_{x^2-y^2}$ orbital. On the contrary the CCs are located at 0.28 Å from the metallic centre and should be related to the lobes of the d_{xy} orbital. If such an association is plausible, one would anticipate a large spin density value at the CDs critical points and a smaller value at the CCs points. Moreover, CDs should be characterized by positive and low $-\nabla^2\rho$ values if compared to the CCs ones, where the magnitude of the Laplacian should be negative and higher in absolute value. The reason of this is because the spin and the electron density have a different behaviour and concentrate/dilute in distinct ways, as already stated in Section 1.8 in Chapter 1.

Results in Table 2.4 confirm these predictions.

The CC point located between the two Cu atoms has, as predicted, a very low spin density ($s < 0.001$ *a.u.*) and a very negative $-\nabla^2\rho$ ($-\nabla^2\rho = -104.95$ *a.u.*), while on the contrary, the same quantities for the CD located along the Cu-N1 axis are 0.413 and 19.33 *a.u.*, respectively, at the CASSCF level of theory. B3LYP functional retrieves qualitatively similar results, with s values smaller at the CD point and larger for the CC point when compared to the multi-reference method. This behaviour is in some way connected to the tendency of DFT functionals to exaggerate both the spin delocalization towards the ligands and the local spin relaxation mechanism. High spin density values at the CD points are purely due to the magnetic orbitals, with a null or close to zero relaxation contribution. CDs for the other metal-ligand axes (Cu-N4 and Cu-N5) result almost identical in magnitudes to what found for the Cu-N1 CD.

Table 2.4 also reports the spin density and electron density Laplacian values evaluated at the (3,+1) $-\nabla^2\rho$ critical points (ring-like critical points). These stationary points were

Table 2.4: CASSCF(6,6)/B3LYP distance from the Cu atom (R_{Cu}), electron density laplacian ($-\nabla^2\rho(\boldsymbol{r})$), spin density ($s(\boldsymbol{r})$) and its magnetic ($s_{mag}(\boldsymbol{r})$) and relaxation ($s_{relax}(\boldsymbol{r})$) components evaluated for some critical points (CPs) along Cu-Cu', Cu-N1, Cu-N4 and Cu-N5 axes. Distances are expressed in Å, laplacian and spin density values in atomic units *a.u.*.

Bond, CP[a]	R_{Cu}	$\nabla^2\rho(\boldsymbol{r})$	$s(\boldsymbol{r})$	$s_{mag}(\boldsymbol{r})$	$s_{relax}(\boldsymbol{r})$
Cu-Cu', CC	0.277/0.278	-104.95/-94.76	0.000/0.011	0.000/0.000	0.000/0.011
Cu-N1, CD	0.429/0.434	19.33/18.04	0.413/0.278	0.412/0.274	0.001/0.004
Cu-N1, R	0.292/0.288	-28.37/-41.13	1.030/0.750	1.029/0.733	0.001/0.017
Cu-N1, CC	1.624/1.606	-2.07/-1.77	0.005/0.019	0.006/0.019	-0.001/0.000
Cu-N1, BCP	0.939/0.967	0.36/0.28	0.007/0.002	0.007/0.003	0.000/-0.001
Cu-N4, CD	0.430/0.434	19.41/18.13	0.416/0.286	0.416/0.282	0.000/0.004
Cu-N4, R	0.292/0.288	-29.27/-42.04	1.051/0.792	1.050/0.774	0.001/0.018
Cu-N4, CC	1.597/1.595	-2.74/-2.38	0.008/0.036	0.008/0.032	0.000/0.004
Cu-N4, BCP	0.925/0.954	0.39/0.28	0.007/0.001	0.007/0.002	0.000/-0.001
Cu-N5, CD	0.429/0.434	19.48/18.17	0.418/0.284	0.418/0.280	0.000/0.004
Cu-N5, R	0.292/0.288	-28.87/-41.96	1.057/0.788	1.056/0.770	0.001/0.018
Cu-N5, CC	1.593/1.591	-2.69/-2.35	0.008/0.033	0.008/0.030	0.000/0.003
Cu-N5, BCP	0.923/0.953	0.39/0.29	0.008/0.001	0.008/0.002	0.000/-0.001

[a] CD: $-\nabla^2\rho$ (3,-3) charge depletion; R: $-\nabla^2\rho$ (3,+1) ring-like; CC: $-\nabla^2\rho$ (3,+3) charge concentration; BCP: ρ (3,-1) bond critical point.

found along the Cu-ligand axis, as for the CDs, but closer ($\approx$ 0.28 Å from the Cu nucleus). At this point the value of $-\nabla^2\rho$ is negative as for the CCs but lower in absolute value, while the total $s(\boldsymbol{r})$ is much larger than the CDs values ($s \approx 1.05$ and 0.78 for CASSCF and B3LYP methods, respectively). This ring-like critical points are also associated with the $d_{x^2\text{-}y^2}$ orbital. These additional data further corroborate the analysis concerning the Cu asphericity.

The s and $-\nabla^2\rho$ values evaluated at the CCs lying in the valence shell charge of the coordinating N atoms are also shown in Table 2.4. These CCs, one for each N atom except for the azido N1 which has two equivalent (3,-3) $-\nabla^2\rho$ maxima, are almost aligned in a *key-lock* arrangement with the corresponding CDs close to the Cu nuclei. At CASSCF level, the spin density values in these CC points are low, at least two order of magnitude if compared to the CD values on the Cu and the Laplacian is negative but close to zero. The major contributions to the spin density come from the magnetic terms. On the other hand, B3LYP predicts s values three times larger than the CASSCF, while the Laplacian values are similar.

Besides the already discussed Laplacian points, Table 2.4 shows also the s and $-\nabla^2\rho$ values evaluated at the bond critical points (BCPs). Here the situation is similar to that described for the CCs located on the N atoms but with some differences. Both methods agree that total $s(\boldsymbol{r})$ is very low, sometimes negligible, and the $-\nabla^2\rho$ are almost zero. It is here worth noting that the DFT functional locates the BCP farther from the Cu atoms than the CASSCF method ($\approx$ 0.96 Å for the B3LYP and 0.93 Å for the CASSCF(6,6), on average). This fact further confirms the DFT excessive spin delocalization towards the

ligands and its description of an increased covalent character of the Cu-N bonding when compared to CASSCF wavefunction.

2.3.5 Source Function at Selected Reference Points

Some of the points reported in Table 2.4 have been chosen as reference points for the classical source function percentage analysis (see Equation 1.33)) together with Cu', N1 and N4' nuclei as additional reference points.

The first reference point chosen is the $-\nabla^2\rho$ (3,+3) CD of Cu' along Cu'-N1 axis (which is equal to the Cu-N1 CD) and the results are reported in Figure 2.11.

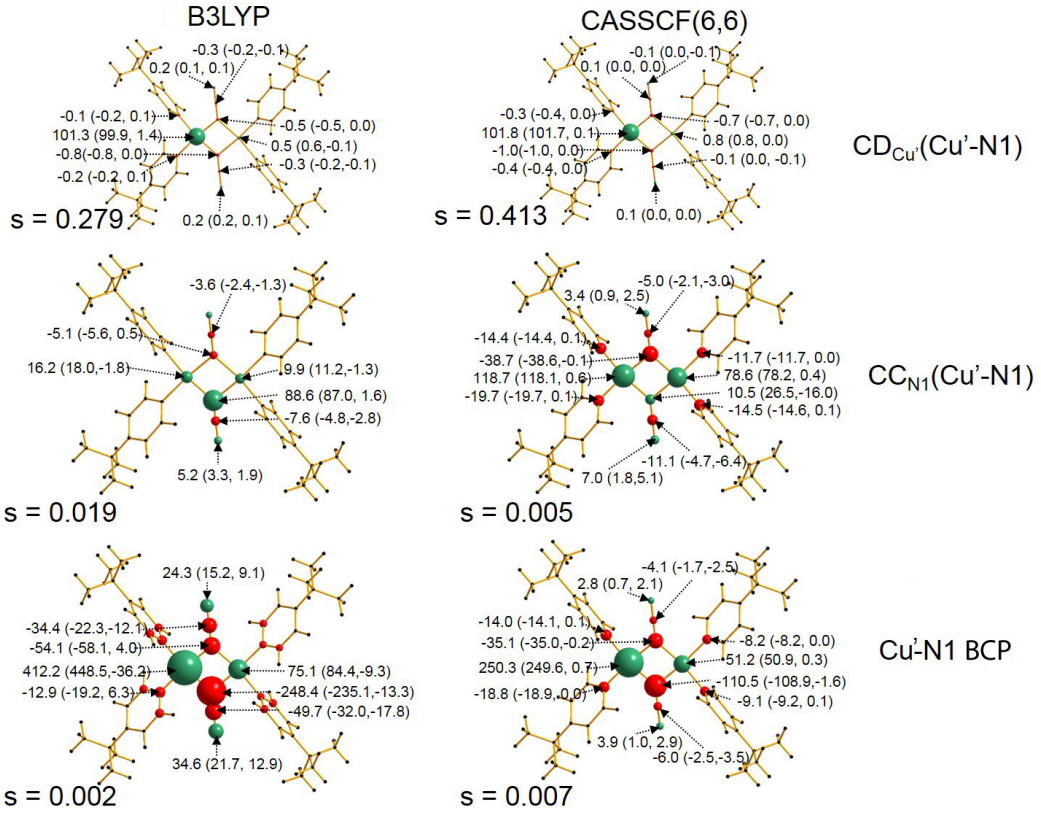


Figure 2.11: B3LYP and CASSCF(6,6) SF percentage contributions to the spin density $s(\boldsymbol{r})$ at three selected reference points along the Cu'-N1 bond in the End-On complex: (i) charge depletion (CD), (ii) charge concentration (CC) close to N1 nucleus, (iii) bond critical point (BCP). Numerical values represent the total(magnetic/relaxation) percentage contributions. Green/red spheres denote α/β effect on the spin density of the selected reference point.

The spin density at this point is dominated by the contribution from Cu' (101.8% and 101.3% at the CASSCF and B3LYP levels, respectively) and the decomposition in magnetic and relaxation contributions shows that the former is the most important term (101.7% and 99.9% at CASSCF and B3LYP levels, respectively). The source contributions from the azido atoms are very small in percentage but, taking into account the spin density value at this reference point, they are not negligible in absolute value. A |1%| value for such contributions amount to ≈ 0.003 and ≈ 0.004 *a.u.* for B3LYP and CASSCF,

respectively, which have the same order of magnitude of the total spin density evaluated at the Cu'-N1 BCP (≈ 0.002 and ≈ 0.007 *a.u.* for B3LYP and CASSCF, respectively). This fact underlines the importance of the spin back-delocalization mechanism from the ligand atoms to the magnetic centre. The major contributions come from the N1 atoms, but also N2 and N3 atoms yield small effects that cooperate positively and negatively in the reconstruction of the density at the selected point. Besides, also the other Cu atom has a non negligible contribution (0.5% for the DFT functional, 0.8% for the CASSCF method). The sign of this source term is positive as for Cu', revealing the presence of a relevant spin delocalization process between the paramagnetic centres and confirming in some way the ferromagnetic coupling present among the metallic centres. The decomposition of these contributions in the classic magnetic and relaxation terms shows that, as expected, the magnetic part gives the main contribution to the total reconstruction (see Figure 2.11).

The spin density at the CC point located along the Cu'-N1 axis and close to the N1 atom is low (0.019 and 0.005 *a.u.* at the B3LYP and CASSCF(6,6), see Table 2.4 and second row of Figure 2.11) when compared with the (3,+3) and (3,+1) points along the same line but close to the metallic centre. The N1 bridging atom that holds the CC reference point and both Cu atoms produce α-effects, while the symmetric N1' and all the other pyridine N atoms tend to contrast the spin density accumulation yielding β-effects. This behaviour highlights a spin delocalization mechanism from one paramagnetic centre to the other through the bridging N1 atom. All the sources of atoms along this ferromagnetic pathway cooperate to the reconstruction of the total spin density of the CC point, while all the other sources from atoms linked through other magnetic pathways oppose. As expected, the source function patterns are dominated by the magnetic components. Both B3LYP and CASSCF(6,6) outcomes are qualitatively in agreement concerning the α and β effect yielded by the single atoms, but they disagree from a quantitative point of view. B3LYP functional predicts the larger source contribution coming from the bridging N1 atom (88.6%), while the two Cu atoms yield smaller contributions (16.2% and 9.9% from Cu' and Cu, respectively). On the contrary, the multi-determinant method predicts a complete opposite situation. In this case the two Cu atoms are the major sources (118.7% and 78.6% from Cu' and Cu, respectively) while the N1 atom only contributes with a smaller percentage equal to 10.5%, which is comparable or even lower in absolute value than the sources coming from other atoms. The prediction by the employed DFT functional is a clear consequence of its tendency to overestimate the spin delocalization to the ligands if compared to the CASSCF outcomes. The CASSCF(6,6) gives a much larger relative weight to the opposing effect from sources related to the atoms not directly involved in the magnetic pathways than B3LYP does. This result suggests that a larger spin delocalization does not necessary imply a relatively larger source function communication among the atoms involved in the spin delocalization.

The positive spin density area around the bridging N pointing towards the Cu(II) cation, showed in Figure 2.2 and symbolized by the CC reference point, may be explained in term of orbital language in two different ways. The first way through a spin delocaliza-

tion mechanism through the mixing of the Cu $d_{x^2-y^2}$ orbitals and the π *gerade* and *ungerade* orbitals of the N_3^- ligands.[87] The second possibility is to consider a superexchange mechanism between the two paramagnetic centres and the bridging N1 when the angle Cu-N-Cu' is around 90°.[100] Both of these descriptions rely on models. On the contrary, the source function analysis provides a rigorous model-free picture of these interaction mechanisms at the selected reference point, in this case the N1 CC point, in terms of delocalised and cooperating sources along the Cu'-N1-Cu pathways.

The last row panels of Figure 2.11 report the atomic source function contributions for the Cu'-N1 BCP reference point. The total spin density in this point is positive but quite close to zero (0.002 and 0.007 *a.u.* for B3LYP and CASSCF(6,6), respectively). The source pattern is dominated for both levels of theory by the Cu magnetic contribution which yields a positive α effect (412.2/448.5% and 250.3/249.6% total/magnetic contributions for B3LYP and CASSCF(6,6) level of theory, respectively). Cu' atom cooperates with the other metallic centre to increase the spin density at the reference point, while the bridging and central N atoms of the azide moiety together with the pyridine nitrogen atoms counteract this effect through their β effect. This Cu-N-Cu' α/β effect alternation implies that these three atoms introduce a spin polarization mechanism in the bonding region at the BCP. B3LYP and CASSCF(6,6) qualitatively agree on the relative weights of the atomic contribution, but they differ from a quantitative point of view and on the relative weight of the magnetic and relaxation contributions. B3LYP relaxation terms are on average larger than the CASSCF(6,6) ones and they usually oppose to the magnetic terms.

The source function contributions evaluated at the Cu', N1 and N4' nuclear positions are shown in Figure 2.12.

The spin density evaluated at the Cu' nucleus position is as low as -0.018 and -0.003 *a.u.* at B3LYP and CASSCF(6,6) level of theory. The source patterns from these two methods are completely different. According to B3LYP functional, the Cu' nucleus itself is the major contributor of the spin density in that point yielding a β effect ($\approx$ 80%), helped by bridging N1 and central N2 azido atoms which give total contributions equal to $\approx$ 7-8% and $\approx$ 4-5%, respectively. On the contrary, the other Cu source, the terminal azido N3 and the pyridinic N atoms oppose to the previous sources, producing an α effect. Magnetic and relaxation contributions are both important in this case, with the latter being the dominant source for the Cu' atom, where the magnetic term is equal to -27.0% and the relaxation one is 106.6%. CASSCF(6,6) instead predicts a more delocalized and cooperative situation. The Cu' source is still the major contribution on absolute value, but Cu and the two bridging N1 and N1' atoms yield comparable, on absolute values, contributions. Giving some numbers, the absolute source value from Cu' is -133.4%, while the contributions from the other cited atoms are $\approx$ -88% and $\approx$ 83%. respectively. Moreover, also the pyridine N atoms bonded to Cu' have a non negligible effect. It is worth noting that the Cu' source evaluated by the CASSCF method has an opposite sign with respect to the DFT prediction, that is CASSCF, contrary to DFT functional, predicts Cu

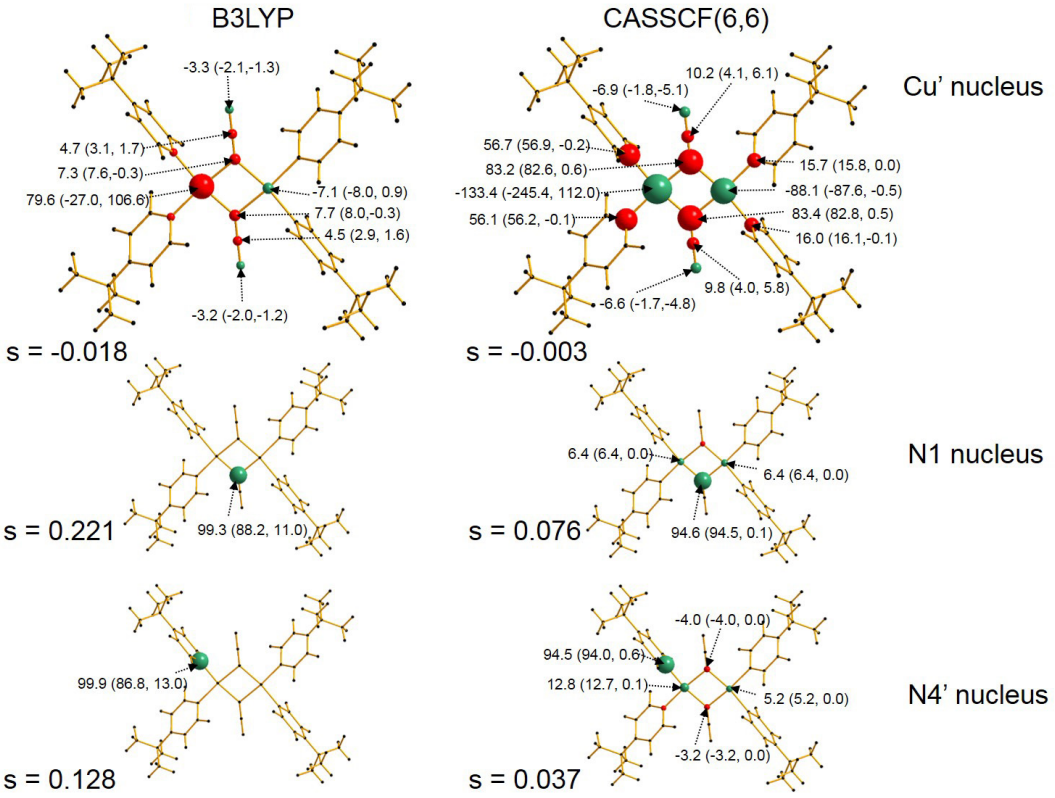


Figure 2.12: B3LYP and CASSCF(6,6) SF percentage contributions to the spin density $s(\boldsymbol{r})$ at three nuclear positions in the End-On complex: Cu', N1 and N4'. Numerical values represent the total(magnetic/relaxation) percentage contributions. Green/red spheres denote α/β effect on the spin density of the selected reference point.

atom to yield an α-effect. At CASSCF level the Cu' atom (together with the symmetric Cu atom) counteracts the negative spin density value at the selected reference point. On the contrary, the N sources co-operate to reduce the $s(\boldsymbol{r})$ value down to the final value.

At the N1 nucleus the spin density value is one order of magnitude larger than at the Cu' nucleus, resulting 0.221 *a.u.* and 0.076 *a.u.* at the B3LYP and CASSCF(6,6) level. The three times larger spin density value obtained with the B3LYP method probably results from the observed exaggeration of spin delocalization on ligands by this model. The B3LYP functional predicts a large source contribution from the N1 bridging atom (99.3%), while the two Cu atoms have a negligible effect on this reference point. The magnetic source is the major contribution (88.2%), but the relaxation term still has a non-zero weight equal to 11.1%. CASSCF(6,6) also predicts the N1 source as the main component (94.6%), but here the two Cu effects are not negligible, being both equal to 6.4%. Again, these outcomes demonstrate that the B3LYP model exaggerates the amount of spin delocalization on the ligands and underestimates spin connection or communication mechanism with respect to complete active space method.

Similar observation may be made for the reconstruction of the unpaired electron density at the N4' nucleus position. Spin density value is higher than for the Cu' nucleus and about

half of that found on N1 nucleus. At CASSCF(6,6) level, both Cu atoms (12.8% and 5.2% for Cu' and Cu, respectively) cooperate with N4' atom (94.5%) to reconstruct the spin density on the selected reference point. On the contrary, in the DFT prediction this mechanism is completely absent and the source contribution only comes from N4' itself that yields 99.9% of the total reconstructed density.

The source function patterns for the reference points along the Cu'-N4', chosen analogously to those along Cu'-N1, are shown in Figure 2.13. They are be discussed here because their results closely resemble those found along Cu'-N1.

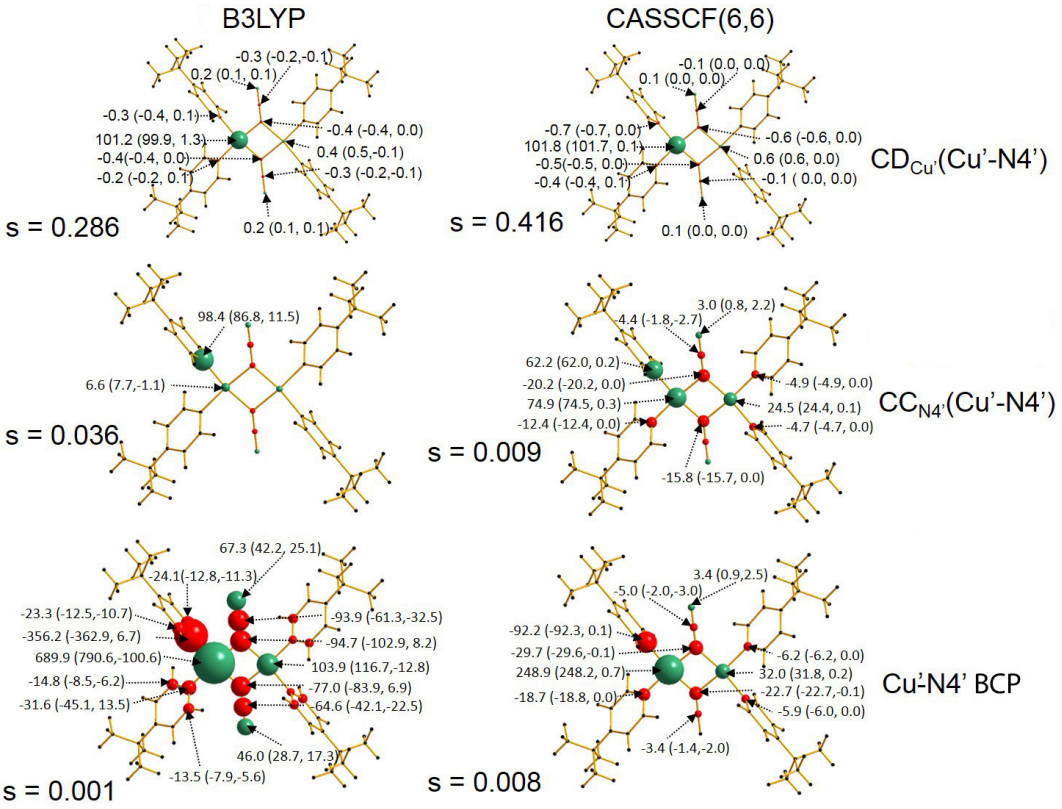


Figure 2.13: B3LYP and CASSCF(6,6) SF percentage contributions to the spin density $s(\boldsymbol{r})$ at three selected reference points along the Cu'-N4' bond in the End-On complex: (i) charge depletion (CD), (ii) charge concentration (CC) close to N4' nucleus, (iii) bond critical point (BCP). Numerical values represent the total(magnetic/relaxation) percentage contributions. Green/red spheres denote α/β effect on the spin density of the selected reference point.

2.3.6 Spin-Resolved Components vs Spin Densities

In their paper, Deutsch *et al.* analyse the spin-resolved ρ_α and ρ_β components of electron density of the End-to-End Cu azido complex, using a spin-split multipolar model to jointly refine X-rays and polarized neutron diffraction data.[82] They found a good agreement between the spin-resolved components and those estimated by B3LYP functional. One may wonder if the agreement between these two components is a sufficient condition to guarantee a similar agreement of their resulting spin density distribution. In this section

this statement is verified for the End-On Cu azido compound.

Figure 2.14 compares contour plots of ρ_α, $-\rho_\beta$ and s evaluated at B3LYP and CASSCF(6,6) level of theory on the same plane studied before, that is the least-square plane of the four ligand N atoms around the Cu atoms.

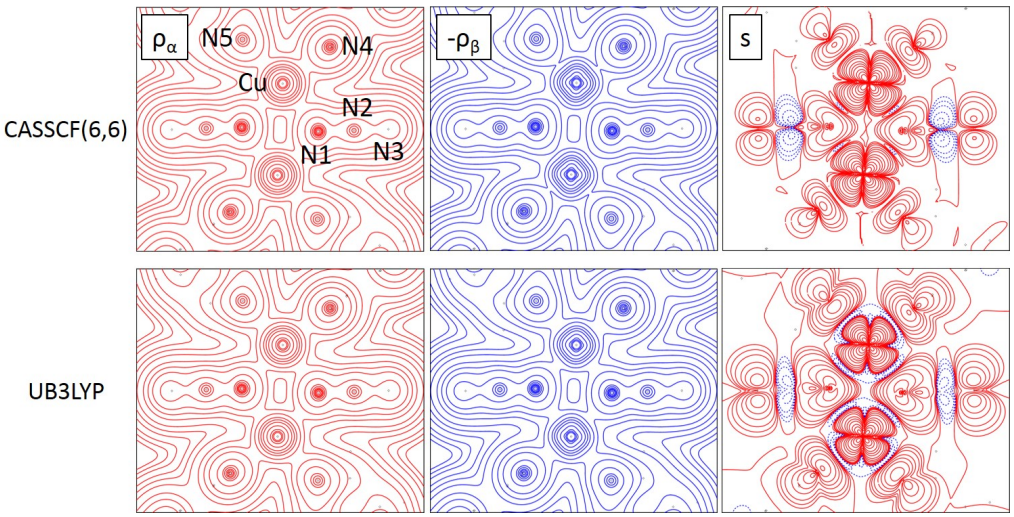


Figure 2.14: CASSCF(6,6) and B3LYP contour plots of spin resolved ρ_α, $-\rho_\beta$ and the spin density $s = \rho_\alpha - \rho_\beta$ in the End-On complex least-square plane of the four N atoms around each Cu atom. Color code and isocontours levels are the same of Figure 2.2.

The two spin resolved densities appear to be similarly described by the two methods employed. ρ_α has a circular shape around the paramagnetic centre, while ρ_β has a more squared shape due to the incomplete $d_{x^2\text{-}y^2}$ orbital. The five α electrons are able to fully fill the d α-spin orbitals, while the four β electrons can not do the same with the d β-spin ones. This asymmetric filling leads to the asphericity of the density and of its Laplacian around the Cu atom, as already discussed in the previous section.

Although the two spin-resolved densities are similar for both methods, this is not true for the spin density, as shown in Figure 2.14. The shape of the ρ_α and ρ_β distributions are defined by all the electrons, most of them described similarly in the two distributions as they formally fill mostly double occupied orbitals. On the contrary, the spin density singles out the difference between the distribution of the few (two in this case) unpaired electrons. It is then clear that this distribution is extremely more sensitive to fine details of the computational method than the simple α and β electron densities and then it is reasonable that different model approaches may significantly differ in the spin density description. Summarizing this section, a good agreement between the spin-resolved densities predicted from different methods does not always reflect in a qualitatively similar agreement for the related spin density.

2.4 End-to-End Cu Azido Complex

2.4.1 Effect of the Active Size on the CASSCF Calculations

DFT functionals are very often used to study molecules and transition metal complexes.[55] One of the main reason of their success is the low computational effort, which allows one to treat relatively large molecules with a good accuracy. Despite this large use, one of the main not tackled challenge for the DFT functionals is the treatment of open-shell systems. The reason of this resides, as already explained in Section 1.6, in the absence of any unpaired electrons term in the Hamiltonian. The most used exchange-correlation functionals, such as for example B3LYP, BLYP, BP86, PBE and M06, are not calibrated to correctly reproduce the SD of magnetic systems, and this leads to non-accurate spin density distributions.[55] *Ab-initio* electron correlation methods, such as the CASSCF, are then fundamental to introduce correlation correction and then recover accurate spin densities. Unfortunately, at least two *Achilles* heels affect the systematic use of this method. The first problem is the dimension of the active space used. In principle, the larger is the active space considered, the more accurate is the final result. This statement is valid only from a theoretical point of view though. Since multi-determinantal methods requires large amount of memory and high computational power to be performed, *ab initio* softwares are only able to manage a limited number of electrons and orbitals. The second critical point is the choice of the active orbitals. No real *a priori* criteria are available for the selection of proper active orbitals and quite always this choice is taken arbitrarily. Moreover, too large active spaces may generally lead to convergence failures and practical recipes to define their most suited size have been put forth.

As already discussed in Section 2.2, for the End-to-End two CASSCF calculations were performed: a CASSCF(6,6) and a CASSCF(10,10). The former uses an active space of 6 electrons distributed in 6 orbitals, resulting in 2 doubly occupied, 2 half-occupied and 2 virtual orbitals. The latter instead consider 10 electrons and 10 orbitals as active, that is 4 doubly occupied, 2 half-occupied and 4 virtual orbitals. The nature of the orbitals used is explained in Section 2.2. The outcomes of the two active-space adopted for the End-to-End system were analysed and compared with the unrestricted Hartree Fock approach (UHF, only HF hereinafter) to highlight the differences in the choice of active space.

HF calculations show a high spin contamination effect. Spin annihilation procedure was therefore necessary to recover the correct spin state (initial spin state: $<S^2> = 2.0946$, spin-annihilated final state $<S^2> = 2.0045$). According with the variational principle, the energy gain obtained moving from HF to multi-configurational methods increases with the active space size. CASSCF(6,6) method stabilizes the system by -0.0575 *a.u.* with respect to HF, and the enlargement of the active space adds an extra-stabilization contribution of -0.0448 *a.u.* (-0.1023 *a.u.* with respect to HF). Energy data of these wavefunctions and the DFT ones are reported in Table 2.5.

This large energy gain moving from the CASSCF(6,6) to CASSCF(10,10) implies that a reasonable convergence can not be ensured for the smaller active space, even though

Table 2.5: Number of configurations and energy values, expressed in Hartree, of the End-to-End Cu azide complex evaluated for all the levels of theory adopted.

Method	Configurations	Energy / Eh
HF	1	-5267.02726069
CASSCF(6,6)	225	-5267.08474188
CASSCF(10,10)	44100	-5267.12957238
B3LYP	1	-5281.35400306
BLYP	1	-5280.92465233

full convergence is not guaranteed at the CASSCF(10,10) level. We selected the magnetic orbitals accordingly to what written in Section 2.2. The magnetic orbitals found for the CASSCF(10,10) method are almost equal to those found at the CASSCF(6,6) level of theory and have occupation values very close to one, nominally 1.032 and 0.977 e.

Table 2.6 compares the QTAIM electron net charges q(Ω) and the electron spin populations SP(Ω) of some relevant atoms, evaluated at the HF and CASSCF levels of theory.

Table 2.6: QTAIM net charges (q) and spin populations (SP) for selected atoms (Ω) at three levels of theory: HF, CASSCF(6,6) and CASSCF(10,10)

	q(Ω)			SP(Ω)		
Basin Ω	HF	CAS(6,6)	CAS(10,10)	HF	CAS(6,6)	CAS(10,10)
Cu	1.459	1.463	1.460	0.891	0.909	0.821
O1	-1.329	-1.330	-1.342	0.021	0.020	0.019
N1	-1.440	-1.441	-1.362	0.033	0.020	0.057
N2	-1.209	-1.210	-1.212	0.028	0.021	0.019
N3	-0.218	-0.204	-0.196	-0.086	0.002	0.023
N4	-0.201	-0.248	-0.259	0.009	0.003	0.003
N5	-0.401	-0.370	-0.357	0.109	0.017	0.034
$\sum_{\text{azide}}$[a]	-0.820	-0.822	-0.812	0.032	0.022	0.060
$\sum'$[a]	-3.339	-3.339	-3.269	0.114	0.083	0.155

[a] $\sum_{\text{azide}}$: sum of q or SP values over the azido atoms (N3, N4 and N5). $\sum'$: sum of q or SP values over the selected atoms, except Cu.

Table 2.6 clearly shows that atomic net charges are more stable against the change of level of theory than the spin populations. The Cu net charge predicted by the CASSCF(10,10) method is as high as 1.460 e and it remains almost unchanged when CASSCF(6,6) and HF is considered (1.463 e and 1.459 e, respectively). On the contrary, the spin population (SP) found on the same atom by the CASSCF method with the largest active space, 0.821 e, undergoes a non-negligible variation, being equal to 0.891 e for the HF level of theory and 0.909 e in the case of CASSCF(6,6). The different impact of the method on net charges and spin populations is in some way intrinsic in the definition of the spin density field. For instance, lets consider an hypothetical case where, for some reason, the α component of the electron density increases and the β one decreases by the same magnitude. The

value of the total electron density, evaluated as the sum of the two components, remains unchanged following these variations while the spin density, defined as the difference of the α and β components, undergoes a large variation. If something similar happens, this may be a symptom of the shortcomings of the wavefunctions used. Looking at the spin populations, the greatest changes occur on the atoms belonging to the azido moiety, namely N3, N4 and N5. The total spin density population of the azido moiety ranges from 0.022 e (CASSCF(6,6)) to 0.060 e (CASSCF(10,10)). What is more interesting, however, is the variation of the single atomic contributions within the group.

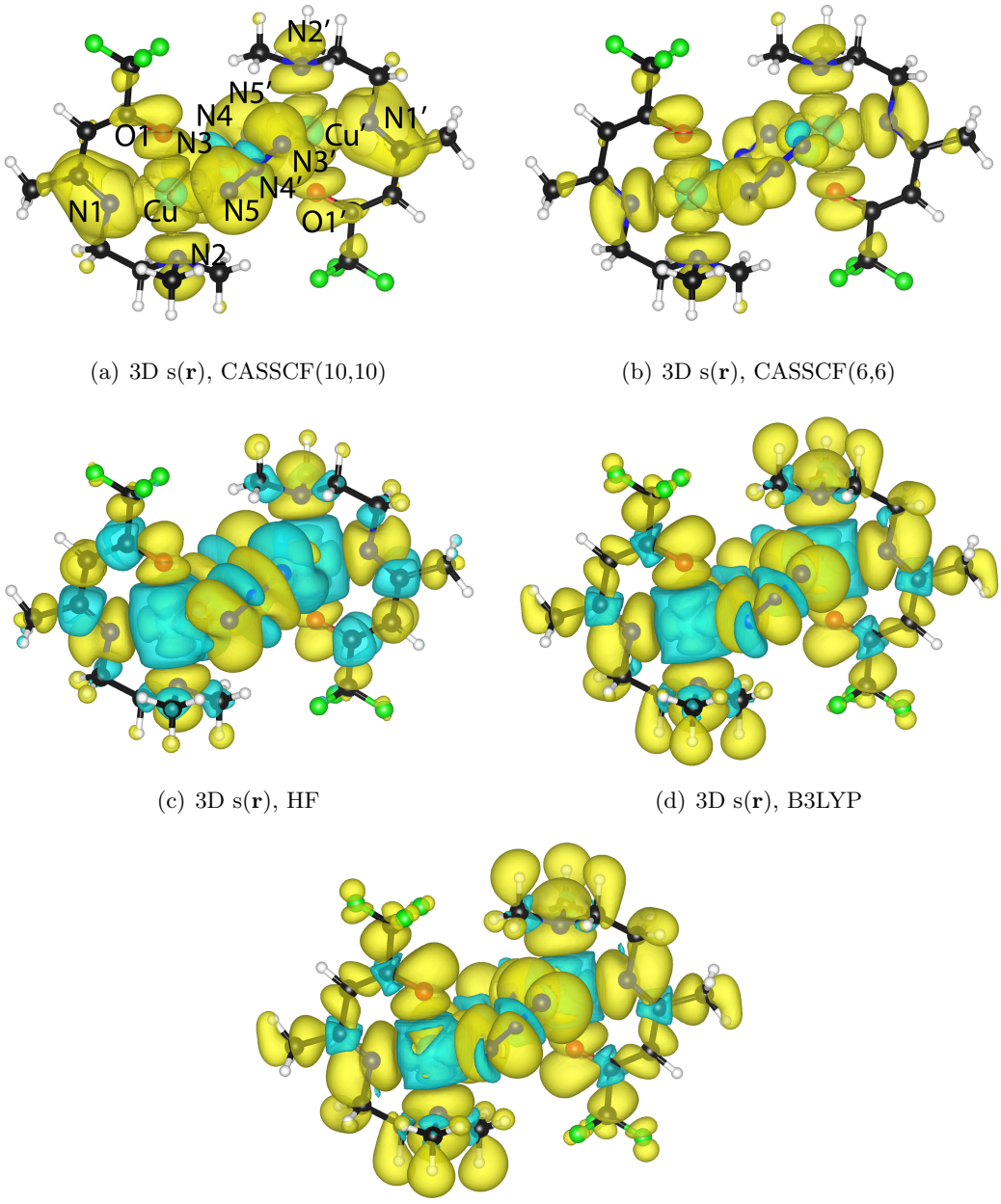
HF level of theory predicts a high positive spin population on the short-bonded N5 atom (0.109 e), while the long-bonded N3 nitrogen bears a large and negative population (-0.086 e). The N4 central atom is predicted to have also in this case a very low spin population, equal to 0.009 e. In this case the HF predicts a sort of spin polarization mechanism within the azido ligand.

The CASSCF(6,6) model smoothes this spin population distribution. In this case, the largest contribution comes from the N5 atom and is equal to 0.17 e, while the N3 and N4 populations are almost close to zero but still positive. CASSCF(6,6) does not predict any spin polarization mechanism along the N_3^- moiety. It is worth noting that although the total populations located on the azido moiety predicted by the HF and CASSCF(6,6) methods are very close to each other, they are an expression of two different behaviours. For the former method, the total population is the result of the sum of large and opposite populations, while for the latter it is due to quite marginal values.

When the active space is expanded to (10,10), the total α delocalization toward the azido atoms increases. All the spin populations of these atoms remains positive but larger in magnitude than for the CASSCF(6,6). The same trend can be seen by taking into account the numbers reported for $\sum'$ in Table 2.6, where the sum of the atomic spin populations for the atoms listed in Table 2.6 (Cu excluded) raises from 0.083 e in CASSCF(6,6) to 0.155 e in the CASSCF(10,10) model. This variation may be probably connected with the addition in the active space of π/π^* natural orbitals centred on the L ligands, which returns a more flexible model and allows a greater α-spin delocalization towards the ligands.

Figure 2.15 shows the 3D spin density isovalue surfaces obtained by CASSCF and HF calculations, together with the same quantities predicted by the two DFT functional used.

All the levels of theory here used retrieve the same $d_{x^2-y^2}$ spin distribution around the Cu atoms. HF predicts extended negative regions, depicted in light blue, around the metallic Cu centres and on the azido moieties, forming in this latter case a clear spin polarization pattern. Small but not negligible negative spin density regions can also be found on the carbon atoms of the L ligands. These features are completely or partially missing in the CASSCF calculations. In the CASSCF(6,6) isosurfaces map, the positive and negative spin density regions on the azido ligands annihilate each other, leading to a contraction of the positive spin envelope located around the N5 and to a more smooth spin distribution. CASSCF(10,10) and CASSCF(6,6) qualitatively agree on the predicted



(a) 3D s(**r**), CASSCF(10,10)

(b) 3D s(**r**), CASSCF(6,6)

(c) 3D s(**r**), HF

(d) 3D s(**r**), B3LYP

(e) 3D s(**r**), BLYP

Figure 2.15: three-dimensional spin density isovalue surfaces plots for all the used methods: CASSCF(10,10), CASSCF(6,6), HF, B3LYP and BLYP. Positive values are reported in yellow, negative values in light blue. Isovalue levels are fixed at $1 \cdot 10^{-4}$ *a.u.*.

spin distribution, but they differ on the N3 and N5 isosurfaces and on the small negative region found on the central N4 atoms. Moreover, the N1 atom in the CASSCF(10,10) method shows a more delocalized spin density isosurface.

The two DFT functionals yield qualitatively similar 3D spin density distributions, but with large differences with respect to CASSCF(10,10) result. As for the HF case in Figure 2.15, DFT functionals predict a negative region, more expanded than in the former case, around the Cu atom. They also lead to a spin polarization within the azido bridges, with negative spin density zones located on the short-bonded N5 and N4 atoms. CASSCF(10,10) and DFT calculations agree in predicting positive spin density regions on the two terminal N3 and N5 azido atoms, but with different shapes and orientation. In B3LYP and BLYP approach, these positive regions assume a well defined π-like shape, oriented in the azido atoms least-square plane, while in the CASSCF(10,10) N5 shows a more spherical distribution and N3 has a less defined π-like, oriented perpendicular to the azido atoms plane.

From now on, only the CASSCF(10,10) and DFT functionals results are used and compared for the analysis of the spin density distribution of this End-to-End azido complex.

2.4.2 Population Analysis

Similarly to what was found for the EO complex in Section 2.3 and what shown in the previous section, spin populations are strongly dependent on the computational method employed and also on the nature of the exchange-correlation functional.[55,69] This difference between the DFT functionals is intrinsic in the DFT theory as we said before, since both the electronic energy and the observables are defined only by the ED distribution.[11] Again, it is notable to stress that the spin density distribution is in principle not needed to define the exchange-correlation functional but is a pure additional variable which we do not have exploited in the DFT Hamiltonians we considered. Classic DFT functionals, like those we used, fail to correctly reproduce the unpaired electron distribution.[55,68,69]

The spin density populations calculated using the QTAIM and the Mulliken's partitioning, plus the data obtained by Aronica *et al.*[88] are reported in Table 2.7.

Using the QTAIM partitioning, CASSCF(10,10) estimates that $\approx$ 82 % of the total 2 e resides on the Cu basins, $\approx$ 6% is located on the azido bridge and the heteroatoms of the ligand L bear $\approx$ 10% of the total population. The Mulliken's populations are close to the QTAIM ones, predicting $\approx$ 84%, 6% and 8% of the spin population to be located on the Cu atoms, the azide groups and the ligands L, respectively. These outcomes are in agreement with those obtained by Aronica *et al.* using a DDCI-3 approach[101,102] ($\approx$ 81%, $\approx$ 6% and $\approx$ 8%, respectively).[88] DFT predictions are instead completely different. B3LYP QTAIM populations indicate that only $\approx$ 64% of the unpaired electrons reside on the Cu atoms, while on the N_3^- bridge and on the ligands they are $\approx$ 11% and $\approx$ 21%, respectively. Similar results are found using the Mulliken's partitioning ($\approx$ 65%, $\approx$ 11% and $\approx$ 21%, respectively) The Cu spin population found by hybrid B3LYP functional is decisively lower

Table 2.7: Spin populations (SP) for selected atoms or group of atoms Ω in the End-to-End azido Cu complex. SP(Ω): QTAIM partitioning; $SP_{Mull}(\Omega)$ and $SP_{Mull,Ar}(\Omega)$: Mulliken's partitioning for our wavefunctions and for the Aronica *et al.*[88]; $SP_{PND}(\Omega)$: polarized neutron diffraction estimate

Basin	SP(Ω)[a]	$SP_{Mul}(\Omega)$[a]	$SP_{Mul,Ar}(\Omega)$[b-c]	$SP_{PND}(\Omega)$[c]
Cu	0.821/0.640/0.528	0.839/0.650/0.534	0.806/0.571	0.719(6)
O1	0.019/0.067/0.083	0.013/0.065/0.082	0.044/0.077	0.043(5)
N1	0.057/0.059/0.078	0.053/0.069/0.079	0.028/0.100	0.044(5)
N2	0.019/0.078/0.089	0.015/0.077/0.090	0.021/0.113	0.076(6)
N3	0.023/0.065/0.091	0.022/0.073/0.100	0.040/0.113	0.033(7)
N4	0.003/0.005/0.010	0.005/-0.001/0.002	0.004/-0.022	0.004(5)
N5	0.034/0.042/0.066	0.031/0.038/0.064	0.039/0.044	0.029(6)
$\sum_{azide}$[d]	0.060/0.112/0.167	0.058/0.110/0.166	0.083/0.135	0.066(18)
$\sum_L$	0.095/0.204/0.250	0.081/0.211/0.251	0.093/0.290	0.163(16)

[a] CASSCF(10,10)/B3LYP/BLYP data.
[b] DDCI-3/DFT population.
[c] Data taken from Table 5 of Reference [88].
[d] $\sum_{azide}$: sum of SP values over the azido atoms (N3, N4 and N5). $\sum_L$: sum of SP values over the heteroatoms belonging to ligand L (O1, N1 and N2).

than that predicted by CASSCF(10,10), while on the contrary the amount of unpaired electrons on the ligands is strongly increased. This behaviour confirms, as for the EO case, that DFT functionals tend to overestimate spin delocalization towards the non-magnetic atoms. A larger discrepancy occurs when BLYP results are considered. In this case, QTAIM Cu spin population decreases to ≈ 53% and the population on the azido bridge and on the hetoroatoms of the L ligand increases to ≈ 17% and ≈ 25%, respectively. At least for this compound, pure BLYP exchange-correlation functional has a higher tendency to delocalize the two unpaired electrons than hybrid B3LYP. These outcomes are opposite to what found on Fe(II) nitrosyl complex by Boguslawski *et al.*[55], where pure DFT functionals have smaller differences, on average, when compared to CASSCF results than hybrid ones.[55] These evidences seem to suggest that no clear and general rules about the accuracy of DFT functionals in predicting the spin density distribution, according to the nature of the exchange-correlation Hamiltonian, can be stated before further investigations case by case.

Polarized neutron diffraction populations, obtained by Aronica refining 125 magnetic structures, are also in this case intermediate between the CASSCF(10,10) and the B3LYP estimates.[88] Cu population (≈ 72%) seems to be closer to the B3LYP outcomes, but on the other hand N_3^- and L heteroatoms have populations closer to the CASSCF results (≈ 7% and ≈ 16%, respectively). Deutsch *et al.* found similar values (Cu population ≈ 74 % of the total 2 electrons) refining the spin-split multipolar model against X-ray and PND data, using 252 magnetic structure factors in the refinement (data not reported here).[82]

All the three levels of computation reported in Table 2.7 qualitatively agree on the spin population magnitude of the central nitrogen atom of the azido moiety (N4), that is very close to zero as found in the EO complex (see Table 2.2). On the contrary, they disagree

on the relative weight of the terminal azido N3 and N5 atoms. CASSCF(10,10) predicts N5 to have a higher population than N3, equal to ≈ 3.4% and ≈ 2.3%, respectively. On the other hand, B3LYP and BLYP reveal an opposite trend, resulting ≈ 4.2% (N5) and ≈ 6.5% (N3) the former and ≈ 6.6% (N5) and ≈ 9.1% (N3) the latter.

A deeper insight on the nature of remarkable differences between the CASSCF and DFT in retrieving the spin population can be obtained by an *in tandem* exam of their electron and spin density properties. In fact, one may wonder if this enhanced spin delocalization for the DFT functionals is only a consequence of the excessive electron delocalization or not.[103]

As shown in Table 2.8, the Cu net charges calculated by DFT methods are significantly lower than the one retrieved by CASSCF(10,10) method. B3LYP and BLYP functionals predict the Cu net charge to be equal to +1.135 *e* and +1.042 *e*, respectively, while CASSCF(10,10) finds a charge q = +1.460 *e*. These results demonstrate that more electron population lies in the Cu DFT basin than in the CASSCF one.

Table 2.8: QTAIM atomic charges (q) for selected atoms or group of atoms Ω in the End-On azido Cu complex at all the used levels of theory

Basin	UHF	CASSCF(6,6)	CASSCF(10,10)	B3LYP	BLYP
Cu	1.459	1.463	1.460	1.135	1.042
O1	-1.329	-1.330	-1.342	-1.07	-1.006
N1	-1.440	-1.441	-1.362	-1.079	-1.001
N2	-1.209	-0.210	-1.212	-0.890	-0.820
N3	-0.218	-0.204	-0.196	-0.227	-0.222
N4	-0.201	-0.248	-0.259	-0.134	-0.122
N5	-0.401	-0.370	-0.357	-0.352	-0.328
$\sum_{azide}$ [a]	-0.820	-0.822	-0.812	-0.713	-0.672
$\sum$ [a]	-3.339	-3.339	-3.268	-3.042	-2.826

[a] $\sum_{\text{azide}}$: sum of q values over the azido atoms (N3, N4 and N5). $\sum$: sum of q values over the selected atoms, Cu included.

Moreover, the average distance between the Cu nucleus and its bond critical points is larger for the DFT level of theory than for the CASSCF. B3LYP shows an increment equal to +0.05 *a.u.* with respect to CASSCF value, while the corresponding enhancement for the BLYP case is +0.07 *a.u.*. As a consequence of this, the DFT functionals also show a Cu volume (evaluated using ρ cutoff equal to 10^{-3} *a.u.*) increase, marking a +7.6 *a.u.* and +10.1 *a.u.* for B3LYP and BLYP, respectively, when compared to the multiconfigurational method. All these changes are a consequence of the exaggerated electron delocalization inherent to DFT.

To help to disentangle the electron density effects from the spin density ones, we evaluated atomic net charges and spin populations for DFT wavefunctions using CASSCF(10,10) QTAIM surfaces. All the numerical results are reported in Table 2.9 and have to be compared with results reported in Table 2.8.

When the CASSCF boundaries are considered, the DFT Cu charge becomes higher

Table 2.9: Net charges (q), spin populations (SP) and their difference (Δ) with respect to the same physical quantity evaluated with the proper boundaries and reported in Tables 2.7 and 2.8 in the End-to-End complex for DFT functionals evaluated using the CASSCF(10,10) QTAIM atomic surfaces.

Basin	B3LYP				BLYP			
Ω	q(Ω)	Δq(Ω)	SP(Ω)	ΔSP(Ω)	q(Ω)	Δq(Ω)	SP(Ω)	ΔSP(Ω)
Cu	1.279	0.144	0.640	0.000	1.233	0.191	0.528	0.000
O1	-1.221	-0.147	0.067	0.000	-1.190	-0.184	0.084	0.001
N1	-1.331	-0.252	0.069	0.000	-1.309	-0.308	0.078	0.000
N2	-1.128	-0.238	0.078	0.000	-1.105	-0.285	0.089	0.000
N3	-0.177	0.050	0.065	0.000	-0.164	0.058	0.090	-0.001
N4	-0.257	-0.123	0.005	0.000	-0.269	-0.147	0.011	0.001
N5	-0.324	0.028	0.042	0.000	-0.300	0.028	0.066	0.000
$\sum_{azide}$ [a]	-0.758	-0.045	0.112	0.000	-0.733	-0.061	0.167	0.000
$\sum_{L}$ [a]	-3.680	-0.637	0.214	0.000	-3.604	-0.777	0.251	0.001

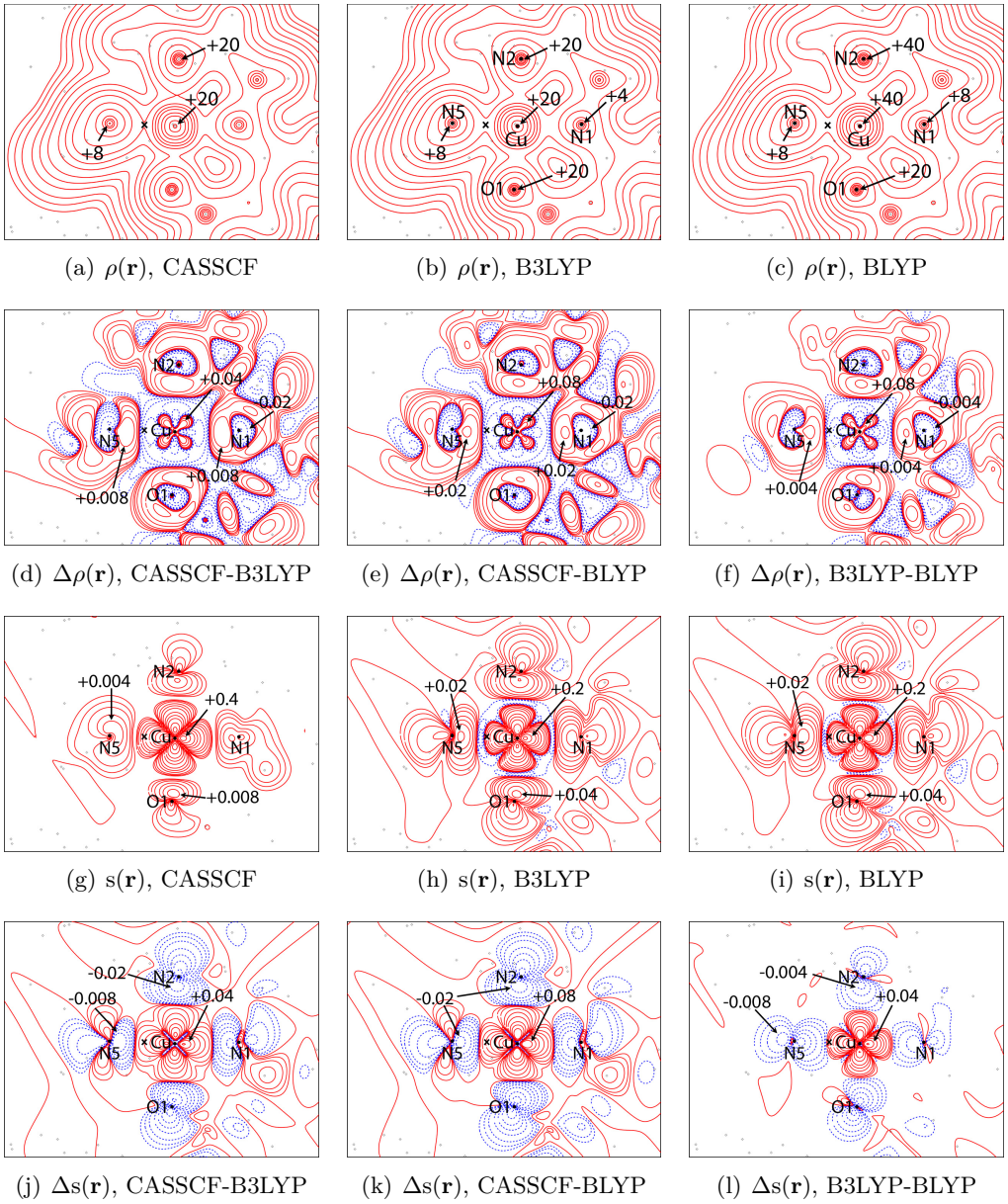
[a] Sum of q, SP or their variations Δ values over the azido atoms N3, N4 and N5 ($\sum_{azide}$) and over the ligand heteroatoms O1, N1 and N2 ($\sum_{L}$).

while the same quantity decreases for the ligand heteroatoms (O1, N1 and N2). Azide atoms behave differently one to each other: N3 and N5 charges increase when the CASSCF surface is used, while N4 decreases. Overall, the net charge on the azide group slightly decreases. Spin population remains almost constant if we consider DFT or CASSCF(10,10) Bader's surfaces. This lack of variation may be due either to a compensation of α and β contributions effect or to the inclusion of a region where the spin density distribution is close to zero.

Spin density contour maps, evaluated in the least-square plane of the square-pyramid base (Cu, N3 and the L heteroatoms O1, N1 and N2) and drawn in Figure 2.17, show that B3LYP and BLYP predict a negative region located around the Cu positive d-shape spin distribution. The excess of delocalization could be responsible for the creation of this polarization mechanism between the paramagnetic and non-paramagnetic centres. Related to this, the DFT approach also enhances covalency of Cu-ligands and consequently reduces the charge transfer.

To get further details in this decomposition, CASSCF-DFT difference maps of electron and spin density distributions in the same plane considered in the previous paragraph were studied (see Figure 2.16).

Both B3LYP and BLYP functionals depict Cu-N and Cu-O bonds as more shared and covalent when compared to CASSCF. This is evident looking at the shape of isodensity contours of ρ around the Cu-N5 BCP. Same results can be obtained looking at numbers in Table 2.8, where one can see that DFT net charges are closer to zero than for the CASSCF method for all the reported atoms but the N3, the azido atom involved in the Cu-N long bond with the formally double occupied d_{z^2} orbital. Indeed, analysing the electron density difference maps, one can notice that CASSCF method predicts less density from the Cu



(a) $\rho(\mathbf{r})$, CASSCF (b) $\rho(\mathbf{r})$, B3LYP (c) $\rho(\mathbf{r})$, BLYP

(d) $\Delta\rho(\mathbf{r})$, CASSCF-B3LYP (e) $\Delta\rho(\mathbf{r})$, CASSCF-BLYP (f) $\Delta\rho(\mathbf{r})$, B3LYP-BLYP

(g) s($\mathbf{r}$), CASSCF (h) s($\mathbf{r}$), B3LYP (i) s($\mathbf{r}$), BLYP

(j) Δs($\mathbf{r}$), CASSCF-B3LYP (k) Δs($\mathbf{r}$), CASSCF-BLYP (l) Δs($\mathbf{r}$), B3LYP-BLYP

Figure 2.16: CASSCF(10,10), B3LYP and BLYP contour plots of electron density ($\rho(\mathbf{r})$, first row), spin density (s($\mathbf{r}$), third row) for the End-to-End complex evaluated in the least-squares plane of the Cu-O (O1) and of the three shorter Cu-N bonds (N1, N2 and N5). Second and fourth row panels show the difference maps of electron and spin density maps relative to CASSCF or between the two adopted DFT functionals. The solid black X indicates the BCP along the Cu-N5 bond and it refers to the portrayed density for the density maps, while it refers to the CASSCF(10,10) wavefunction for the difference maps. Colour code and isosurface levels are the same of Figure 2.2.

nucleus to the Cu-ligands BCPs (reported along the Cu-N5 bond as a black dot) than the DFT functional. The opposite occurs in the region between these BCPs and the ligand nuclei. Moreover, some compensating mechanisms are evident close to the ligand nuclei and in the d_{xy} direction for the Cu nucleus. These compensating regions have an opposite charge transfer if compared to that observed along the Cu-ligand bonds.

The spin density difference maps (Figure 2.16) shows that the reported increase of the covalency of the metal-ligands interactions is accompanied by a decrease of the spin density on the Cu and an increase on the ligands. In particular, the CASSCF spin density distribution is higher in the region between the Cu nucleus and the Cu-ligands BCPs than DFT ones, while it is lower between these BCPs and the ligand nuclei.

From the results shown in Tables 2.7-2.8 and in Figure 2.16, it is clear that the lower electron population and higher spin population on Cu in CASSCF(10,10) with respect to DFT functionals, are related to a preferential decrease of its β over its α density component. On the contrary, for the ligand atoms we found an increase of the electron population on going from CASSCF to DFTs, meaning that in DFT ρ_α is preferentially delocalized to the ligands over ρ_β, which also accounts of the DFT higher spin population on the ligands. These outcomes are confirmed by the data reported in Table 2.10, where the α and β populations were evaluated accordingly to Bader's partitioning.

Table 2.10: α and β populations (N_α and N_β) for the CASSCF(10,10), B3LYP and BLYP wavefunctions for some atoms (Ω). Variations (ΔN_α and ΔN_β) between CASSCF(10,10) and DFT outcomes are also reported

	CASSCF(10,10)		B3LYP		BLYP		CASSCF-DFT	
Ω	N_α	N_β	N_α	N_β	N_α	N_β	ΔN_α[a]	ΔN_β[a]
Cu	14.18	13.36	14.25	13.61	14.24	13.71	-0.07/-0.06	-0.27/-0.35
O1	4.68	4.66	4.57	4.50	4.54	4.46	0.11/0.14	0.16/0.20
N1	4.21	4.15	4.07	4.00	4.04	3.96	0.14/0.17	0.15/0.19
N2	4.12	4.10	3.98	3.91	3.95	3.87	0.14/0.19	0.17/0.23
N5	3.70	3.66	3.70	3.66	3.70	3.63	0.00/0.00	0.00/0.03

[a] B3LYP/BLYP results. ΔN_α and ΔN_β are the difference between CASSCF and DFT populations.

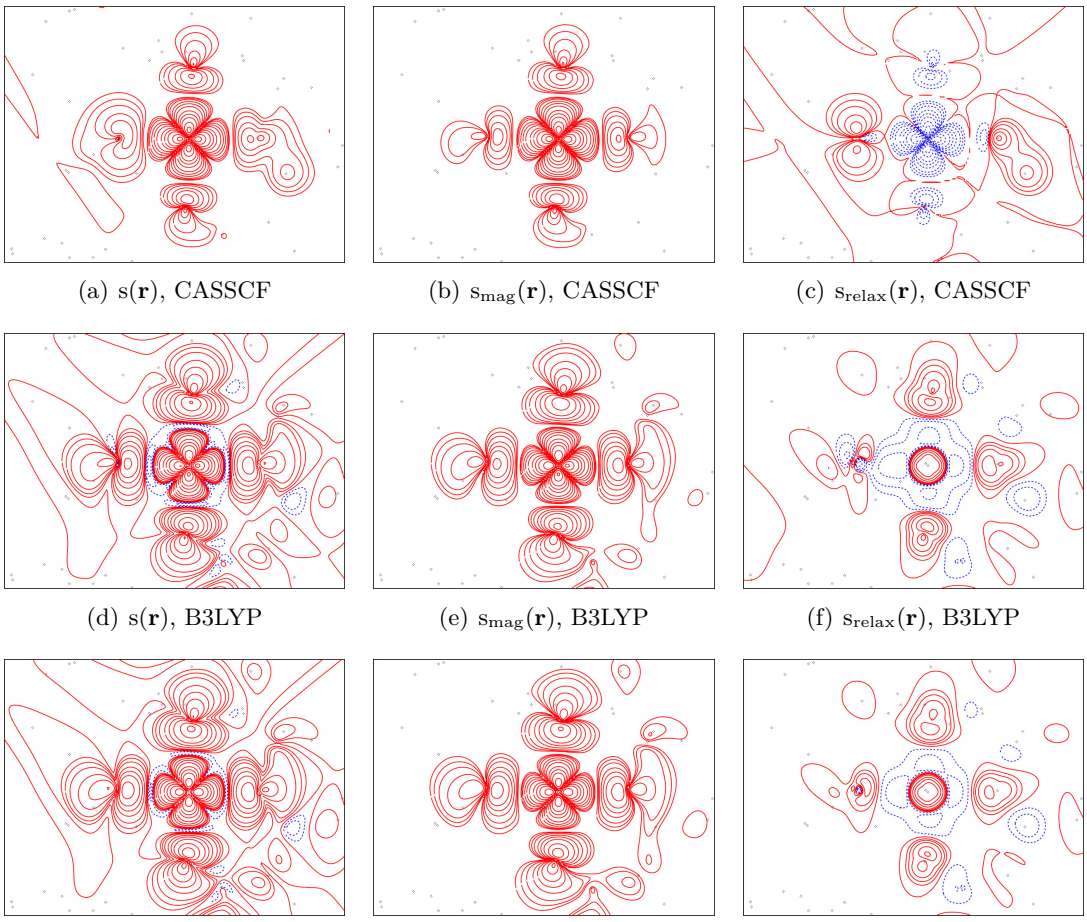
Both Cu ΔN_α and ΔN_β are negative (-0.07/-0.06 and -0.27/-0.35, respectively), reflecting the higher electron density population (or lower net charge) predicted by DFT. The drop of the β component for the Cu atom is effectively higher, in absolute value, than the α one, confirming that there is preferential decrease of β electrons on going from DFT to CASSCF level of theory. For the ligand atoms ΔN_β increase more than ΔN_α, except for N5 where the difference is negligible. Although the effect is less pronounced for the ligand atoms than for the Cu, because of the compensating effects shown before, the global trend is in agreement with what stated before.

In conclusion, it has been shown that the wrong tendency of DFT to highly delocalize the electron density, causing a decreased charge separation between Cu and the ligands, also implies a different redistribution of ρ_α and ρ_β. This latter events bring to an enhanced

spin delocalization in DFT when compared to complete active space results, *phenomenon* already largely discussed in literature.[102,104]

2.4.3 Total, Magnetic and Relaxation Spin Density Maps

Spin density contour maps and their magnetic and relaxation components have been evaluated in the least-squares plane of the Cu-O (O1) and of the three shorter Cu-N bonds (N1, N2 and N5) (Figure 2.17) and in the N_3^- least-squares plane (Figure 2.18).



(a) s(**r**), CASSCF (b) s_{mag}(**r**), CASSCF (c) s_{relax}(**r**), CASSCF

(d) s(**r**), B3LYP (e) s_{mag}(**r**), B3LYP (f) s_{relax}(**r**), B3LYP

(g) s(**r**), BLYP (h) s_{mag}(**r**), BLYP (i) s_{relax}(**r**), BLYP

Figure 2.17: CASSCF(10,10), B3LYP and BLYP spin density maps (s(**r**), first panel) and their magnetic (s_{mag}(**r**), second panel) and relaxation (s_{relax}(**r**), third panel) contributions for the End-to-End complex evaluated in the least-squares plane of the Cu-O (O1) and of the three shorter Cu-N bonds (N1, N2 and N5). Color code and isosurface levels are the same of Figure 2.2.

The spin density distribution in the pyramidal square base is similar to what was found for the End-On azido complex reported in Section 2.3.2 for all the used methods. The total spin density distribution around the Cu atoms displays the same $d_{x^2-y^2}$ shape, where the four positive lobes are directed towards the atoms N1, N2, N5 and O1, which are the among belonging to the base of the square-pyramid. The apical N3 atom is above the reported plane and directed towards the double-occupied d_{z^2} orbital. The two adopted

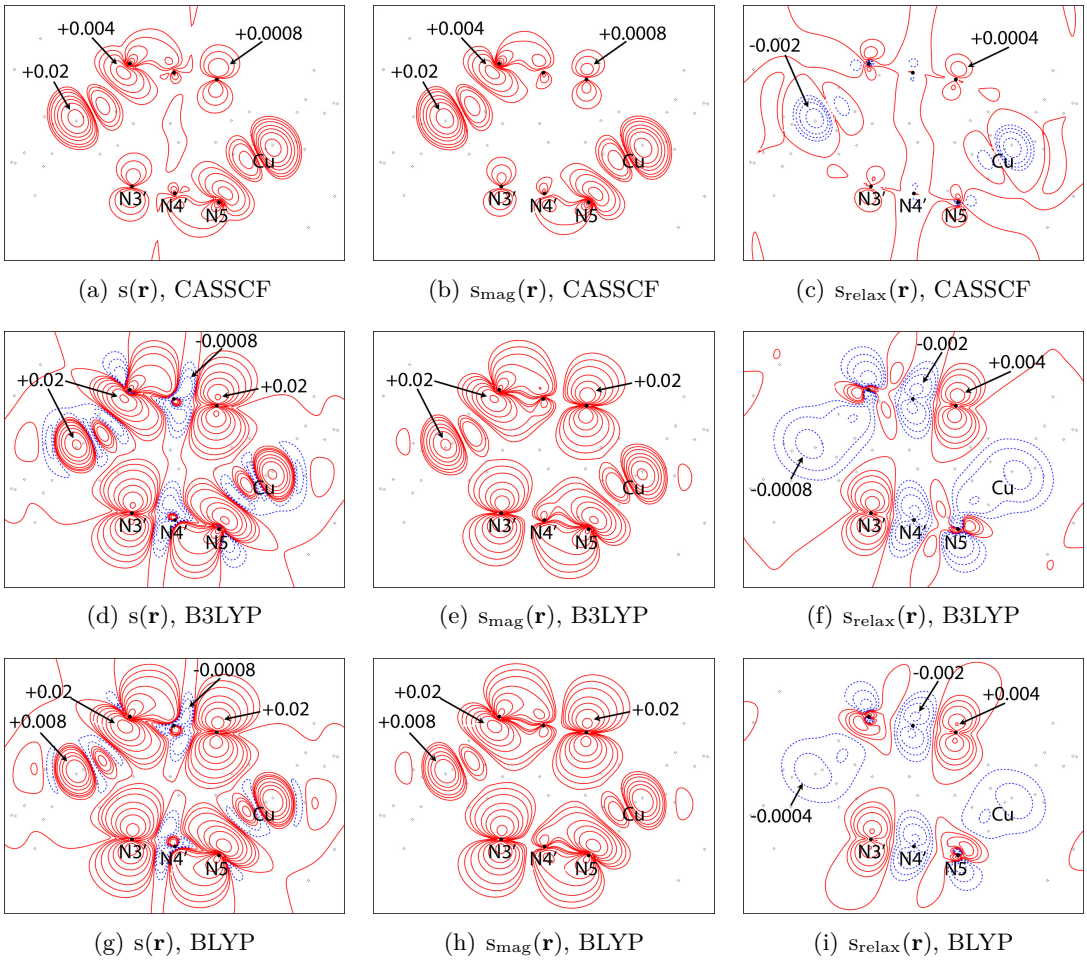
DFT functionals, the hybrid B3LYP and the pure BLYP, predict similar spin densities in this plane. A negative region is found between the $d_{x^2-y^2}$ distribution around the Cu nucleus and the ligand atoms. On the contrary, CASSCF(10,10) does not predict this region and no polarization mechanism from the metallic centre toward the ligands atoms is found. These outcomes perfectly fit what found in the EO complex study case, where similar features were recovered.

Comparison of the highest contour isosurfaces values reveals that, as already stated in Section 2.3.2 for the EO system, the DFT method exaggerates spin delocalization. B3LYP and BLYP functionals both find maximum peaks of +0.2 e for the $d_{x^2-y^2}$ distribution and +0.02 e for the spin density located around the coordinating atoms. The same quantities for the CASSCF(10,10) level of theory are +0.4 e and +0.004 e, respectively. These results are in agreement with the spin populations reported in Table 2.7.

The decomposition of the total spin density into magnetic and relaxation terms substantiate that the former is the dominant contribution. CASSCF(10,10) and both DFT exchange-correlation functionals retrieve qualitatively similar magnetic spin densities but they disagree on the spin concentration/delocalization degree. The magnetic terms retrieve the main features, that is the Cu $d_{x^2-y^2}$ angular shape and the main distributions centred on the ligands heteroatoms. On the other hand, the relaxation terms are extremely sensitive to the method used. As in the case of the End-On complex, DFT calculations display a positive, spherically shaped relaxation spin density centred on the Cu nucleus surrounded by a diffuse negative region. Weak and positive spin distributions are also present on the ligands atoms. On the contrary, the CASSCF(10,10) spin relaxation density shows a $d_{x^2-y^2}$ distribution around the paramagnetic centre as for the magnetic term but with opposite sign and lower absolute value. This feature overlaps completely the magnetic one, leading just to a small decrease of the total spin density and not to the appearance of the negative spin density region found in the DFT maps. A π-like distribution is retrieved on the azidic N5 atom and another relevant distribution is located on the N1 atom. These distributions have a strong role in the shape definition of the total spin density around the ligands atoms.

The spin density contours evaluated in the azido least-square plane is shown in Figure 2.18.

The two DFT functionals lead to a spin polarization mechanism through the azido bridges, with positive spin density regions around the N3 and N5 atoms and a negative region around the N4 atom. Again, this feature is similar to what observed in the azido moiety of the End-On system shown in Figure 2.2, where a polarization mechanism occurs within the azido moiety. In particular, it is worth noting that N3 and N5 atoms have different unpaired electrons distributions. Spin density around the N5 and N5' atoms have a T-like shape as found for the ligand N atoms in the EO complex (Figure 2.2). On the contrary, spin density distribution located on the N3 atom in the End-to-End system has a π-like shape and it is similar to what found for the terminal N3 in the End-On azido (Figure 2.2). This fact suggests that the Cu and N3 atoms hardly share any



(a) s(**r**), CASSCF (b) s_{mag}(**r**), CASSCF (c) s_{relax}(**r**), CASSCF

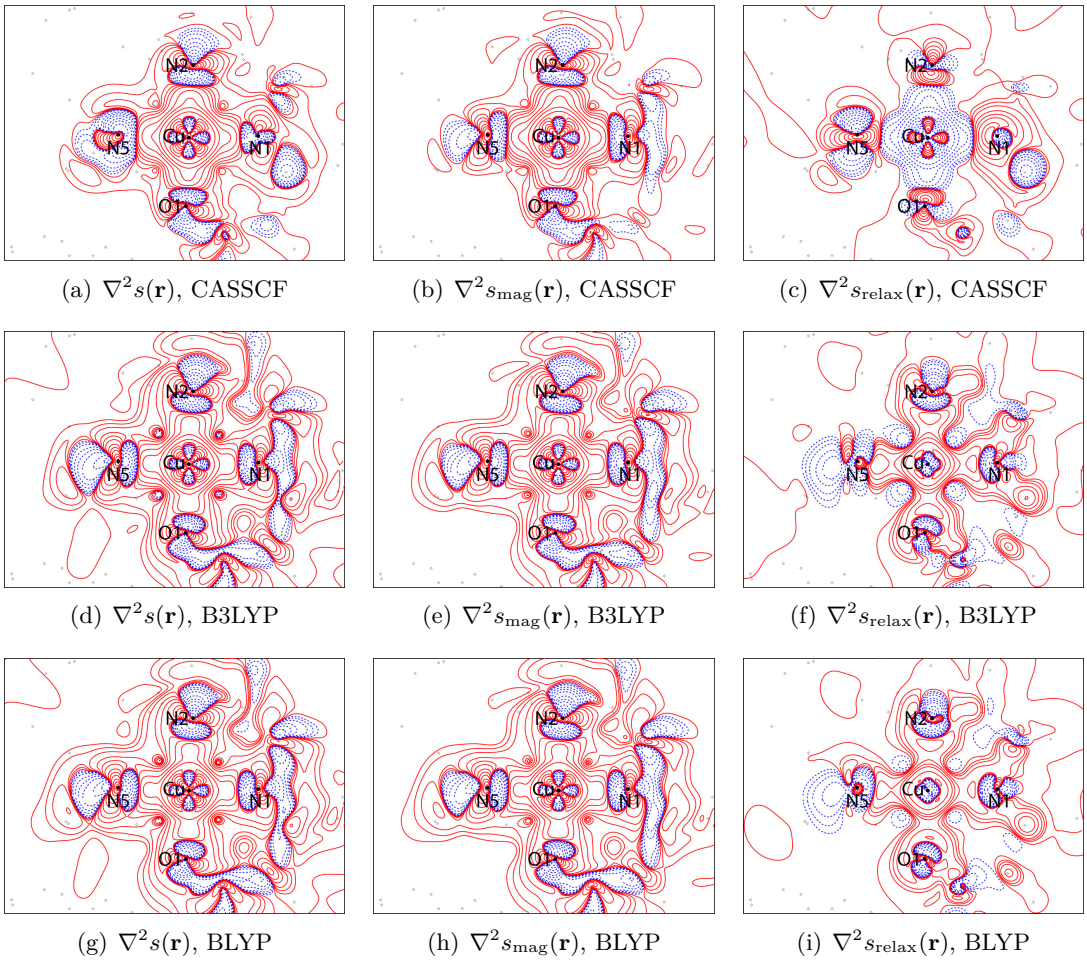
(d) s(**r**), B3LYP (e) s_{mag}(**r**), B3LYP (f) s_{relax}(**r**), B3LYP

(g) s(**r**), BLYP (h) s_{mag}(**r**), BLYP (i) s_{relax}(**r**), BLYP

Figure 2.18: CASSCF(10,10), B3LYP and BLYP spin density maps (s(**r**), first panel) and their magnetic (s_{mag}(**r**), second panel) and relaxation (s_{relax}(**r**), third panel) contributions for the End-to-End complex evaluated in the least-square plane of the azido moieties. Color code and isosurface levels are the same of Figure 2.2.

significant spin information. The magnetic contributions (second column panels in Figure 2.18) are responsible for the positive parts of the total spin density and represent the most important spin density terms since they are able to reconstruct the main features of the total distribution. The relaxation components have lower absolute values and they are responsible for the negative region around the N4 atom. Contrary to DFT functionals, CASSCF(10,10) does not depict any polarization mechanism nor negative spin density regions within the azido bridge (see first panel in Figure 2.18). The N3 and N5 distributions are qualitatively similar to what already found for the B3LYP and BLYP methods, but they differ from a quantitative point of view. Maxima peak values in the CASSCF case on the azido N atoms are lower than for DFT, confirming once more the higher spin delocalization predicted by the latter method. The maxima values obtained is +0.004 *a.u.* around the N5/N5' atoms, while in the DFT cases the values obtained are one order of magnitude higher and equal to +0.02 *a.u.*. Also in this case, the magnetic term is the most relevant part while the relaxation contribution is almost negligible.

Figure 2.19 shows the spin density Laplacian distribution in the least-square plane of the Cu-O and of the three shorter Cu-N bonds.



(a) $\nabla^2 s(\mathbf{r})$, CASSCF (b) $\nabla^2 s_{\mathrm{mag}}(\mathbf{r})$, CASSCF (c) $\nabla^2 s_{\mathrm{relax}}(\mathbf{r})$, CASSCF

(d) $\nabla^2 s(\mathbf{r})$, B3LYP (e) $\nabla^2 s_{\mathrm{mag}}(\mathbf{r})$, B3LYP (f) $\nabla^2 s_{\mathrm{relax}}(\mathbf{r})$, B3LYP

(g) $\nabla^2 s(\mathbf{r})$, BLYP (h) $\nabla^2 s_{\mathrm{mag}}(\mathbf{r})$, BLYP (i) $\nabla^2 s_{\mathrm{relax}}(\mathbf{r})$, BLYP

Figure 2.19: CASSCF(10,10), B3LYP and BLYP spin density Laplacian maps ($\nabla^2 s(\mathbf{r})$, first panel) and their magnetic ($\nabla^2 s_{\mathrm{mag}}(\mathbf{r})$, second panel) and relaxation ($\nabla^2 s_{\mathrm{relax}}(\mathbf{r})$, third panel) contributions for the End-to-End complex evaluated in the least-squares plane of the Cu-O (O1) and of the three shorter Cu-N bonds (N1, N2 and N5). Color code and isosurface levels are the same of Figure 2.2.

The general discussion is similar to what already said in Section 2.3.2 for the End-On complex. The large positive $d_{x^2\text{-}y^2}$ distribution found around the Cu atom in the spin density maps depicted in Figure 2.17 is due to the concentrated d-like negative lobes found in the spin density Laplacian maps by all the methods used. The most evident difference among the multi-reference and the DFT maps is the spin density Laplacian distribution around the azido N5 atom. CASSCF(10,10) predicts the N5 nucleus to reside on an island of positive density completely surrounded by a pool of SD negative Laplacian distribution. DFT functionals also predict the N5 nucleus in a positive region, but in this case the negative SD Laplacian regions are divided in a triangular-like shaped region and an elliptical one.

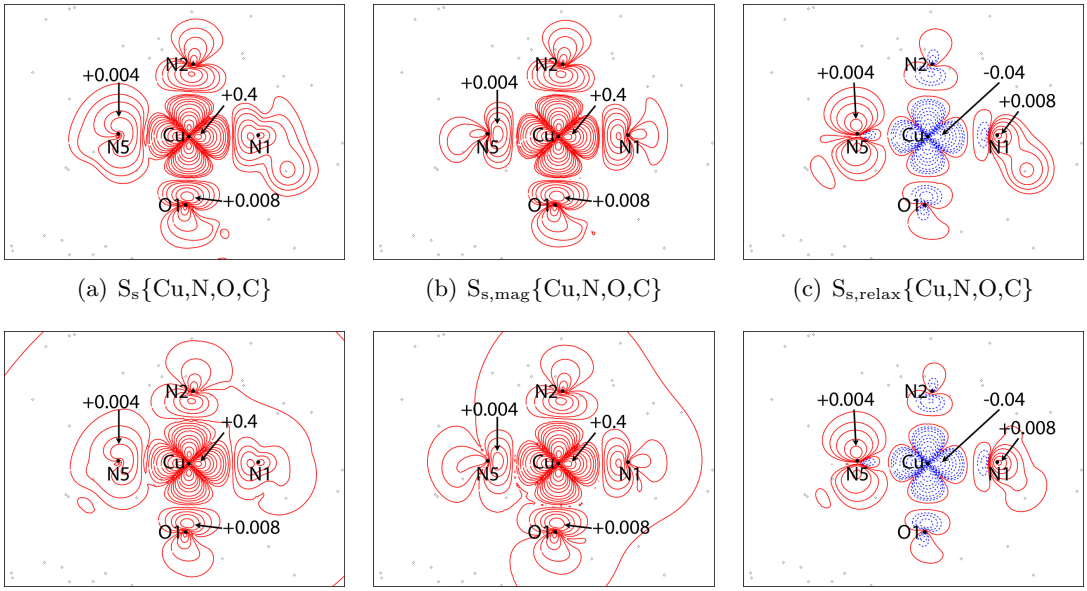
The decomposition of the Laplacian in its magnetic and relaxation terms clearly ex-

plains the different distribution found at the CASSCF level of theory. All the three panels (second column in Figure 2.19) are qualitatively similar and differ only for the spatial extension of the positive/negative areas. The relaxation parts of the SD Laplacian distribution are found to be completely different among CASSCF and DFT methods. What is worthy to be noted here is that the relaxation part in the multi-configurational method has a strong role in defining the shape of the SD Laplacian distribution on N5 and the other ligand atoms. This effect and, in general, the effect of the relaxation part is less evident in the DFT maps.

2.4.4 Source Function Partial Density Reconstruction

The source function spin density partial reconstruction is useful, as in the case of the End-On system, to have an easy to grasp picture on the *cause-effect* relationship between the atomic basin and the reconstructed density.

Figures 2.20 to 2.22 shows the partial reconstructed maps in the pyramidal plane obtained using {Cu,N,O,C} and {Cu,N,O} subsets, where the syntax here used is the same as explained in Sections 1.9 and 2.3.3 and refers to all the atoms of these elements in the system.

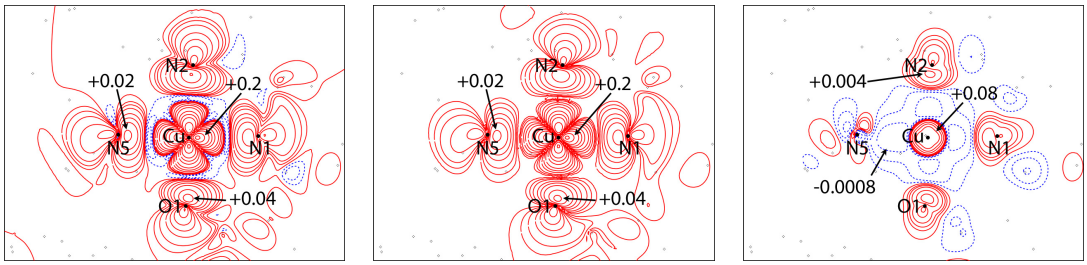


(a) S_s{Cu,N,O,C} (b) $S_{s,mag}${Cu,N,O,C} (c) $S_{s,relax}${Cu,N,O,C}

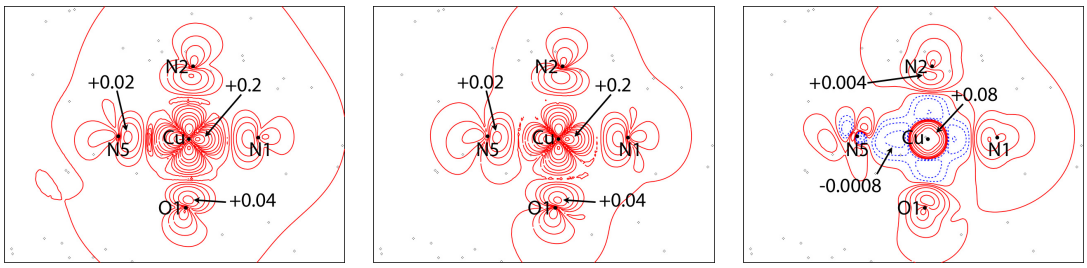
(d) S_s{Cu,N,O} (e) $S_{s,mag}${Cu,N,O} (f) $S_{s,relax}${Cu,N,O}

Figure 2.20: CASSCF(10,10) source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-to-End complex evaluated in the least-squares plane of the Cu-O (O1) and of the three shorter Cu-N bonds (N1, N2 and N5). Subsets containing Cu, N, O and C atoms ({Cu,N,O,C}) and without the C contributions ({Cu,N,O}) have been considered. Colour code and isosurface levels are the same of Figure 2.2.

When the source contributions from the copper, nitrogen, oxygen and carbon atoms are considered (Cu,N,O,C subset), the reconstructed density is almost indistinguishable from



(a) S_s\{Cu,N,O,C\} (b) $S_{s,mag}$\{Cu,N,O,C\} (c) $S_{s,relax}$\{Cu,N,O,C\}

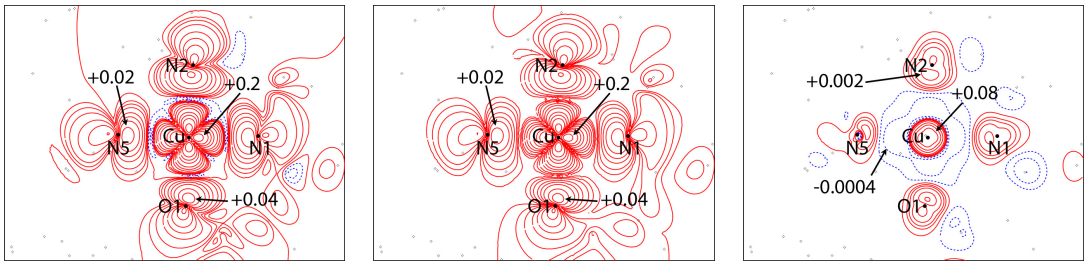


(d) S_s\{Cu,N,O\} (e) $S_{s,mag}$\{Cu,N,O\} (f) $S_{s,relax}$\{Cu,N,O\}

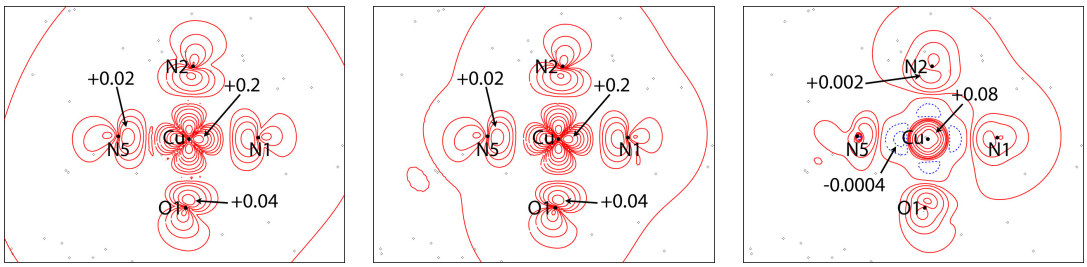
Figure 2.21: B3LYP source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-to-End complex evaluated in the least-squares plane of the Cu-O (O1) and of the three shorter Cu-N bonds (N1, N2 and N5). Subsets containing Cu, N, O and C atoms ({Cu,N,O,C}) and without the C contributions ({Cu,N,O}) have been considered. Colour code and isosurface levels are the same of Figure 2.2.

the real one (compare Figures 2.20 to 2.22 with Figure 2.17) both for the CASSCF(10,10) and DFT levels of theory. Some small differences occur close to the borders of the picture due to numerical finite accuracy of the integration. When the contributions from the C atoms are excluded from the reconstruction ({Cu,N,O} subset, second rows of Figures 2.20 to 2.22), few remarkable differences appear. As for the EO system, the negative regions around the Cu nucleus disappear/contract for both DFT functionals. This fact remarks once more that the source function descriptor is able to disclose the nature of specific features and the effect on them from the various atoms in the system. In addition to such difference, relative to CASSCF, the positive spin density regions located close to the ligands atoms undergo a strong contraction, even if the values of the maxima remain almost the same. These effects can be explained looking at the magnetic and relaxation densities. The negative region around the Cu in the relaxation map undergoes a strong contraction (more evident in the BLYP functional than in the B3LYP) and at the same time there is a similar effect on the positive region around the ligands in the magnetic map. On the contrary, CASSCF(10,10) partial reconstructed maps are only lightly affected by the removal of C contributions, showing a small contraction of the positive region around the Cu nucleus in the total maps and some other small differences in the magnetic and relaxation contributions.

The source contribution coming from the Cu (S_s\{Cu\}) shows the usual $d_{x^2-y^2}$ distribution around the nucleus.



(a) $S_s\{Cu,N,O,C\}$ (b) $S_{s,mag}\{Cu,N,O,C\}$ (c) $S_{s,relax}\{Cu,N,O,C\}$



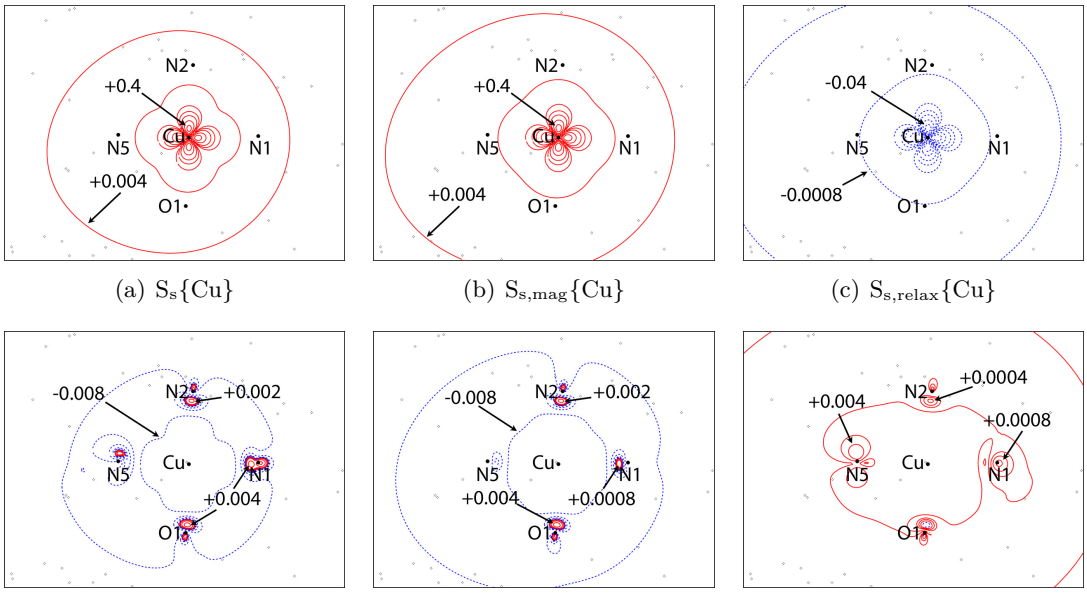
(d) $S_s\{Cu,N,O\}$ (e) $S_{s,mag}\{Cu,N,O\}$ (f) $S_{s,relax}\{Cu,N,O\}$

Figure 2.22: BLYP source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-to-End complex evaluated in the least-squares plane of the Cu-O (O1) and of the three shorter Cu-N bonds (N1, N2 and N5). Subsets containing Cu, N, O and C atoms ({Cu,N,O,C}) and without the C contributions ({Cu,N,O}) have been considered. Colour code and isosurface levels are the same of Figure 2.2.

The CASSCF Cu source contribution is higher in magnitude and more contracted with respect to BLYP, which is in turn more contracted than the B3LYP functional. These results comply well with the higher spin delocalization trend of the DFT functional when compared to multi-reference methods. All the reported methods agree that the $S_{s,mag}\{Cu\}$ is the dominant contribution and they all found qualitatively similar sources. As already seen in the total spin density maps, the relaxation terms are instead quite different from each other. CASSCF method predicts a relaxation term qualitatively similar to the magnetic distribution but with an opposite β effect on the total density. On the other hand, the DFT functionals retrieve a positive spherical relaxation distribution centred on the Cu nucleus and a diffuse β effect region around it.

The source contributions from the four coordinating atoms belonging to the pyramid square base (namely N1, N2, O1 and N5) are similar to those found for the End-On Cu complex when the subset composed by all the N atoms was considered (Figures 2.6 and 2.7). CASSCF(10,10) $S_s\{N,O\}$ has strongly contracted positive α effect regions located around the ligand nuclei, surrounded by a diffuse β effect region. DFT functionals, instead, predict a more expanded positive T-like shaped region because of the already discussed exaggeration of α-spin electron delocalization. In particular, the BLYP exchange-correlation functional shows more expanded regions of spin density sources than the B3LYP model.

This contrasting behaviour between the two DFT functionals for the {Cu} and {N,O} subsets almost cancel out when summed up, leading to similar $S_s\{Cu,N,O\}$ contributions



(a) $S_s\{Cu\}$ (b) $S_{s,mag}\{Cu\}$ (c) $S_{s,relax}\{Cu\}$

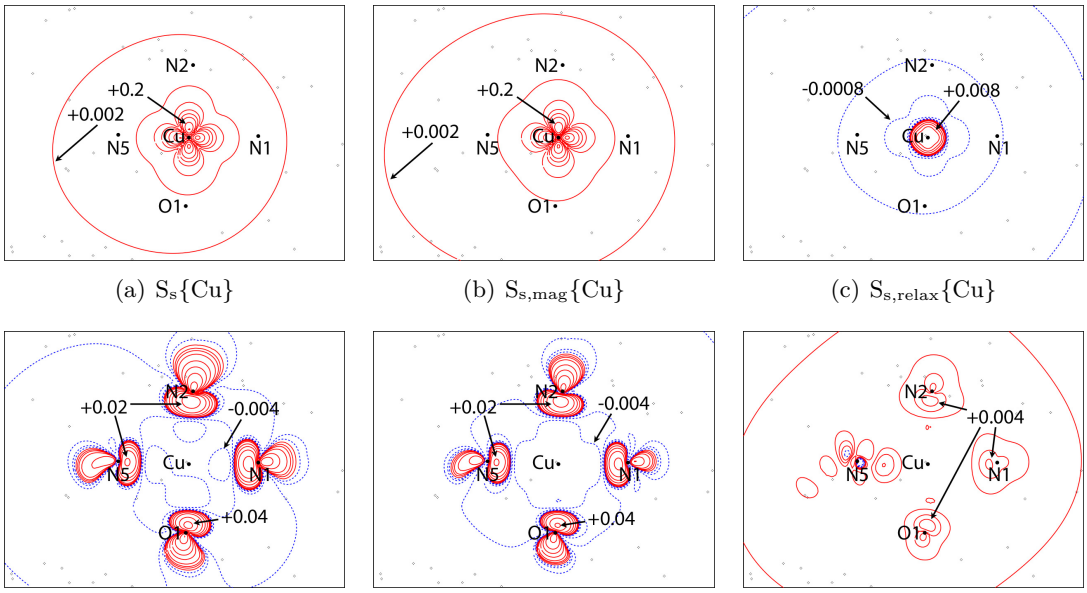
(d) $S_s\{N,O\}$ (e) $S_{s,mag}\{N,O\}$ (f) $S_{s,relax}\{N,O\}$

Figure 2.23: CASSCF(10,10) source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-to-End complex evaluated in the least-squares plane of the Cu-O (O1) and of the three shorter Cu-N bonds (N1, N2 and N5). Subsets containing Cu atoms ({Cu}) and heteroatoms ({N,O}) have been considered. Colour code and isosurface levels are the same of Figure 2.2.

(Figures 2.20 to 2.22). It is worth noting that the large positive spin density regions located around the azido N5 atom have different natures. In the case of CASSCF(10,10), these regions are dominated by the sources coming from the Cu atom, while in DFT cases the major contributions come from the N5 sources. This difference highlight that the two methods predict a different degree of spin comunication among atoms. In the case of the CASSCF method, the spin density distribution around the Cu atom and the ligand atoms is an effect of the co-operation between different atomic sources and sinks, while in the case of DFT the shapes of the spin density around the atoms are mainly due to the atoms themselves.

The same source atomic contributions decomposition has been performed in the azido least-square plane. $S_s\{Cu,N,O,C\}$, $S_s\{Cu,N,O\}$ and $S_s\{Cu\}$ contributions are reported in Figures 2.26 (CASSCF), 2.27 (B3LYP) and 2.28 (BLYP), together with their magnetic and relaxation terms.

Similar conclusions as above are found when the source from all the atoms (except the H atoms) is considered. The removal of the C source has a similar effect for the DFT functionals, where the negative region located on the central N4 is reduced or eliminated due to the contraction of the magnetic and relaxation contributions. The effects on the CASSCF(10,10) reconstructed density are lower and mainly focused on the contraction of the positive spin density areas. All the three methods predict qualitatively similar positive and very delocalized $S_s\{Cu\}$ contributions, dominated by the magnetic parts and



(a) $S_s\{Cu\}$ (b) $S_{s,mag}\{Cu\}$ (c) $S_{s,relax}\{Cu\}$

(d) $S_s\{N,O\}$ (e) $S_{s,mag}\{N,O\}$ (f) $S_{s,relax}\{N,O\}$

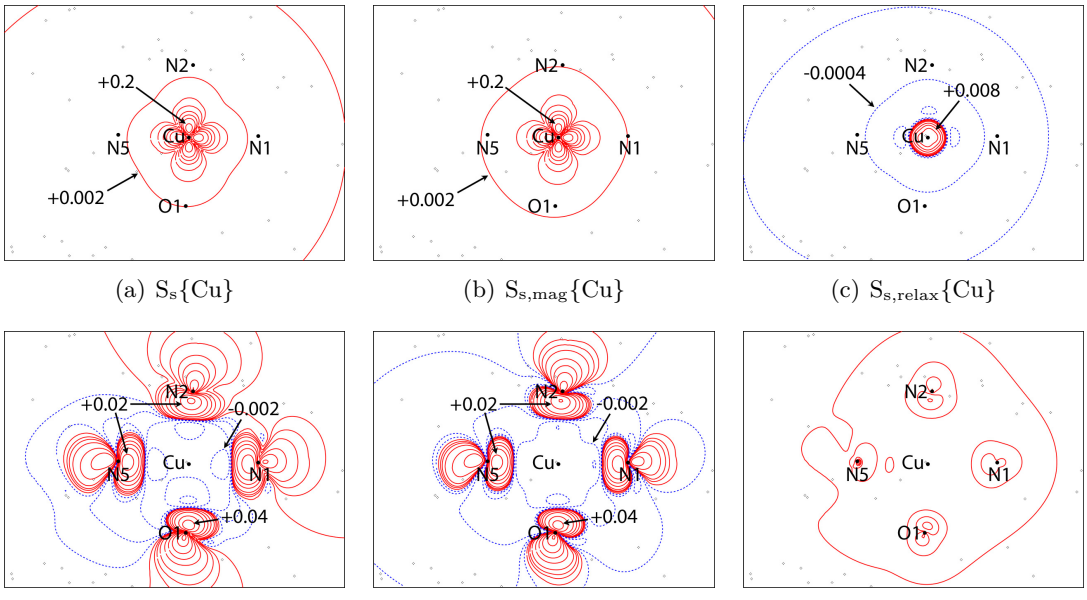
Figure 2.24: B3LYP source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-to-End complex evaluated in the least-squares plane of the Cu-O (O1) and of the three shorter Cu-N bonds (N1, N2 and N5). Subsets containing Cu atoms ({Cu}) and heteroatoms ({N,O}) have been considered. Colour code and isosurface levels are the same of Figure 2.2.

with opposing relaxation terms.

More interesting results come from the comparison and the decomposition of the $\{N_{azido}\}$ source contributions reported in Figures 2.29 (CASSCF), 2.30 (B3LYP) and 2.31 (BLYP).

The CASSCF(10,10) method predicts highly contracted positive spin densities on terminal N3 and N5 atoms and an overall diffuse negative spin density. The magnetic term mainly reconstructs the spin density around the N5 and N4 atoms, while the relaxation term has a large influence in determining the region around N3. The decomposition of $S_s\{N_{azido}\}$ into $S_s\{N5\}$, $S_s\{N4\}$ and $S_s\{N3\}$ (Figures 2.29 to 2.31) highlights that the positive N5 source function contribution is mainly located on itself and it is directed towards the Cu atom, with a diffuse β effect spread all over the space. On average, the N5 atom behaves similarly to the N1 in the End-On complex, with the difference in the shape of the positive spin density originated by the different coordination of the two atoms (Figures 2.29 to 2.31 for the EE complex and Figures 2.8 and 2.9 for the EO complex). Comparing the corresponding pictures, one can notice that the N3 atom behaves instead as the terminal N3 in the EO complex, with an α effect everywhere. N4 source opposes to these contributions producing β effect everywhere in the space. This holds true for all the methods used, even if large differences on the degree of spin density delocalization are evident between CASSCF and DFT methods.

As a conclusion of this section, we have demonstrated that the source function partial



(a) $S_s\{Cu\}$ (b) $S_{s,mag}\{Cu\}$ (c) $S_{s,relax}\{Cu\}$

(d) $S_s\{N,O\}$ (e) $S_{s,mag}\{N,O\}$ (f) $S_{s,relax}\{N,O\}$

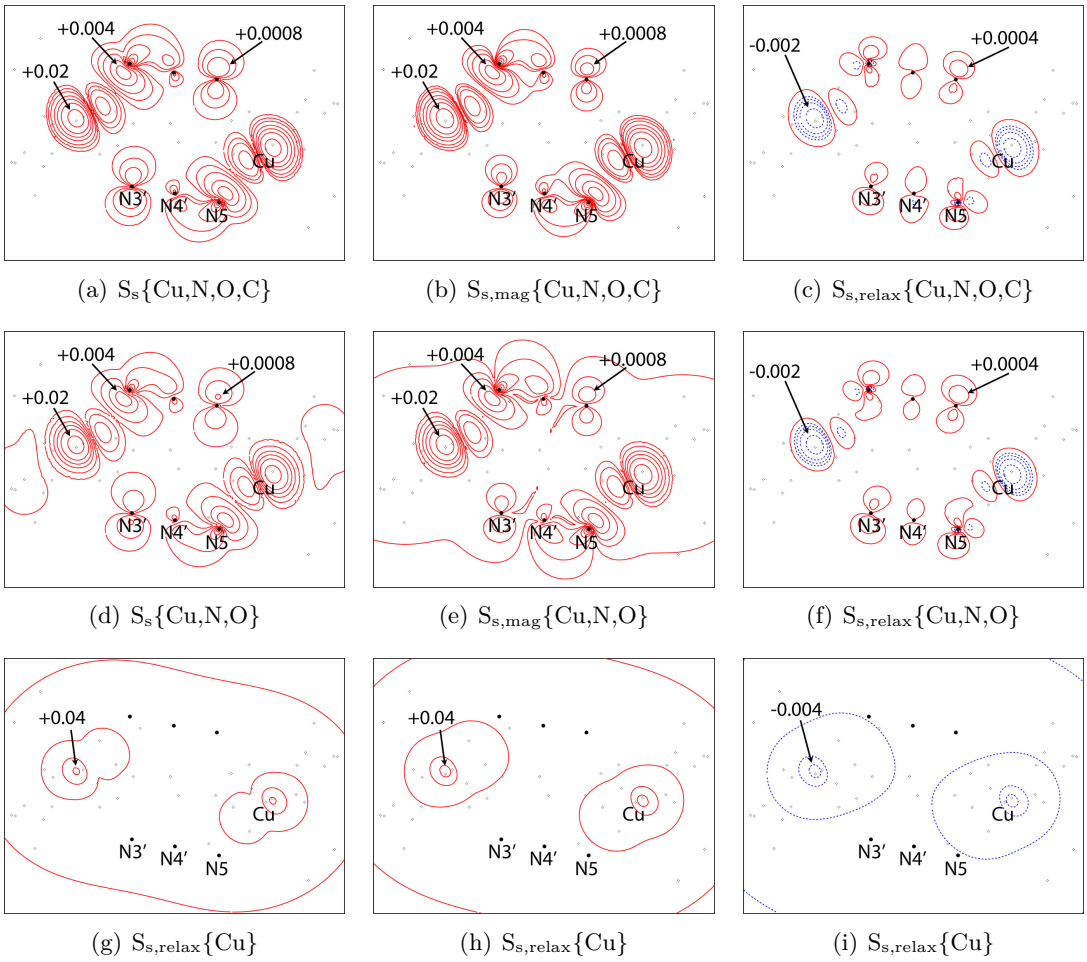
Figure 2.25: BLYP source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-to-End complex evaluated in the least-squares plane of the Cu-O (O1) and of the three shorter Cu-N bonds (N1, N2 and N5). Subsets containing Cu atoms ({Cu}) and heteroatoms ({N,O}) have been considered. Colour code and isosurface levels are the same of Figure 2.2.

spin density reconstruction is a valid tool to highlight the nature of the different behaviour between post-HF methods, such as the CASSCF, and the DFT methods. The former predict localized spin density distributions, with large amount of the unpaired electrons localized on the Cu atoms and close to the ligand nuclei. On the other hand, the two DFT functionals overdelocalize the spin density, producing spread partial densities. It is worth to highlight once more that, although the DFT methods yield a more delocalized scenario, on average the communication among the atoms is less evident than for the CASSCF outcomes. This fact clarifies how the degree of localization/delocalization of a density does not strictly correlate with the degree of communication among the atoms.

2.4.5 Cu 3d electrons asphericity

It has been already shown in section 2.3.4 that the incomplete filling of d orbitals is reflected into a departure of the Laplacian of the electron density around the metallic centre from spherical symmetry. A similar behaviour was also expected for this compound.

In the EE azido complex, the ligands molecules coordinate the Cu metallic atoms in an almost square-pyramidal arrangement. According to the crystal field theory, this geometry together with the d^9 electronic configurations, predicts that the singly occupied orbital is the $d_{x^2-y^2}$, identically to what found in the End-On azido complex. The $d_{x^2-y^2}$ orbital is oriented towards the atoms belonging to the pyramidal square base, that are the three



(a) S_s{Cu,N,O,C} (b) $S_{s,mag}${Cu,N,O,C} (c) $S_{s,relax}${Cu,N,O,C}

(d) S_s{Cu,N,O} (e) $S_{s,mag}${Cu,N,O} (f) $S_{s,relax}${Cu,N,O}

(g) $S_{s,relax}${Cu} (h) $S_{s,relax}${Cu} (i) $S_{s,relax}${Cu}

Figure 2.26: CASSCF(10,10) source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-to-End complex evaluated in the least-square plane of the azido moieties. Subsets containing Cu, N, O and C atoms ({Cu,N,O,C}), without the C contributions ({Cu,N,O}) and only the Cu atoms ({Cu}) have been considered. Colour code and isosurface levels are the same of Figure 2.2.

N1, N2 and O1 belonging to the L ligand and the short-bonded azido N5 atom. The other nitrogen atom of the azide, N3, points along the z-axis direction and it interacts with the doubly occupied d_{z^2} Cu orbital. This inhomogeneity can explain the different bond length of Cu-N5 (2.000 Å) and Cu-N3 (2.356 Å).

Analogously to what found for the End-On complex, the $-\nabla^2\rho$ distribution around the Cu atom departs from the spherical symmetry and it shows (3,+3) $-\nabla^2\rho(\boldsymbol{r})$ minima located along the bond paths linking the atoms of the ligands and (3,-3) maxima lying in between such paths.[24,98,99] The former critical points are associated to charge depletion regions (CDs) of the electron density, while the latter to charge concentration zones (CCs). (3,-1) ring critical point (R) and bond critical points (BCPs) are also found along the Cu-ligand axis. Distribution of the reference points, except for the CCs point, is shown in Figure 2.32.

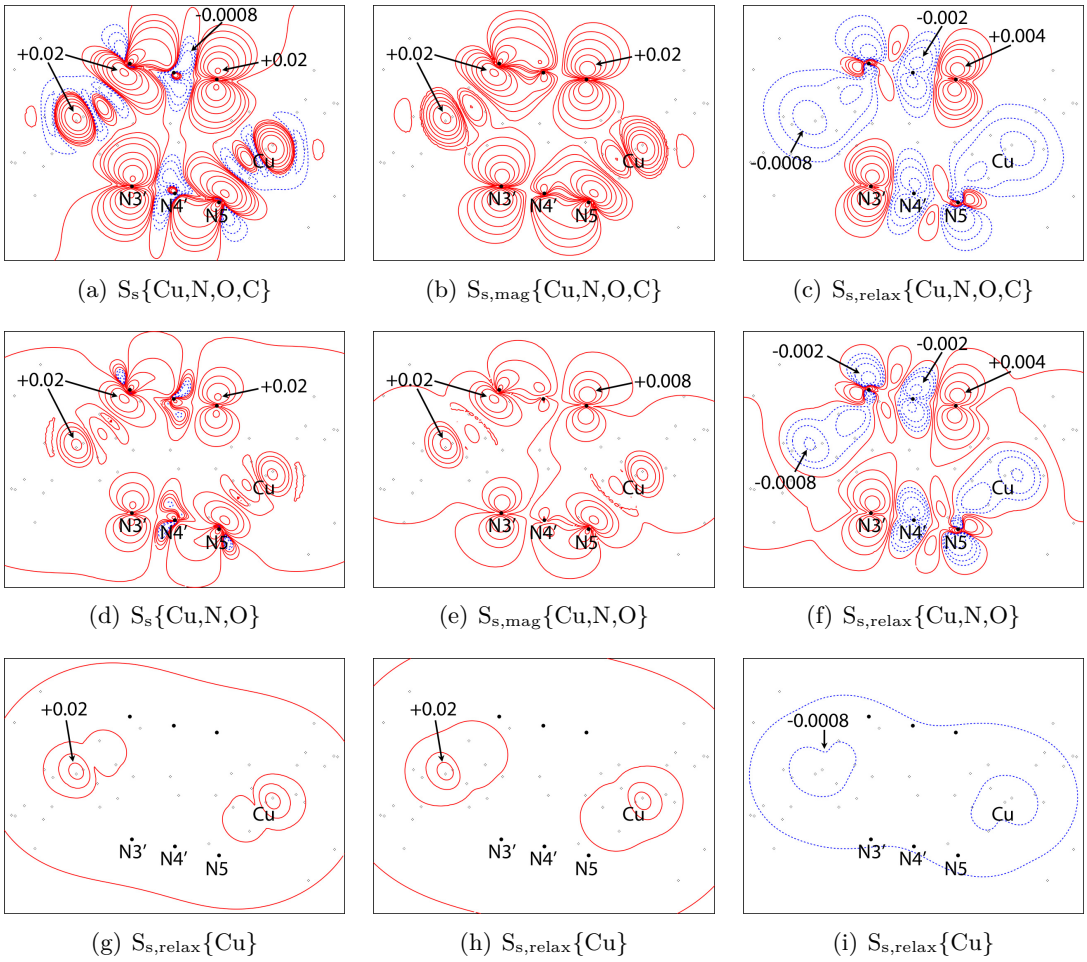


Figure 2.27: B3LYP source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-to-End complex evaluated in the least-square plane of the azido moieties. Subsets containing Cu, N, O and C atoms ({Cu,N,O,C}), without the C contributions ({Cu,N,O}) and only the Cu atoms ({Cu}) have been considered. Colour code and isosurface levels are the same of Figure 2.2.

The analysis performed on these reference points confirms what expected from a simple model analysis. The two CD points located along the Cu-N5 and Cu-N3 have different behaviours since they interact with different orbitals, with a large positive spin density at the Cu-N5 CD and a negligible one at the Cu-N3 CD. Data confirming these outcomes are reported in Tables 2.11 and 2.12.

The $-\nabla^2\rho$ value at the Cu-N5 CD is more positive than at the Cu-N3 CD, both for CASSCF(10,10) (19.10 *a.u.* and 13.65 *a.u.*, respectively) and DFT functionals (17.93/17.30 *a.u.* and 14.14/14.26 *a.u.*, respectively). This fact reveals that the electron density is more depleted at the Cu-N5 CD than at the Cu-N3 symmetric reference point. As expected, the spin density at Cu-N5 CD point is high, being equal to 0.363 *a.u.* at the CASSCF level, while the same quantity is negligible for the Cu-N3 CD (0.000 *a.u.*). The same statement remains qualitatively true for both hybrid B3LYP and pure BLYP functionals, where the

Table 2.11: CASSCF(10,10) distances (R_{Cu}), minus electron density Laplacians ($-\nabla^2\rho(\boldsymbol{r})$), spin densities ($s(\boldsymbol{r})$), magnetic and relaxation components of the spin density ($s_{mag}(\boldsymbol{r})$ and $s_{relax}(\boldsymbol{r})$) evaluated at the selected reference points along the Cu-ligands bonds. If not otherwise stated, all the values are reported in atomic units

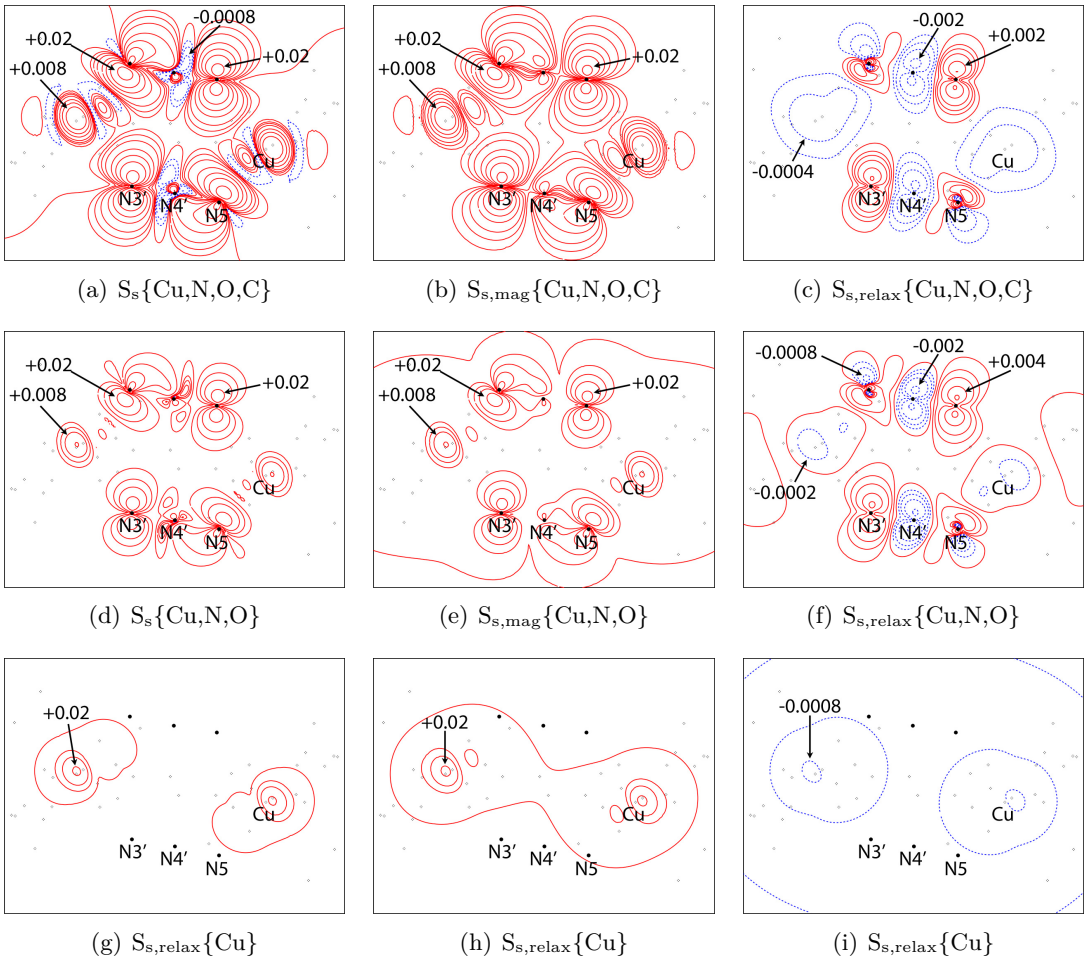
Bond	CPs[1]	R_{Cu}Å	$-\nabla^2\rho(\boldsymbol{r})$	$s(\boldsymbol{r})$	$s_{mag}(\boldsymbol{r})$	$s_{relax}(\boldsymbol{r})$
Cu-N5	CD	0.43	19.10	0.363	0.404	-0.041
	R	0.29	-29.48	0.913	1.016	-0.103
	BCP	0.94	0.35	0.007	0.007	-0.001
Cu-N3	CD	0.45	13.65	0.000	0.000	0.000
	R	0.28	-69.88	0.001	0.001	0.000
	BCP	1.16	0.15	0.000	0.000	0.000
Cu-N2	CD	0.43	19.39	0.390	0.434	-0.044
	R	0.29	-27.33	0.969	1.078	-0.109
	BCP	0.94	0.35	0.007	0.008	-0.001
Cu-N1	CD	0.43	19.18	0.364	0.405	-0.041
	R	0.29	-30.27	0.926	1.030	-0.104
	BCP	0.92	0.43	0.007	0.008	-0.001
Cu-O1	CD	0.43	19.43	0.392	0.436	-0.044
	R	0.29	-27.37	0.978	1.088	-0.110
	BCP	0.91	0.50	0.008	0.009	-0.001

[1] CD: $-\nabla^2\rho$ (3,+3) charge depletion; R: $-\nabla^2\rho$ (3,+1) ring; BCP: ρ (3,-1) bond critical point.

Table 2.12: B3LYP/BLYP distances (R_{Cu}), electron density Laplacians ($-\nabla^2\rho(\boldsymbol{r})$), spin densities ($s(\boldsymbol{r})$), magnetic and relaxation components of the spin density ($s_{mag}(\boldsymbol{r})$ and $s_{relax}(\boldsymbol{r})$) evaluated at the selected reference points along the Cu-ligands bonds. If not otherwise stated, all the values are reported in atomic units

Bond	CPs[a]	R_{Cu}(Å)	$-\nabla^2\rho(\boldsymbol{r})$	$s(\boldsymbol{r})$	$s_{mag}(\boldsymbol{r})$	$s_{relax}(\boldsymbol{r})$
Cu-N5	CD	0.43/0.44	17.93/17.30	0.282/0.226	0.279/0.223	0.004/0.002
	R	0.29/0.29	-40.36/-46.59	0.770/0.637	0.753/0.625	0.016/0.012
	BCP	0.97/0.98	0.26/0.24	0.003/0.001	0.003/0.001	-0.002/-0.001
Cu-N3	CD	0.45/0.45	14.14/14.26	0.001/0.003	0.001/0.003	0.002/0.002
	R	0.28/0.28	-68.73/-68.06	0.013/0.022	0.003/0.013	0.011/0.009
	BCP	1.18/1.19	0.14/0.14	0.000/0.000	0.000/0.000	0.000/0.000
Cu-N2	CD	0.43/0.44	17.99/17.32	0.293/0.232	0.289/0.230	0.003/0.002
	R	0.29/0.29	-41.19/-48.46	0.796/0.653	0.781/0.642	0.015/0.010
	BCP	0.97/0.98	0.25/0.24	0.002/0.001	0.004/0.002	-0.002/0.000
Cu-N1	CD	0.43/0.44	18.04/17.43	0.284/0.226	0.280/0.224	0.003/0.002
	R	0.29/0.29	-41.95/-48.66	0.785/0.648	0.769/0.637	0.016/0.011
	BCP	0.94/0.95	0.33/0.32	0.002/0.001	0.004/0.001	-0.002/0.000
Cu-O1	CD	0.43/0.44	18.0/17.40	0.292/0.231	0.289/0.229	0.003/0.002
	R	0.29/0.29	-41.69/-48.91	0.795/0.653	0.780/0.642	0.015/0.010
	BCP	0.94/0.94	0.40/0.39	0.003/0.002	0.005/0.002	-0.002/0.000

[a] CD: $-\nabla^2\rho$ (3,+3) charge depletion; R: $-\nabla^2\rho$ (3,+1) ring-like; BCP: ρ (3,-1) bond critical point



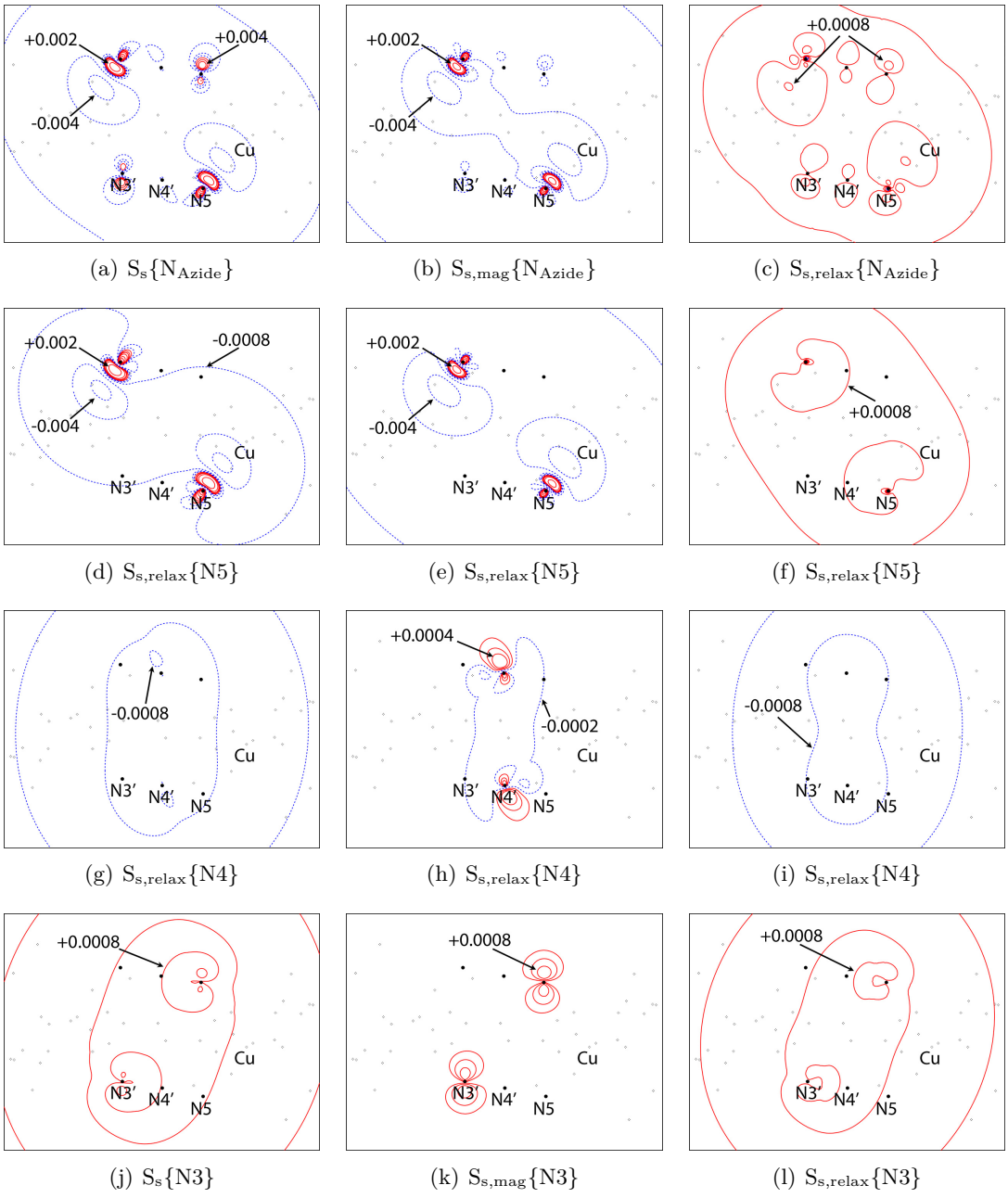
(a) $S_s\{Cu,N,O,C\}$ (b) $S_{s,mag}\{Cu,N,O,C\}$ (c) $S_{s,relax}\{Cu,N,O,C\}$

(d) $S_s\{Cu,N,O\}$ (e) $S_{s,mag}\{Cu,N,O\}$ (f) $S_{s,relax}\{Cu,N,O\}$

(g) $S_{s,relax}\{Cu\}$ (h) $S_{s,relax}\{Cu\}$ (i) $S_{s,relax}\{Cu\}$

Figure 2.28: BLYP source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-to-End complex evaluated in the least-square plane of the azido moieties. Subsets containing Cu, N, O and C atoms ({Cu,N,O,C}), without the C contributions ({Cu,N,O}) and only the Cu atoms ({Cu}) have been considered. Colour code and isosurface levels are the same of Figure 2.2.

spin density value for Cu-N5 CD is smaller (0.282 and 0.226 *a.u.*, respectively) than the one found by the CASSCF method. The spin density evaluated at the Cu-N3 CD by the two DFT functionals is close to the one predicted by CASSCF (0.001 and 0.003 *a.u.* for B3LYP and BLYP, respectively). A similar trend is found for the (3,+1) R $-\nabla^2\rho$ critical points but with larger magnitude of the spin density values for both points. This is due to the fact that also these points are found along all the Cu-ligand axes, which means that they are associated to the $d_{x^2-y^2}$ orbital, but located closer (0.28-0.29 Å) than the CDs. Both CDs and Rs points clearly confirm the difference between the two Cu-N_{azido} bonds, where the Cu-N5 is the one involved with the magnetic $d_{x^2-y^2}$ orbital and Cu-N3 with the non-magnetic d_{z^2}. It is worth noting that for the Cu-N5 reference points the magnetic term of the spin density is the dominant one. The relaxation terms are very low in the case of the DFT functionals and they contribute ≈ 1-2% to the total density. On the contrary,



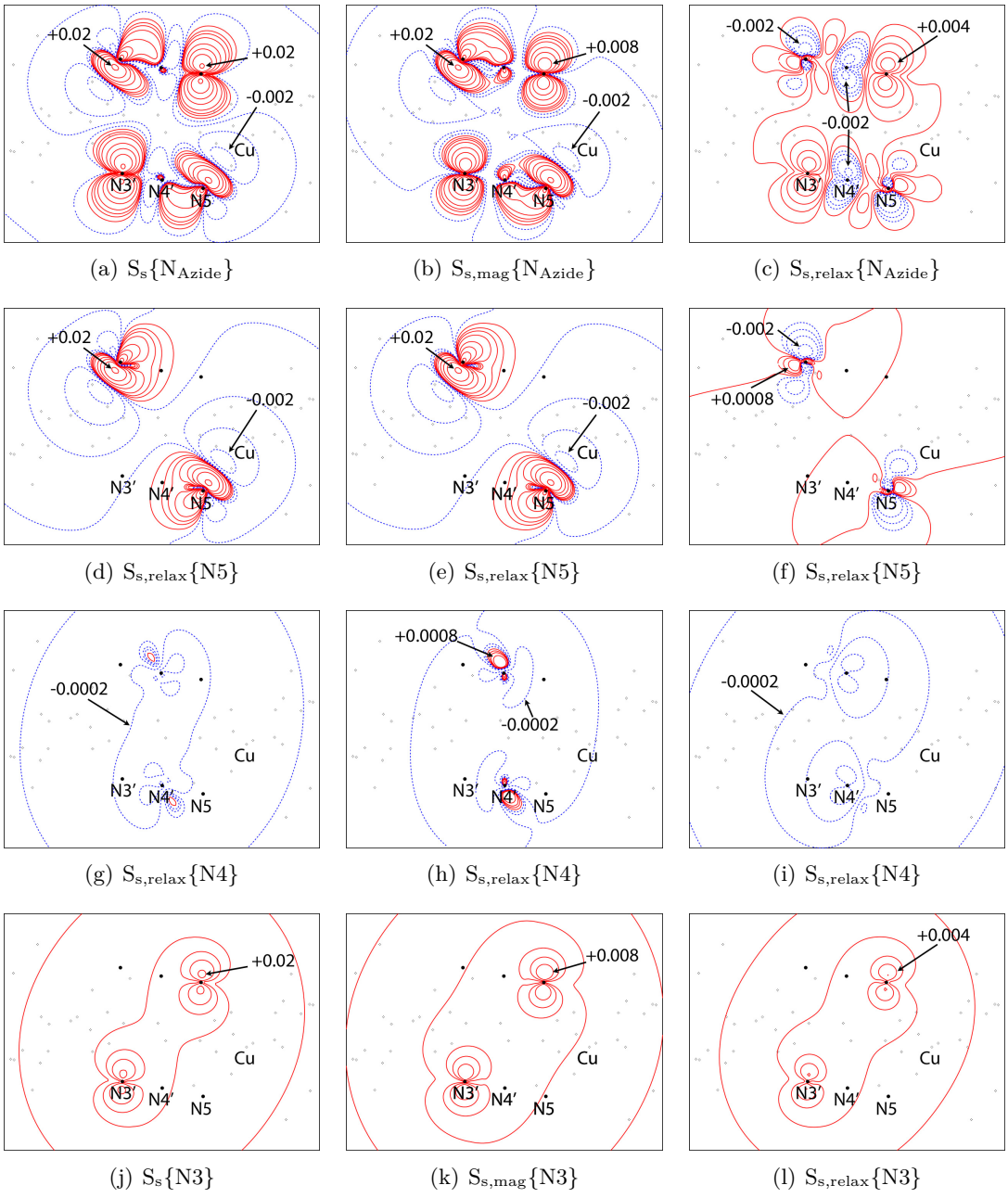
(a) $S_s\{N_{Azide}\}$ (b) $S_{s,mag}\{N_{Azide}\}$ (c) $S_{s,relax}\{N_{Azide}\}$

(d) $S_{s,relax}\{N5\}$ (e) $S_{s,relax}\{N5\}$ (f) $S_{s,relax}\{N5\}$

(g) $S_{s,relax}\{N4\}$ (h) $S_{s,relax}\{N4\}$ (i) $S_{s,relax}\{N4\}$

(j) $S_s\{N3\}$ (k) $S_{s,mag}\{N3\}$ (l) $S_{s,relax}\{N3\}$

Figure 2.29: CASSCF(10,10) source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-to-End complex evaluated in the least-square plane of the azido moieties. Subsets containing the N atoms of the azido groups($\{N_{Azido}\}$) and the pairs of symmetric atoms N5+N5', N4+N4' and N3+N3' ({N5}, {N4} and {N3}, respectively) have been considered. Colour code and isosurface levels are the same of Figure 2.2.



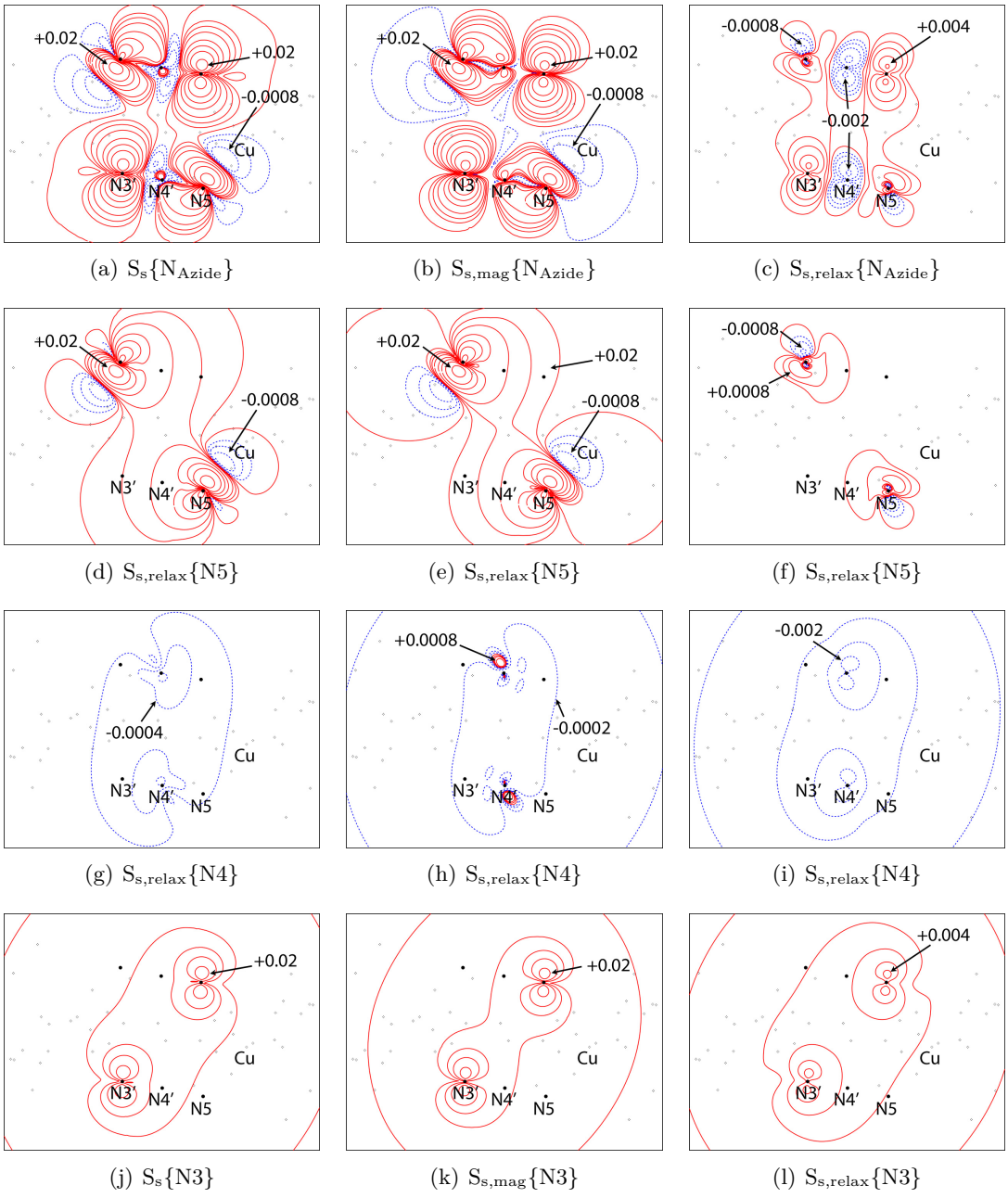
(a) $S_s\{N_{Azide}\}$ (b) $S_{s,mag}\{N_{Azide}\}$ (c) $S_{s,relax}\{N_{Azide}\}$

(d) $S_{s,relax}\{N5\}$ (e) $S_{s,relax}\{N5\}$ (f) $S_{s,relax}\{N5\}$

(g) $S_{s,relax}\{N4\}$ (h) $S_{s,relax}\{N4\}$ (i) $S_{s,relax}\{N4\}$

(j) $S_s\{N3\}$ (k) $S_{s,mag}\{N3\}$ (l) $S_{s,relax}\{N3\}$

Figure 2.30: B3LYP source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-to-End complex evaluated in the least-square plane of the azido moieties. Subsets containing the N atoms of the azido groups($\{N_{Azido}\}$) and the pairs of symmetric atoms N5+N5', N4+N4' and N3+N3' ({N5}, {N4} and {N3}, respectively) have been considered. Colour code and isosurface levels are the same of Figure 2.2.



(a) $S_s\{N_{Azide}\}$ (b) $S_{s,mag}\{N_{Azide}\}$ (c) $S_{s,relax}\{N_{Azide}\}$

(d) $S_{s,relax}\{N5\}$ (e) $S_{s,relax}\{N5\}$ (f) $S_{s,relax}\{N5\}$

(g) $S_{s,relax}\{N4\}$ (h) $S_{s,relax}\{N4\}$ (i) $S_{s,relax}\{N4\}$

(j) $S_s\{N3\}$ (k) $S_{s,mag}\{N3\}$ (l) $S_{s,relax}\{N3\}$

Figure 2.31: BLYP source function partial reconstructed spin density maps (first panel) and their magnetic (second panel) and relaxation (third panel) contributions for the End-to-End complex evaluated in the least-square plane of the azido moieties. Subsets containing the N atoms of the azido groups($\{N_{Azido}\}$) and the pairs of symmetric atoms N5+N5', N4+N4' and N3+N3' ({N5}, {N4} and {N3}, respectively) have been considered. Colour code and isosurface levels are the same of Figure 2.2.

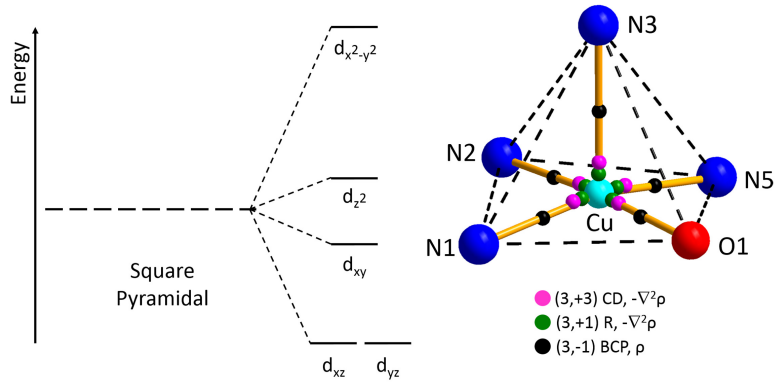


Figure 2.32: Cu atom 3d electron asphericity in the End-to-End azido Cu complex. Left: d orbital energy order for a metal in a square-pyramidal geometry. Right: critical points distribution around the Cu atom. Reference points colour code is: pink: $-\nabla^2\rho$ (3,+3) charge depletion (CD); green: $-\nabla^2\rho$ (3,+1) ring-like (R); black: ρ (3,-1) bond critical point (BCP).

the CASSCF level of theory predicts a larger effect of the relaxation, $\approx 11\%$ on absolute value. It is also interesting noting that while in both B3LYP and BLYP the magnetic and relaxation terms cooperate to increase the spin density in the Cu-N5 CD and R points, the CASSCF contributions oppose each others. On average, it is clear from these outcomes that CASSCF and DFT levels of theory show large quantitative differences, confirming once again the big accuracy difference between these methods. Also B3LYP and BLYP disagree from a quantitative point of view, showing that the spin results strongly depend on the adopted functional.

In agreement with what previously found for the End-On Cu azide, the spin density at the BCP is negligible or null for both kinds of bonds. This can be explained considering that the electrons tend to maximize their coupling in bonding regions. The BCPs distances found by the CASSCF method are slightly lower than those found for the DFT functionals (for example, 1.16 *a.u.* and 1.18/1.19 *a.u.* for CASSCF(10,10) and B3LYP/BLYP, respectively). These BCPs shifts toward the metal found in the correlated wavefunction are similar to those observed in the End-On complex and may be related to the excessive α-spin delocalization and consequent Cu-ligand bond covalency predicted by the DFT methods.

ρ and $-\nabla^2\rho$ values of the BCP reference points compare well with the results obtained by Deutsch *et al.* using a spin-split multipolar model.[82] On the contrary, the $s(\boldsymbol{r})$ values found by Deutsch *et al.* are more than halved with respect to those found at CASSCF(10,10) level, highlighting as already discussed in Section 2.3.6 that a good agreement between the spin density resolved components is not sometimes sufficient to retrieve good spin density values.

The last parts of Tables 2.11 and 2.12 show the analogous results for the Cu-N1, Cu-N2 and Cu-O1 reference points. As expected, these outcomes all resemble those found for the Cu-N5 bond and not those along the Cu-N3 bond.

2.4.6 Source Function at Selected Reference points

As for the End-On complex in Section 2.3.5, some of the critical points reported in Tables 2.11 and 2.12 have been used as reference points for analysing the percentage spin density SF contributions. As shown in Tables 2.11 and 2.12, Cu-N5 and Cu-N3 linkages represent two very distinct Cu-ligand interactions and their associated CD critical points have thereby selected as interesting reference points in the SF analysis.

Such an analysis could be very useful to get further insights on the mechanism of spin transmission and communication between the paramagnetic and the non-magnetic centres. Atomic source function percentage contributions on these two reference points are reported in Figure 2.33.

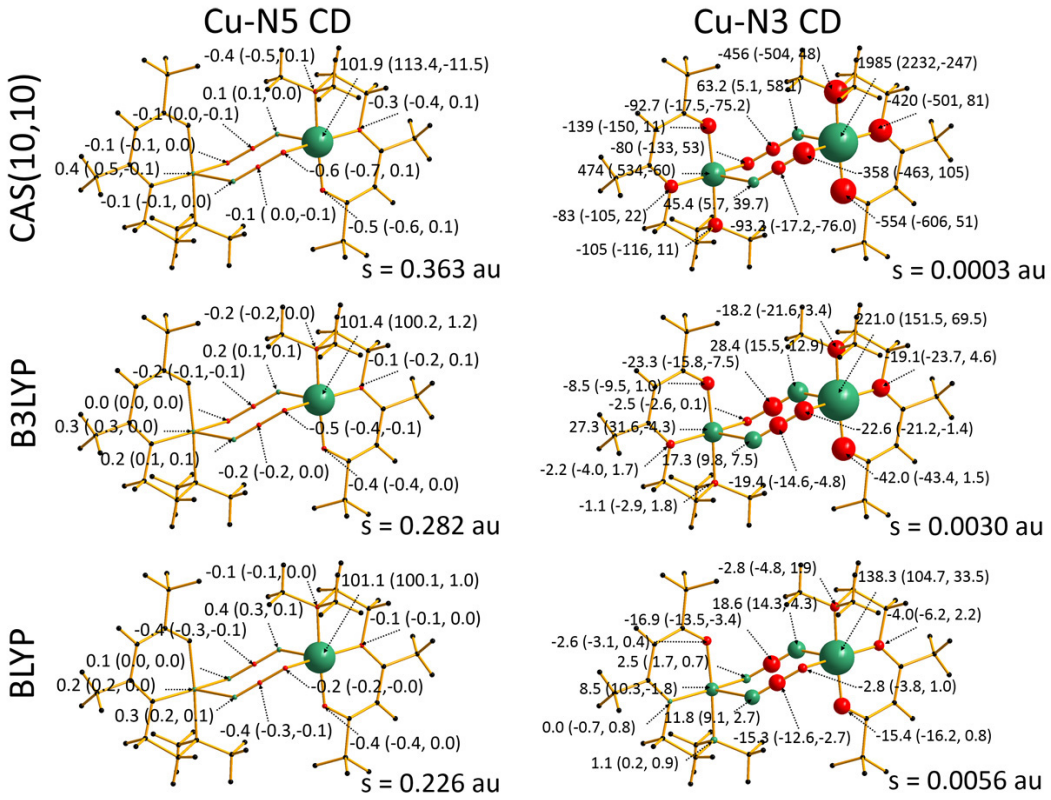


Figure 2.33: CASSCF(10,10), B3LYP and BLYP SF percentage contributions to the spin density $s(\boldsymbol{r})$ at charge depletion (CD) reference points along the short Cu-N5 bond (Cu-N5 CD) and the long Cu-N3 bond (Cu-N3 CD). Numerical values represent the total(magnetic/relaxation) percentage contributions. Green/red spheres denote α/β effect on the spin density of the selected reference point.

The source function contributions to the spin density at the Cu-N5 CD point (left panels in Figure 2.33) clearly show that the Cu source is the dominant one. All the used methods only agree from a qualitative point of view. DFT models found that the total Cu contribution is strictly close to 101% (101.4% and 101.1% for B3LYP and BLYP, respectively). The magnetic term is the dominant one and it is responsible for almost the totality of the spin density (100.2% and 100.1% for the hybrid and pure functionals,

respectively), while the relaxation terms yield only about 1% of the total density. In the CASSCF case, the Cu magnetic contribution is higher than the total spin density (113.1%), but it is counteracted by a non-negligible relaxation contribution (-11.5%), leading to an overall 101.9% contribution from the paramagnetic centre. The percentage contributions coming from all the other atoms are lower than 1% for all the used level of theory. However, it is important to remind that on absolute values these contributions account for $\approx 0.002/0.003$ e, which is higher than the total spin density evaluated in other points, like the Cu-N3 CD. From a qualitatively point of view, all the methods agree on the relative sign of the other atomic contributions. All the coordinating atoms of the L ligand (N1, N2 and O1) yield a β effect to the total spin density at the Cu-N5 CD. The azido nitrogen atoms N5 and N4 behave in the same way, while the long-bonded N3 and the Cu' atoms yield α effect contributions. This overall situation is similar to what found for the End-On azide for the Cu'-N1 CD reference point (see Figure 2.11, Section 2.3.5).[68] The α effect yielded by the far Cu' atom agrees with the two copper atoms being weakly ferromagnetically coupled.

The S_s% contribution evaluated at the CD along the longer Cu-N3 bond highlights a different *scenario*. Even if the total spin density evaluated in this reference point is quite small, the contributions are highly delocalized over all the complex with large and opposing values. The $S_s\{Cu\}$% source still remains the most predominant one, but now contrary to what we saw for the Cu-N5 CD point, it does not account for the totality of the reconstructed density. All the adopted levels of theory predict α and β effects pattern similar to those found at the Cu-N5 CD reference point. However, percentage sources greatly differ since the values of the SD at the chosen reference point are about 2-3 order of magnitude lower than at the Cu-N5 CD and also vary among the methods adopted (0.0003 *a.u.* for CASSCF(10,10), 0.0030 *a.u.* for B3LYP and 0.0056 au for BLYP). CASSCF(10,10) found a total spin density value one order of magnitude lower than the DFT functionals (0.0003 *a.u.* for CASSCF and 0.0030 and 0.0056 *a.u.* for B3LYP and BLYP, respectively). Both magnetic and relaxation terms play an important role in the determination of the total density. B3LYP sources are, on average, higher than those of the BLYP functional despite the spin density predicted by the former functional is lower than the one obtained by the latter. This fact shows that there may be a more favoured spin communication between the atoms in the hybrid B3LYP functional than in BLYP.

We have also evaluated the percentage source function contributions when the N3 and N5 nuclei positions are taken as reference points. Figure 2.34 shows the percentage atomic contributions at these points for all the used methods through the classical sphere representation.

The spin density evaluated at the N5 nucleus position is 3-4 times smaller than the same quantity evaluated on the N1 atom of the End-On complex (compare Tables 2.11 and 2.12 with Table 2.4). CASSCF(10,10) and B3LYP/BLYP methods find $s(\boldsymbol{r})$ values equal to 0.024 and 0.058/0.051 *a.u.*, respectively. On the End-On compound instead, $s(\boldsymbol{r})$ value from CASSCF(6,6) and B3LYP are equal to 0.076 and 0.221 *a.u.*, respectively (see

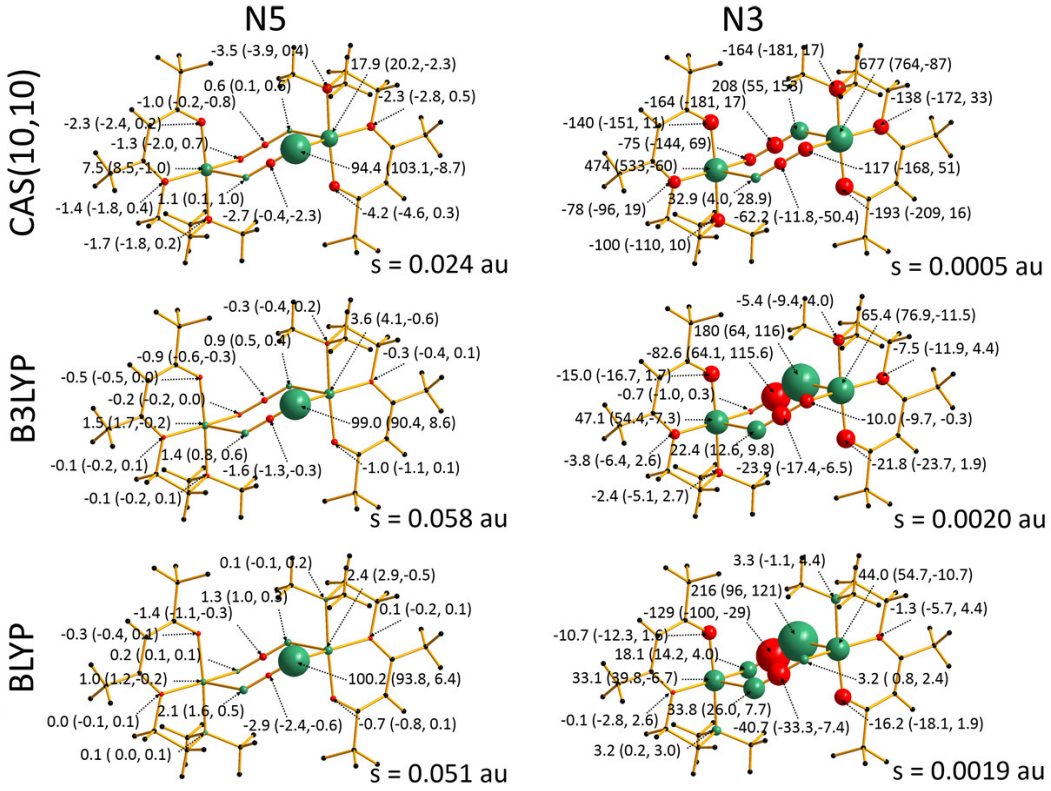


Figure 2.34: CASSCF(10,10), B3LYP and BLYP SF percentage contributions to the spin density $s(\boldsymbol{r})$ at N5 and N3 nuclei positions. Numerical values represent the total(magnetic/relaxation) percentage contributions. Green/red spheres denote α/β effect on the spin density of the selected reference point.

Section 2.3.5). This is clearly a sign of the less efficient α-spin delocalization found in the End-to-End complex with respect to the End-On, that also reflects the lower coupling constant J in the former complex (see Section 2.1). The role of N5 in the spin coupling mechanism between the two Cu atoms is mediated in the EE complex by the other azido nitrogen atoms N4 and N3 and it is therefore less efficient. On the contrary, in the EO complex the analogous N1 is directly bonded to the metallic centres. The atomic sources evaluated at the N5 nucleus positions are in agreement to what found in Figure 2.12 for the analysis of the N1 point in the End-On complex (Section 2.3.5). The main contribution comes from N5, that is from the atom hosting the reference point. The N5 sources account $\approx$ 95-100% of the total reconstructed density (Figure 2.34). The magnetic terms are, also in this case, the most important ones in the reconstruction but, differently to what found for the CDs point above, the relaxation contributions are predicted to be non-negligible by all the used methods. While the CASSCF(10,10) predicts $S_{s,mag}\{N5\}$% and $S_{s,relax}\{N5\}$ to oppose each other in the reconstruction of the unpaired electron density (+103.1% and -8.7%, respectively), the B3LYP and the BLYP predict cooperative magnetic and relaxation sources (90.4% and 8.6% for the former and 93.8% and 6.4% for the latter). The other atomic contributions have low percentage values. The two Cu atoms and the N3

and N3' atoms produce α effects at the N5 nucleus, while the middle azido nitrogen N4 and the other ligand heteroatoms yield negative β effects. The percentage values of the two Cu source are obviously different because of geometrical factors and depend on the level of theory used. The CASSCF method predicts larger $S_s\{Cu\}$ and $S_s\{Cu'\}$ contributions that together account for the 25.4% of the total spin density at the N5 nucleus. On the contrary, B3LYP and BLYP functionals retrieve that the overall Cu sources only account for 3-5%, although the ratio between Cu and Cu' sources remain constantly equal to $\approx$ 2.4 as found for the CASSCF(10,10). These outcomes reveal that either hybrid or pure DFT underestimate spin connection and communication efficiency relative to the CASSCF method.

The spin density values at the N3 nucleus are instead similar to those found for the N3 terminal azido nitrogen in the End-On complex. CASSCF(10,10) calculates a very small $s(\boldsymbol{r}) = 0.0005$ *a.u.* value, while the B3LYP and BLYP functionals predict spin densities 4 times larger, equal to 0.0022 *a.u.* and 0.0019 *a.u.*, respectively.

The negligible spin density at the N3 nucleus point is the result of concurring and opposing atomic sources. The two metallic atoms Cu and Cu', together with the symmetric equivalent N3', increase the spin density at the chosen point producing α effects, while almost all the other atoms counteract with β effects. Contrary to what found for the N5 nucleus, here the three levels of theory employed disagree on the nature of the main atomic contribution. According to CASSCF(10,10), the two Cu atoms are the main contributors to the reconstruction of the spin density at N3 nucleus position (see Figure 2.34). On the contrary, both DFT functionals predict (with different percentage values) that the major effect comes from the N3 atom. This different behaviour can be once more expressed as a limit of DFT to correctly describe the unpaired density distribution with respect to the multi-configurational methods.

In this section it was demonstrated once more how the source function descriptor is able to catch subtle features and to explain differences when different level of theory are used. DFT functionals tend to overestimate the α-spin delocalization from the paramagnetic centres towards the ligand atoms, and on the other hand they underestimates the communication between the atoms and then the ability of the metallic centres to transmit their spin information to the non-magnetic atoms.

2.5 Conclusions

In literature, spin transmission mechanisms and magnetic interactions between paramagnetic atoms through diamagnetic ligands are usually explained in terms of spin delocalization, polarization or superexchange mechanisms. Models based on orbital descriptions[87,88,105] or topological analysis of the spin density[105] are used to correlate the spin density distribution with the magnetic properties, and to get further insights on the mechanism of spin transmission.

In this chapter and in the related works[68,69] we have presented some recent applications

on the use of the source function descriptor to gain insight on the spin information transmission mechanism in magnetic systems. We applied the source function tool to analyse the spin density distribution of two different complexes: an End-On[87] and an End-to-End[88] dinuclear Cu(II) azido complexes. In both systems, the ferromagnetic interaction among the two Cu atoms is mediated by the azido moiety.

The source function descriptor, together with the distribution of the electron density Laplacian critical points, enables us to highlight not only how the magnetic centres contribute to determine the local spin delocalization and polarization at any point, but also how the other atoms cooperate in a positive or negative way to this. This feature is a peculiarity of the source function tool, that allows to see the electron or spin density distribution in both the directions: from the magnetic to the non-magnetic centres and *viceversa*. The source function is able to clarify the role of atoms or group of atoms to determine the spin density at any point within a *cause-effect* relationship. Moreover, the introduction of a new way to represent this descriptor, the source function partial reconstructed maps, enables one to obtain an immediate visualization of the role played by a subset of atoms in a molecular plane or in the space without depending on the choice of reference points. The decomposition of the atomic sources into the magnetic and relaxation contributions helps to gain more chemical insights and to disclose the origin of the unpaired electrons distribution and the discrepancies related to the choice of different levels of theory.

In both the studied complexes, it is evident that the spin density distributions are strongly dependent on the theoretical approach used. CASSCF method predicts that QTAIM and Mulliken spin populations are mainly localized on the Cu atom while DFT functionals yield a more delocalized picture. The different computational methods retrieve spin densities maps with similar features but also with remarkable differences, like for instance the negative region around the Cu atoms found in the B3LYP/BLYP functionals and not present in the CASSCF calculation or the different mechanism of spin transmission along the azido moieties. The use of the source function partial reconstructed maps reveal that the origin of these negative spin density regions are due to the C atoms of the ligand groups. The analyses highlight that DFT functionals overestimate spin delocalization from the magnetic centres towards the non-magnetic ones, while comparatively they underestimate spin transmission and communication among atoms. This latter evidence is clear from the analysis of the atomic decomposition of the partial reconstructed density maps. As a consequence of this excessive spin delocalization, B3LYP and BLYP Hamiltonians overestimate spin polarization mechanism between the Cu metallic centre and the ligand atoms. The comparison of the hybrid B3LYP and the pure BLYP in the End-to-End azido complex reveals that these effects are more evident in the latter than in the former, contrary to what Boguslawski *et al.* found for the Fe(II) nitrosyl complex.[55] In turn, these facts also affect the partial covalent nature of metal-ligand interaction which eventually dampens spin transmission in the system.

Data from polarized neutron diffraction techniques seem to be intermediate among the

CASSCF and DFT outcomes for both studied systems, but on average close to the former. Despite practical and technical limitations of the PND experiment, the experimental results seem to be reasonably accurate. This, together with the source function analysis of the *ab initio* wavefunction, suggests that high-quality calculations are required to proper describe the spin density distribution.

The Cu 3d Laplacian asphericity in both End-On and End-to-End complexes presents analogies. Charge depletion critical points are found close to Cu nuclei along the Cu-ligand bonds, while charge concentration points are located in between the former. In the EE complex, Cu-N3 (long) and Cu-N5 (short) bonds of the Cu centre with the azido moieties show different topological properties due to different orbital occupation predicted by the ligand field theory. In the EO system, the Cu-N1 critical points have similar topological properties with respect to the Cu-N5 bond in the EE. These differences and analogies are reflected in the source function percentage analysis. This analysis is able to highlight the different degree of localization/delocalization of the atomic contributions in the two cases, revealing a strongly localized scenario for the EE Cu-N5 bond and a more delocalized panorama for the Cu-N3. Moreover, the SF percentage analysis confirms that the Cu-N5 bond of the EE is similar to the Cu-N1 bond in the EO complex, while no spin information seems to be transmitted along the Cu-N3 bond.

As a last achievement, we demonstrated in the End-On case that a visual perfect agreement of the two components of the electron density distribution, ρ_α and ρ_β, evaluated with two different approaches is not a sufficient condition to obtain a corresponding good agreement between their spin density distributions.

Chapter 3

Source Function Electron Density Study of DNA Base Pairs[i]

Life is one of the most fascinating topic humanity has ever wondered about since the origin of history. Philosophers, alchemists and in the end scientists, have tried to explain its origin and how it is possible that life can generate other life with some inherited characteristic but at the same time with such a large variability. Now we know that the deoxyribonucleic acid, the DNA, is the main player in the hereditary mechanism, but this discovery was not achieved so long ago.

In 1944 one of the fathers of the quantum mechanics, Erwin Schrödinger, published a book entitled "What is Life?" where he introduced for the first time the idea of an "aperiodic crystal" that contains genetic information in its configuration of chemical bonds.[107]

Although the first DNA complex with proteins was obtained in 1896 by Friedrich Miescher,[108] Schrödinger did not have the benefit of knowing the real nature of the genetic material when he published his book. This latter was only elucidated in 1944, more or less simultaneously to Scrödinger's book publication, by Avery *et al.* during their studies on the Griffith's pneumococcal transforming principle.[109] By the 1945, the chemical composition of the genetic material was known but no information neither on the nucleic acid complementarity nor on the three-dimensional structure of the DNA molecule were available. Chargaff and co-workers in 1948 first reported the rules on the pyrimidine and purine bases complementarity.[110–112] They discovered that the amount of guanine base in DNA should be equal to cytosine and that the same relation should be valid for the adenine-thymine bases.[110–112] Few years later, in 1953, Watson and Crick (WC) built, on the basis of X-ray diffraction scattering images obtained by Frankling and Gosling,[113] their famous model: the double antiparallel α helix DNA.[114–118] Thanks to the Chargaff's rules, Watson and Crick were able to assign the nitrogenous bases using their complementarity and disposing them close to the helical axis, bonded through hydrogen interactions. The phosphate groups were placed at the periphery of the double-helix on the basis of a simple electrostatic repulsion argument. The WC model was in opposition to the other one

[i]This chapter fully reports the results already published in reference [106]

proposed by Pauling and Corey, where the phosphate groups were supposed to be closer to the centre of the double helix and the nitrogenous bases far from it.[119,120] Watson and Crick's pairing in DNA requires adenine (A) to be bonded to thymine (T) by two hydrogen bonds, and guanine (G) to cytosine (C) through three hydrogen interactions. Recently, Parathasarathi *et al.* reported for the first time theoretical evidences for a third hydrogen bond, weaker that the other two, in the AT pair (C-H$\cdots$O interaction).[121]

The hydrogen bonding in the WC model strictly determines specificity and complementarity with an automatic satisfaction of the Chargaff's rules. These interactions are responsible, together with the π-π stacking between adjacent stacked base pairs, for the high stability of the double helix structure and then of the genetic code itself.[122–125] The asymmetric double well potential of the hydrogen bonds in the DNA strands blocks, or slows it down at physiological temperature, the classical possible hopping of an hydrogen atom between the two bases in the same pair through the creation of a potential energy barrier. This expedient allows the DNA to be extremely stable and to be replicated with a very high degree of precision. However, since the hydrogen atoms are quite small, some quantum tunnel effects can show up during the replication.[126–136] If this event occurs immediately prior to cell division, it can lead to mutation potentially dangerous for the organism.[126–129,137–141] This mechanism of genetic error is known in literature as the Löwdin mechanism.[126–136] Single proton tunneling is usually very improbable because of the unfavorable charge separation, while the concerted double proton tunneling in opposite directions has a higher probability to occur because of the preserved charged neutrality.[126,127,142] External applied electromagnetic field can accelerate the double proton hopping and then the induction of mutation.[136–140,143–146]

Nuclei acids (DNA and RNA) are double or single strands constructed by three building blocks: nitrogenous bases, sugars and phosphate groups. The backbone is formed by alternating sugar-phosphate molecules and the bases are covalently bonded to sugars. Only 5 nitrogenous bases are used in DNA and RNA and they can be classified in two groups: purines (adenine (A) and guanine (G)) and pyrimidines (thymine (T), uracil (U) and cytosine(C)). Although these three components co-operate, very often the base pairs are studied separately from the other components of the nucleic acids. This approximation lead to the possibility to use simpler models and then to reduce calculation efforts. The experimental proof that this common approach can be pursuit was obtained by Candeias and Steenken in 1992.[147] In their work, they used 193 nm photons to provoke a ionization-induced photolysis of nucleotides in aqueous solution and then to simulate the effect of ionizing radiation on nucleic acids. Candeias and Steenken found out that the (deoxy)ribose, the phosphate and the (deoxy)ribose-phosphate are all efficiently ionized by this radiation, but their extinction coefficients are two order of magnitude lower than the nitrogenous bases. Moreover, they found similar results performing the same experiment on a model pyrimidine (1,3,5,6-ter-tramethyluracil) whether free or incorporated in nucleotides.[147]. From these outcomes they concluded that the major site where the ionization occurs is not neither the sugar nor the phosphate group, but the base moieties.[147]

Another proof on the possibility to study the nitrogenous bases alone came from DFT studies performed by Schaefer and co-workers.[148] They demonstrated that the addition of the sugar to the base moiety has minimal effect on the AT base pairing energies, independently on the charge (neutral or anionic) of the studied system.[148] The energy variation upon sugar addition is equal to 1 kcal/mol at B3LYP/DZP++ level of theory, revealing that distant groups have negligible effects on the hydrogen bonds strength.[148] These two works demonstrate that studying isolated Watson-Crick base pairs has experimental and theoretical justifications.

The interest on the five nitrogenous bases that are able to encode genetic information in the DNA strands has been a hot topic since the middle of XX century, namely the period when the quantum mechanics was blossoming.[149] A full characterization of physical and chemical properties of the DNA bases was performed across these years using orbitals models[128,137–141], electron density and related DFT properties[76,121,142,150–153].

In the following section the results already published in reference [106] are presented. The aim of the work is to study the role played by the atoms in the system to the construction of the hydrogen bonds and the effect that single ionizations have on these interactions. This study was performed using the tools of quantum theory of atom in molecules (QTAIM)[6–8] and the source function descriptor.[62,154–156] The source function analysis allows one to unravel the single atomic contributions to the density reconstruction at any selected point. If the chosen point is the bond critical point of the hydrogen bonds in the DNA base pair, one can obtain insights on these interactions. The hypothesis tested is whether atoms far from the studied bonds can contribute significantly to the binding strength. This kind of analysis can be used to predict the effect that other possible substituents could have on the hydrogen bonds.[157] Also, the effects of single-ionization on the DNA bases were studied to understand how this DNA damage mechanism affects the stability of the base pairs and how atomic contributions change.

3.1 Methods

The geometries of the two Watson-Crick neutral base pair dimers (AT = adenine-thymine; GC = guanine-cytosine) were fully optimized using DFT B3LYP/6-31+G(d,p) level of theory. Harmonic vibrational frequencies analysis confirms that the systems converged in a minimum of the potential energy surface. Energies and electron densities were further refined by performing single point calculations at the B3LYP/6-311++G(d,p) level of theory on the optimized geometry. Singly ionized radical species (doublet cations, AT+ and GC+, and anions, AT- and GC-) were obtained using the same approach at the unrestricted DFT level of theory.

Model systems for the three different hydrogen bonds, namely N-H$\cdots$N, N-H$\cdots$O and C-H$\cdots$O, were evaluated at the same level of theory of DNA base pairs and at fixed donor-acceptor distances. Distances were selected to cover the range found in the DNA base pairs investigated at their neutral and ionized geometries and they are reported in

Table 3.1. All other geometrical parameters were optimized within the Cs symmetry constraint. Electronic structure calculations and geometry optimizations were performed using Gaussian09 software package.[89]

Table 3.1: Donor-acceptor distances (in Å) for the model hydrogen bond systems. Distances were chosen accordingly to the distances found in the DNA base pairs.

Models	Distances / Å
$H_2N\text{-}H\cdots NH_3$	2.8 / 3.0 / 3.2
$H_2N\text{-}H\cdots OH_2$	2.6 / 2.8 / 3.0 / 3.2
$H_3C\text{-}H\cdots OH_2$	3.6 / 3.8 / 4.0 / 4.2

Topological analysis of the electron density ρ together with the evaluation of the source function contributions were performed using an unpublished modified version of the AIMPAC program package.[92–95] PLOTDEN2016 in-house code has been used to plot the source function partial reconstructed maps. PLOTDEN2016 is an updated and unpublished version of the PLOTDEN2013 code (not published but briefly described in the supporting information of Reference [66]). If not otherwise specified, the Diamond[96] code was employed to draw all the atomic source function percentage ball-and-stick pictures.

3.2 Hydrogen Bonds Models

Table 3.2 lists the atomic SF% values for the electron density at the hydrogen bond BCPs for the three model pairs described in Section 3.1: $H_2N\text{-}H\cdots NH_3$, $H_2N\text{-}H\cdots OH_2$ and $H_3C\text{-}H\cdots OH_2$.

The source function descriptor has been previously used to study hydrogen bonds and a detailed summary can be found in Reference [63]. In general, increasing energy, covalency and local character of the hydrogen bond studied are highlighted by increased source function contributions from the H-bond triad of atoms (hydrogen, donor and acceptor atoms) and the consequent decrease of those from the other atoms.[63] However, the global picture is more complicated than this. The contribution associated to the hydrogen atom directly involved in the interaction undergoes very wide variations according to the nature of the hydrogen bond.[63,70] Moreover, the specific features characterizing the low-barrier hydrogen bonds (LBHB)[158] or the mechanisms controlling the peculiarities of the resonance-assisted hydrogen bonds (RAHB)[63,70,159] are not easy to interpret.

Gatti and co-workers rationalized the behaviour of the source function descriptor on a number of paradigmatic hydrogen bond systems.[70] The evolution of the source function atomic contributions, evaluated at the bond critical point, of a system evolving from a typical weak isolated hydrogen bond (IHB) to a charge-assisted hydrogen bond (CAHB) was studied using a linear dimer of water molecules changing the O-O distance.[70,159] In that study they found out that while the source contributions from the whole water molecules remain almost constant and comparable along the reaction path, large variations

Table 3.2: Source Function percentage (SF%) contributions to the electron density (ρ_{BCP}, in *a.u.*) at the hydrogen bond BCP in the $H_2N\text{-}H\cdots NH_3$, $H_2N\text{-}H\cdots OH_2$ and $H_3C\text{-}H\cdots OH_2$ models as a function of the donor-acceptor distance (in Å).

$R_{D\cdots A}$	$R_{D\text{-}H}$	$R_{H\cdots A}$	ρ_{BCP}	SF%(H)[a]	SF%(H+A)	SF%(H+D+A)[a]	$SF\%_{mol(D)}$[a]
				$H_2N\text{-}H\cdots NH_3$			
3.2	1.009	2.191	0.0180	-45.7	-49.0	17.3	46.9
3.0	1.010	1.990	0.0274	-23.3	-11.8	36.6	46.8
2.8	1.010	1.790	0.0429	-5.9	17.5	57.0	46.6
				$H_2N\text{-}H\cdots OH_2$			
3.2	1.005	2.195	0.0140	-65.6	-52.9	27.4	47.5
3.0	1.004	1.996	0.0215	-38.0	-14.1	48.1	47.3
2.8	1.002	1.798	0.0339	-16.8	15.9	63.8	47.2
2.6	0.997	1.603	0.0546	-0.9	38.4	75.4	47.1
				$H_3C\text{-}H\cdots OH_2$			
4.2	1.091	3.109	0.0025	-241.1	-342.1	-246.4	49.6
4.0	1.091	2.909	0.0037	-165.2	-223.3	-148.8	49.3
3.8	1.091	2.709	0.0056	-111.4	-139.5	20.9	49.0
3.6	1.090	2.510	0.0084	-72.9	-79.9	41.7	48.7

[a] Source contributions from: SF%(H) - hydrogen atom; SF%(H+A) - hydrogen and acceptor atoms; SF%(H+D+A) - hydrogen, donor and acceptor atoms; $SF\%_{mol(D)}$ - proton-donor molecule.

occur for the single atomic contributions. The percentage source contribution from the hydrogen atom directly involved in the interaction increases from large negative values at the equilibrium distance ($O\cdots O$ = 3.020 Å) to slightly positive values when the oxygen atoms are close to each other. On the other hand, the sum of the percentage sources produced by the hydrogen atom and its acceptor range from negative, at equilibrium distance, to positive (less than 50 %) at distance typical of CAHBs ($O\cdots O$ = 2.250 Å)[159], becoming zero at distances ($O\cdots O$ = 2.750 Å) typical of the long chains of $O\text{-}H\cdots O$ bonds in water and alcohols dominated by σ-bond cooperativity.[70] To understand how much this situation is multi-coloured, it is worth noting that for a standard covalent bond the sum of the percentage sources coming from the atoms directly involved in the bond is usually as high as ≈ 80-90 %. If one adds also the contributions from the donor oxygen atoms, the SF% reach values similar to those found for covalent bonds when donor-acceptor distance is as low as 2.5 Å. The source function descriptor in this case describes the hydrogen bonds as a three-centres interaction, as also recently corroborated by Popelier[160] using interacting quantum atoms (IQA) theory[19,161–164] and QTAIM delocalization indices (DI)[74].

At equilibrium distances, the total sources from the atoms not directly involved in the hydrogen bonds can reach the 50 % of the total density at the BCP, while for very short distances this SF% contribution reduces with a consequent increases of the SF%(H+A+D). The former large delocalized pictures has been associated to the dominant electrostatic nature of the hydrogen bond at equilibrium distance, while the latter more localized situation coincides with a partial covalent character, such as in the CAHB case.[63,70]

A profound chemical meaning underlies this large variation of the hydrogen source

function contributions.[63,70] Equation 1.27 in Section 1.7 shows that the source function contribution in a point is strictly correlated with the Laplacian distribution in the space and within the basins. When the atom is isolated, the $\nabla^2\rho$ has a spherical distribution and the SF% values are positive everywhere. For a polyatomic system, the Laplacian distribution of an atom is usually distorted from the spherical symmetry and it can happen that the source contribution from an atom to a specific reference point is negative, that is the local sinks dominate over the local sources. When a weak or moderate hydrogen bond is present in the system, usually the $\nabla^2\rho$ of the H atom is notably asymmetric and inhomogeneous, with a large positive area close to and surrounding the hydrogen bond and a large negative region along the O-H bond. This high asymmetrical distribution of the Laplacian leads to retrieve a negative atomic source when evaluated at the bond critical point of the hydrogen bond. When the distance between the donor and acceptor atoms decreases, the H$\cdots$O hydrogen bond length gets closer and closer to the length of the covalent O-H bond and the $\nabla^2\rho$ becomes more symmetric along the two bonding directions. This fact leads the SF%($\boldsymbol{r}$,H) to become null or positive. It is worth noting that the source function is a very sensitive tool and it is able to see how this subtle changing in the Laplacian distribution or in the basin shape affects the whole system.

In the same work, Gatti and co-workers compared the source function patterns recovered from the two approaching water molecules models with those obtained on a series of prototypical O-H$\cdots$O hydrogen bonds complexes[70] : positively and negatively charge assisted (+CAHB and -CAHB, respectively), resonance assisted (RAHB), polarization assisted (PAHB) and isolated (IHB) hydrogen bonds according to Gilli and Gilli's classification.[159] The comparison revealed similar features, such as the trend against the O$\cdots$O distance, with the exception of the SF%(H) that assumes very different values with respect to the water-water model. The analysis of the RAHBs shows a significantly decreased SF%(H+D+A) contribution relative to the expected values at the hydrogen bond BCP. These deviations highlight once more that the source function descriptor is able to catch very subtle features such as the electron delocalization effects.[65,156] Similar trends are found for systems where either the donor or the acceptor atoms are not oxygen atoms.[156]

A graphical representation of the values reported in Table 3.2 and all the other source function percentage contributions is depicted in Figure 3.1.

Data from Table 3.2 confirm the same trend previously seen by Gatti *et al.*.[70]. Contributions from the donor molecule ($SF\%_{mol(D)}$) and consequently from the acceptor one remain almost unchanged at all the studied distances, independently on the model. On the contrary, the contributions from selected groups of atoms undergo important variations. The trend against the distance is similar for all the studied models but the single values are sensitive to the nature of the atoms involved in the bond. The SF% contribution from the hydrogen atoms is everywhere negative and its absolute value decreases when the donor-acceptor distance reduces. SF%(H) reaches values close to zero in both strong hydrogen interactions at the lowest distance (-5.9 % and -0.9 % for the N-H$\cdots$N and N-H$\cdots$O bond, respectively) while it remains largely negative for the weak C-H$\cdots$O (-72.9%

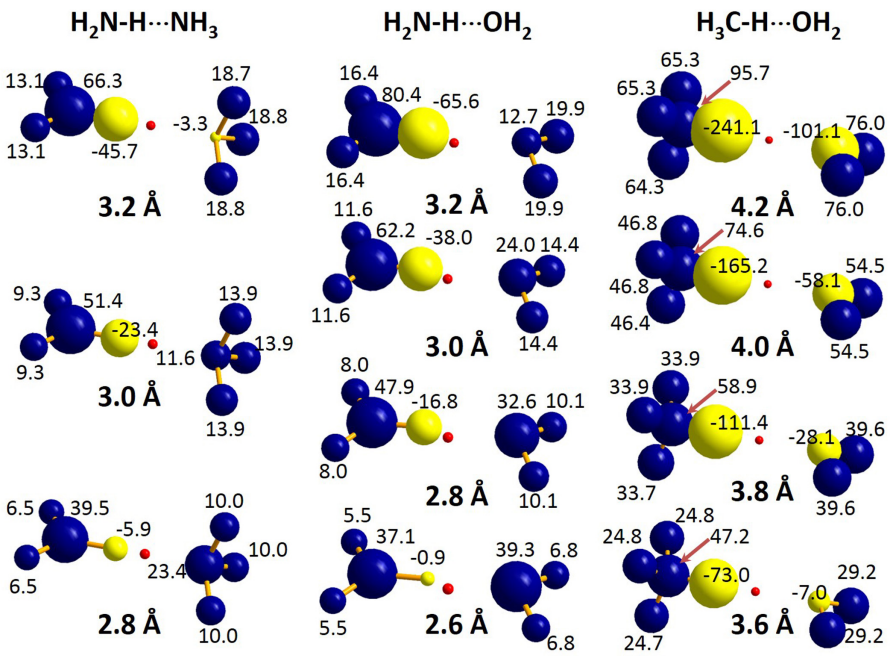


Figure 3.1: source function percentage atomic contributions (SF%) to the electron density at the bond critical point (BCP) of the hydrogen bonds as function of donor-acceptor distance in the models $H_2N\text{-}H\cdots NH_3$, $H_2N\text{-}H\cdots OH_2$ and $H_3C\text{-}H\cdots OH_2$. BCPs are indicated as red dots. Atomic SF% are depicted as sphere whose volume is proportional to the percentage contribution. Blue indicates positive (source) SF%, yellow negative (sink) SF%.

at 3.6 Å). The SF%(H+D+A) contribution is positive for the N-H$\cdots$N and N-H$\cdots$O at any distance and for the C-H$\cdots$O only at the shortest distances. As for the SF%(H), the SF%(H+D+A) increases with the shortening of the donor-acceptor distance. Only for the N-H$\cdots$O hydrogen bond at the shortest distance 2.6 Å the source contribution is close to those usually found for the covalent bonds (75.4 %, close to the usual covalent bond values $\approx$ 80-90 %), while for all the other bonds the sources are much smaller (57.0 % and 41.7 % at the shortest distances for the N-H$\cdots$N and C-H$\cdots$O, respectively).

Ball and stick representations in Figure 3.1 serve as a visual reference to interpret the SF% values found for the DNA base pairs (next Sections). Figure 3.1 shows that all the complexes exhibit a fairly delocalized pattern of sources, where the delocalization degree is inversely proportional to the strength of the hydrogen bonds. The delocalization degree order is then: C-H$\cdots$O > N-H$\cdots$N > N-H$\cdots$O.

3.3 DNA base pairs

3.3.1 Energies and Bond lengths

Figure 3.2 shows the AT and GC base pairs optimized geometry in their neutral and singly ionized anionic (AT- and GC-) and cationic (AT+ and GC+) forms.

Both in AT and GC pairs three hydrogen bonds hold together the purine and the

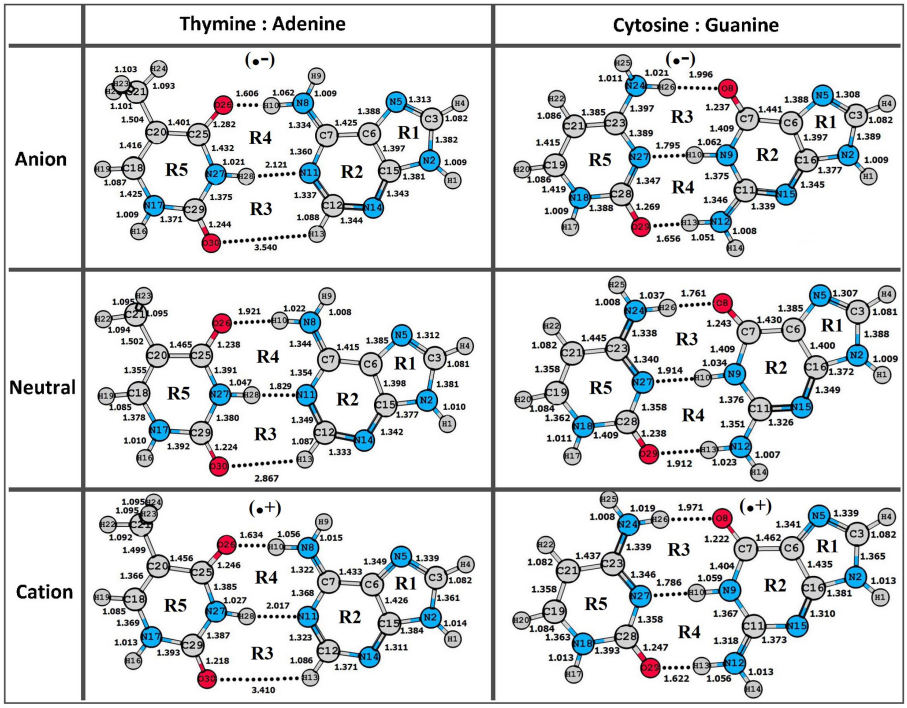


Figure 3.2: optimized structures for the AT and GC base pairs in their neutral and charged ionized forms along with atomic and rings labels. Bond lengths are reported in Å

pyrimidine bases. As already reported at the beginning of this chapter, the former pair have two strong (N-H$\cdots$N and N-H$\cdots$O) and one weak (C-H$\cdots$O) hydrogen bonds, while in the latter pair all the interactions are strong. Asensio and co-workers[165] evaluated the energetic stability given by each of these interactions at the MP2/D95(d,p) level of theory after correction with counterpoise basis set superposition error method[166]. They estimated the strength of single hydrogen bonds equal to the interaction energy of the DNA bases when the latter are disposed in perpendicular planes along the axis of the hydrogen bond of interest.[165] The values obtained by Asensio are similar to what Matta and co-workers found using the topological-based method proposed by Espinosa *et al.*.[153,167]

The two methods predict the total interaction energy of the AT pair equal to 13 (Matta *et al.*[153]) and 15 (Asensio *et al.*[165]) kcal/mol. Both methods confirm that the C-H$\cdots$O is the weakest hydrogen bond and that it contributes only $\approx$ 1 kcal/mol to the total energy.[153,165] The other two bonds present instead higher interaction energies and the two methods agree on the numerical values. The central N-H$\cdots$N accounts for $\approx$ 5-6 kcal/mol and the other N-H$\cdots$O contributes $\approx$ 6 kcal/mol.[153,165]

In the GC pair, both methods predict an interaction energy of the central N-H$\cdots$N bond approximatively equal to 9 kcal/mol. The contribution of the two N-H$\cdots$O lateral hydrogen bonds are predicted to be $\approx$ 11 and 5 kcal/mol by Matta *et al.*, while Asensio *et al.* obtained values $\approx$ 5 and 13 kcal/mol for the same hydrogen bonds, respectively.[153,165]

Tables 3.3 and 3.4 report the donor-acceptor ($d_{D\ldots A}$), donor-hydrogen ($d_{D\text{-}H}$) and hydrogen-acceptor ($d_{H\ldots A}$) bond lengths for the neutral and ionization forms, together

with their absolute variations ΔR and the values of the electron density at the BCPs of AT and GC pairs. A graphical representation of these values are depicted in Figure 3.2.

Table 3.3: adenine-thymine (AT) donor-acceptor ($R_{D\cdots A}$), donor-hydrogen ($R_{D\text{-}H}$) and hydrogen-acceptor ($R_{H\cdots A}$) hydrogen bond distances for neutral, positively charged and negatively charged forms, together with difference of these bond lengths (Δ) between the ionized and the neutral forms and the electron density at the bond critical point (BCP).

Bond	$R_{D\cdots A}$[a]	$R_{D\text{-}H}$[a]	$R_{H\cdots A}$[a]	$\Delta R_{D\cdots A}$[a,b]	$\Delta R_{D\text{-}H}$[a,b]	$\Delta R_{H\cdots A}$[a,b]	$\rho_{BCP}(\boldsymbol{r})$[c]
			N8-H10$\cdots$O26				
AT	2.940	1.022	1.921	//	//	//	0.0262
AT+	2.690	1.056	1.634	-0.250	0.034	-0.287	0.0519
AT-	2.665	1.062	1.606	-0.275	0.040	-0.315	0.0564
			N27-H28$\cdots$N11				
AT	2.877	1.047	1.830	//	//	//	0.0406
AT+	3.036	1.027	2.017	0.159	-0.020	0.187	0.0257
AT-	3.136	1.021	2.121	0.259	-0.026	0.291	0.0209
			C12-H13$\cdots$O30				
AT	3.689	1.087	2.868	//	//	//	0.0043
AT+	4.129	1.086	3.410	0.440	0.001	0.542	no BCP
AT-	4.249	1.088	3.540	0.560	0.002	0.672	no BCP

[a] Distances in Å;
[b] ΔR = R(ionized) - R(neutral);
[c] $\rho_{BCP}(\boldsymbol{r})$ in atomic units (*a.u.*).

In the AT pair, the effects of the ionization process are similar if the system acquires or expels the electron. The central N-H$\cdots$N bond elongates. The donor-acceptor distance moves from 2.877 Å in the neutral form to 3.036/3.136 in the positive and negative ionized species with a total variation of +0.159/0.259 Å. On the contrary, the N-H$\cdots$O hydrogen bond shrinks and the distance $d_{D\cdots A}$ decreases ≈ 0.25/0.28 Å. The major difference that occurs in the ionization process is found for the weak C-H$\cdots$O interaction, where the bond strongly elongates from 3.689 Å in the neutral form to 4.129 and 4.249 Å in the positive and negative ionized forms, respectively. This long elongation leads to the rupture, as it is shown in the next section, of the C-H$\cdots$O bonds.

The ionization process causes similar shortening-stretching phenomenon in the GC pair. Contrary to what found for the AT pair, in this case the central N-H$\cdots$N shrinks by a factor ≈ 0.1 Å. The external N12-H13$\cdots$O29 bond behaves similarly to the central bond and the donor-acceptor distance decreases from 2.935 Å to 2.678 and 2.706 Å for the GC+ and GC- species, respectively, while the other N-H$\cdots$O is weakened as a consequence of an elongation of ≈ 0.2 Å.

The values of the electron density evaluated at the BCP and reported in Tables 3.3 and 3.4 agree with the explained trend. ρ_{BCP} almost doubles when the AT N8-H10$\cdots$O26 and the GC N12-H13$\cdots$O29 shrink, while it increases from 0.0330 *a.u.* in the neutral form to 0.0443 and 0.0438 *a.u.* in the ionized forms for GC central N9-H10$\cdots$N27 bond. Consequently, elongations of the interatomic bonds lead to a decrease of the ρ_{BCP} values.

Table 3.4: guanine:cytosine (GC) donor-acceptor ($R_{D\cdots A}$), donor-hydrogen ($R_{D\text{-}H}$) and hydrogen-acceptor ($R_{H\cdots A}$) hydrogen bond distances for neutral, positively charged and negatively charged forms, together with the difference of these bond lengths (Δ) between the ionized and the neutral forms and the electron density at the bond critical point (BCP)..

Bond	$R_{D\cdots A}$[a]	$R_{D\text{-}H}$[a]	$R_{H\cdots A}$[a]	$\Delta R_{D\cdots A}$[a,b]	$\Delta R_{D\text{-}H}$[a,b]	$\Delta R_{H\cdots A}$[a,b]	$\rho_{BCP}(\boldsymbol{r})$[c]
			N12-H13$\cdots$O29				
GC	2.935	1.022	1.913	//	//	//	0.0268
GC+	2.678	1.056	1.622	-0.257	0.034	-0.291	0.0538
GC-	2.706	1.051	1.656	-0.229	0.029	-0.257	0.0511
			N9-H10$\cdots$N27				
GC	2.947	1.034	1.914	//	//	//	0.0330
GC+	2.844	1.059	1.786	-0.103	0.025	-0.128	0.0443
GC-	2.857	1.062	1.795	-0.090	0.028	-0.119	0.0438
			N24-H26$\cdots$O8				
GC	2.798	1.038	1.761	//	//	//	0.0387
GC+	2.984	1.019	1.971	0.186	-0.019	0.210	0.0235
GC-	3.017	1.021	1.996	0.219	-0.017	0.235	0.0236

[a] Distances in Å;
[b] ΔR = R(ionized) - R(neutral);
[c] $\rho_{BCP}(\boldsymbol{r})$ in atomic units (*a.u.*).

3.3.2 Source Function Contributions at Hydrogen Bond BCPs

The variations on the bond lengths that occur because of the ionization processes is mirrored in the behaviour of the source function percentage values. In Tables 3.5 and 3.6 the SF% values evaluated at the BCPs of the hydrogen bonds in AT and GC base pairs are reported.

The SF% contributions from the hydrogen atoms clearly highlight the trend already seen for the intermolecular interactions. In both the neutral DNA pairs the SF%(H) assume negative values for all the hydrogen bonds. The hydrogen atoms always behave as sinks of electron density. In the AT pair, SF%(H) ranges from -137.8 % in the weak C12-H13$\cdots$O30 interaction to -3.4 % for the central N27-H28$\cdots$N11. In the GC, the situation is more compact and ranges from -23.9 % for the N12-H13$\cdots$O29 to -7.1 % for the N24-H26$\cdots$O8. When the bonds elongate upon the uptake or removal of an electron, the source function contributions from the hydrogen atoms become more negative and their sink-effect is increased. The SF%(H) value decreases from -3.4 % to -34.4 % when the stretching of the AT central N27-H28$\cdots$N11 bond is considered. Similarly, in the N24-H26$\cdots$O8 case the source contributions of the hydrogen reach the values of -30.0 % and -28.7 % for the GC+ and GC- forms, respectively. On the contrary, when the interaction is strengthened the SF%(H) increases and becomes less negative, almost null or largely positive depending on the considered BCP. For instance, the shrinking of the N8-H10$\cdots$O26 in AT is associated with a growth of the SF%(H) from -24.8 % in the neutral form to 5.5 % in the negative charged one with a total increase of $\approx$ 30 percentage

Table 3.5: Source function percentage (SF%) contributions to the electron density at the hydrogen bond BCP in the adenine:thymine base pair in its neutral (AT) and ionized (AT+ and AT-) forms.

Bond	$R_{D\cdots A}$[a]	SF%(H)[b]	SF%(H+A)[b]	SF%(H+D+A)[b]	$SF\%_{Ade}$[b]
			N8-H10$\cdots$O26		
AT	2.940	-24.8	-3.6	40.8	47.3
AT+	2.690	3.0	38.7	67.1	46.3
AT-	2.665	5.5	43.2	69.7	46.4
			N27-H28$\cdots$N11		
AT	2.877	-3.4	13.9	40.5	53.6
AT+	3.036	-23.4	-23.5	13.7	51.8
AT-	3.136	-34.4	-43.8	-1.4	52.8
			C12-H13$\cdots$O30		
AT	3.689	-137.8	-217.2	-178.5	47.1
AT+	4.129	//	//	//	//
AT-	4.249	//	//	//	//

[a] Distances in Å;
[b] Source contributions from: SF%(H) - hydrogen atom; SF%(H+A) - hydrogen and acceptor atoms; SF%(H+D+A) - hydrogen, donor and acceptor atoms; $SF\%_{Ade}$ - adenine molecule.

Table 3.6: Source function percentage (SF%) contributions to the electron density at the hydrogen bond BCP in the guanine:cytosine base pair in its neutral (GC) and ionized (GC+ and GC-) forms.

Bond	$R_{D\cdots A}$[a]	SF%(H)[b]	SF%(H+A)[b]	SF%(H+D+A)[b]	$SF\%_{Gua}$[b]
			N12-H13$\cdots$O29		
GC	2.935	-23.9	-2.0	42.0	46.2
GC+	2.678	3.8	39.8	67.5	46.7
GC-	2.706	2.0	37.0	65.9	46.8
			N9-H10$\cdots$N27		
GC	2.947	-12.5	-2.7	28.7	45.7
GC+	2.844	-0.1	18.7	44.1	46.6
GC-	2.857	-0.1	19.0	44.2	46.6
			N24-H26$\cdots$O8		
GC	2.798	-7.1	23.7	57.6	52.5
GC+	2.984	-30.0	-12.2	35.5	51.8
GC-	3.017	-28.7	-10.7	36.4	52.3

[a] Distances in Å;
[b] Source contributions from: SF%(H) - hydrogen atom; SF%(H+A) - hydrogen and acceptor atoms; SF%(H+D+A) - hydrogen, donor and acceptor atoms; $SF\%_{Gua}$ - guanine molecule.

units. This variation is less extended in the central N9-H10$\cdots$N27 hydrogen bond where the source contribution in GC is equal to -12.5 % and in GC+/GC- reaches the value of -0.1 %.

Another feature connected to the shrinking/elongation of the donor-acceptor distances is the localization/spread of the source contributions. The SF%(H+D+A) source percentage values clearly summarize quantitatively these evidences. In the two strong hydrogen bonds that shorten in the AT and GC pairs, namely N8-H10$\cdots$O26 and N12-H13$\cdots$O29, the atoms triad contribution at the BCP increases from $\approx$ 41-42 % in the neutral form to 66-70 % in the ionized case. The strengthening of these bonds are then associated to a increased localization of the sources, with a consequent reduction of the other atomic percentage contributions. On the contrary, the bond elongation is followed by the decrease of the SF%(H+D+A) value. For instance, at the N24-H26$\cdots$O8 BCP in the GC pair, the SF%(H+D+A) reduces by almost 20 %. In some cases, like for the central N27-H28$\cdots$N11 bond in AT, the triad contribution decreases from the positive 40.4% in the neutral AT down to -1.4% in the AT- form. Negative values for the SF%(H+D+A) source reveal a completely delocalized situation, where the atoms directly involved in the bond are not responsible for the electron density at the BCP.

In general, hydrogen bonds in AT and GC show a fairly delocalized source. The highest SF%(H+D+A) value of 69.7 %, that is the most localized situation, is found for the AT-NH-H10$\cdots$O26 bond, which is far from the usual values found for the covalent bonds ($\approx$ 80-90 %.) On the other hand, the most delocalized situation is reached in the AT C12-H13$\cdots$O30 bond, where the (H+D+A) contribution assumes a value of -178.5 %. In this case, the three atoms globally behave as sinks of electron density and they counteract the role of the other atoms in the reconstruction of the density at the bond critical point.

Another interesting evidence from Tables 3.5 and 3.6 is that the total source contributions from the purine molecules (in turn adenine or guanine) are similar and they are always close to 50% within $\pm$4 %. These outcomes hold true for both AT and GC pairs and they are independent from the strength of the hydrogen bond considered. For instance, the contribution from the adenine base to the strong N8-H10$\cdots$O26 bond BCP is 47.3 % in the neutral form, and similarly the same quantity for the weak C12-H13$\cdots$O30 bond BCP is 47.1 %. This reveals that there is no clear dominant monomer type responsible for the binding in the dimer, neither in the neutral nor in the ionized forms. However, as already found for the hydrogen bond models, the slightly minor contributions always belong to the donor molecule. For instance, in the GC pair the bonds where the guanine (purine) base acts as a donor, namely the N12-H13$\cdots$O29 and the central N9-H10$\cdots$N27 bonds, are associated with $SF\%_{Gua}$ lower than 50 %, while in N24-H26$\cdots$O8, the guanine is the acceptor molecule and its total source contribution is greater than 50 %.

Figures 3.3 and 3.4 depict and summarize the atomic sources and the data reported in Tables 3.5 and 3.6 with the classical ball-and-stick model.

These figures give a clearer evidence of the localization/delocalization effects connected to the hydrogen bonds shrinking/elongation. For instance, in the AT N27-H28$\cdots$N11

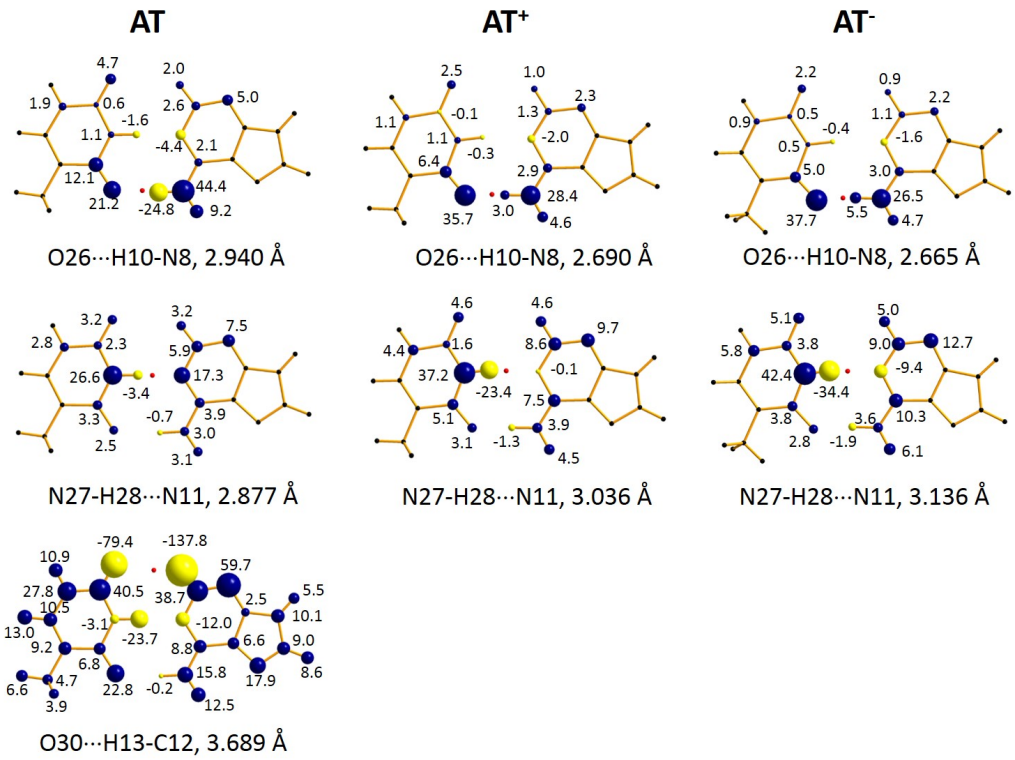


Figure 3.3: source function percentage atomic contributions (SF%) to the electron density at the bond critical point (BCP) of the hydrogen bonds in the adenine:thymine base pair in its neutral (AT) and ionized (AT+ and AT-) forms. BCPs are indicated as red dots. Atomic SF% are depicted as sphere whose volume is proportional to the percentage contribution. Blue indicates positive (source) SF%, yellow negative (sink) SF%. The orientation of the dimer is rotated 180° along the horizontal central axis with respect to Figure 3.2.

bond, it is evident that the contributions from distant atoms are larger in the ionized forms than in the neutral form. On the other hand, an average decrease of the source values belonging to atoms not directly involved in the hydrogen bonds is found in the AT N8-H10$\cdots$O26 bond. A more evident difference in the delocalization spread of the source contributions is observed comparing the scenario for the two strong hydrogen bonds (N8-H10$\cdots$O26 and N27-H28$\cdots$N11) and the weak one (C12-H13$\cdots$O30) in the AT pair. In the latter, the interatomic bond has a larger delocalized nature which is evident from the large atomic contributions (large spheres) from all the atoms of the system. In this case, the lowest atomic source, in absolute value, does not fall below 4 % except for one hydrogen and one carbon atom. The quite low electron density evaluated at the BCP of this bond is due to the large but opposite source contributions from the atoms that counteract each other. It is interesting to notice that for all the pairs, the majority of the atoms contributes positively to the reconstruction of the electron density while only few atoms close to the intermolecular region behave as sinks of electrons and counteract the effect of the others.

Comparison between the SF data found for the DNA base pairs reported in Figures

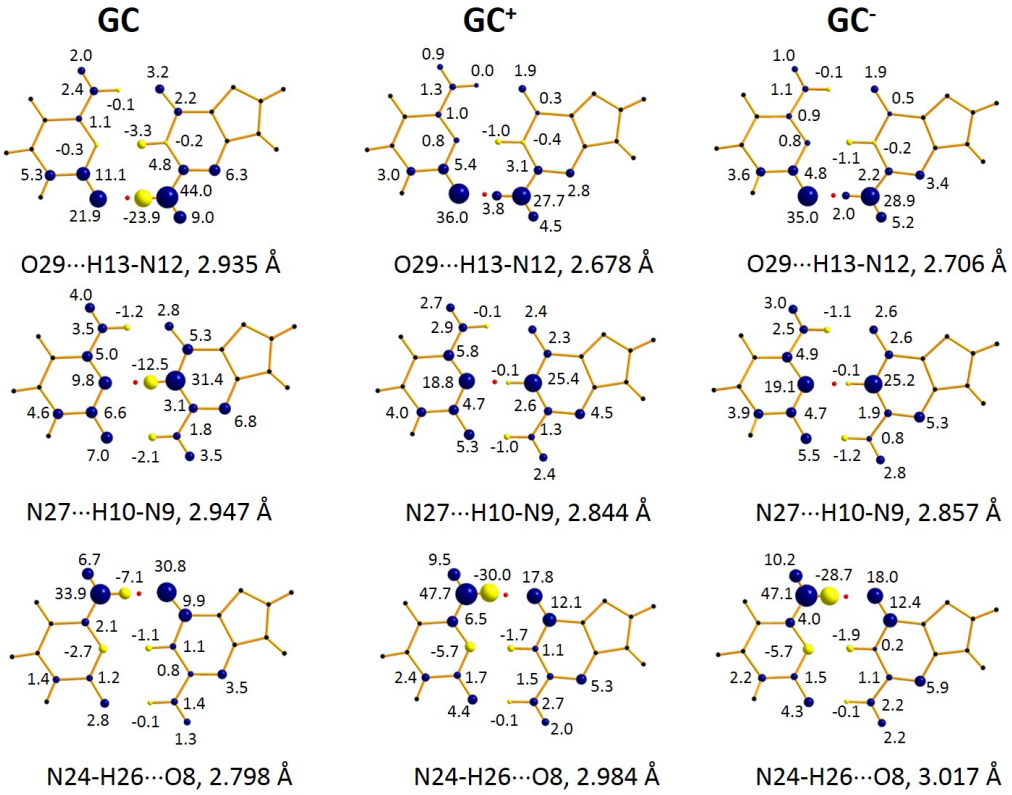


Figure 3.4: source function percentage atomic contributions (SF%) to the electron density at the bond critical point (BCP) of the hydrogen bonds in the guanine:cytosine base pair in its neutral (GC) and ionized (GC+ and GC-) forms. BCPs are indicated as red dots. Atomic SF% are depicted as sphere whose volume is proportional to the percentage contribution. Blue indicates positive (source) SF%, yellow negative (sink) SF%. The orientation of the dimer is the same of Figure 3.2

3.3-3.4 and Tables 3.5-3.6 with those obtained for the model systems and shown in Figure 3.1 and Table 3.2 highlights important differences. The first variation concerns the donor-hydrogen distance. In the N-H$\cdots$N and N-H$\cdots$O models, $d_{D\text{-}H}$ assumes a maximum value of 1.010 Å, while in the DNA base pairs the same values are always larger than 1.019 Å (lowest value in GC N23-H26$\cdots$O8) and can reach values as large as 1.062 Å (highest values in AT- N8-H10$\cdots$O26 and GC- N9-H10$\cdots$N27). The weak C-H$\cdots$O bond has instead an opposite trend, where $d_{D\text{-}H}$ in the model systems is slightly larger ($d_{D\text{-}H}$(model)=1.090-1.090 Å) than in the DNA pairs ($d_{D\text{-}H}$(DNA)=1.086-1.088 Å). The SF%(H) contributions for the DNA pairs (Tables 3.5 and 3.6) are less negative or more positive than in the model systems (Table 3.2) at comparable donor-acceptor distance. N-H$\cdots$N and N-H$\cdots$O hydrogen bonds are clearly strengthened in AT and GC with respect to the models. The same trend is observed for the SF%(H+A), where the values for the AT and GC pairs are generally higher than for the model complexes at similar $R_{D\ldots A}$ distances. On the contrary, SF%(H+D+A) are always lower in the DNA systems than in the proposed models. This is specially true if one considers the central N-H$\cdots$N bonds in both AT and GC pairs. In the AT system, the SF%(H+D+A) for the neutral form is 40.5 % ($d_{D\ldots A}$

= 2.877 Å) while the same value for the H_2N-H$\cdots NH_3$ model at $d_{D\cdots A}$ = 2.8 Å is 57.0 %. Similarly, the source function contribution for the same hydrogen bond associated with AT- is negative and equal to -1.4 %, while the analogous in the model is as high as 17.3 %. In the GC central N9-H10$\cdots$N27 bond the situation is the same, where the SF%(H+D+A) values for GC (28.7 %) and GC- (44.2 %) have to be compared with the reference contributions of 36.6 % and 57.0%, respectively. The fact that SF%(H+D+A) contributions in the DNA pairs are always lower than in the reference models is another sign of the enhanced delocalization of these bonds. The different delocalization can be appreciated comparing Figures 3.3 and 3.4 with 3.1. Regarding the C-H$\cdots$O bond in AT, it is largely weakened with respect to the model system. This is evident looking at the very negative SF%(H) contribution (-137.8 %) and comparing it with the same value predicted for the model (-72.9 % at 3.6 Å donor-acceptor distance). SF%(H+A) and SF%(H+D+A) highlights the same behaviour, being more negative in the AT pair than in the H_3C-H$\cdots OH_2$ model, where the SF%(H+D+A) becomes even positive. The same evidences can be obtained comparing the ρ_{BCP} values in AT (ρ_{BCP} = 0.0043 *a.u.*) and in the model (ρ_{BCP} = 0.0059 *a.u.*). Eventually, this leads to the lack of the BCP in the ionized AT systems at donor-acceptor distances where a BCP is present in the H_3C-H$\cdots OH_2$ model.

A coarser scenario can be studied evaluating the source function contributions from the atoms belonging to specific ring in the system, as labelled in Figure 3.2, to the reconstruction of the electron density at the BCP of the hydrogen bonds. The numerical values are reported in Tables 3.7 and 3.8.

Table 3.7: source function percentage (SF%) ring contributions to the electron density at the hydrogen bond BCP in the adenine:thymine base pair in its neutral (AT) and ionized (AT+ and AT-) forms.

Bond	$R_{D\cdots A}$[a]	SF%(R1)[b]	SF%(R2)[b]	SF%(R3)[b]	SF%(R4)[b]	SF%(R5)[b]
			N8-H10$\cdots$O26			
AT	2.940	6.4	10.0	4.8	53.1	21.4
AT+	2.690	3.3	5.7	3.6[c]	75.2	11.4
AT-	2.665	2.9	5.7	3.2[c]	76.1	10.1
			N27-H28$\cdots$N11			
AT	2.877	7.4	37.8	53.9	54.6	38.0
AT+	3.036	11.4	28.6	33.0[c]	32.1	53.2
AT-	3.136	13.3	26.3	21.5[c]	17.3	62.2
			C12-H13$\cdots$O30			
AT	3.689	48.2	106.1	-188.1	26.6	88.5
AT+	4.129	//	//	//	//	//
AT-	4.249	//	//	//	//	//

[a] Distances in Å;
[b] SF%(RX), with X=1-5: source function contributions from atoms belonging to ring RX (see Figure 3.2 for rings labels);
[c] R3 is not a topological ring because of C12-H13$\cdots$O30 bond rupture.

Table 3.8: source Function percentage (SF%) ring contributions to the electron density at the hydrogen bond BCP in the guanine:cytosine base pair in its neutral (GC) and ionized (GC+ and GC-) forms.

Bond	$R_{D\cdots A}$[a]	SF%(R1)[b]	SF%(R2)[b]	SF%(R3)[b]	SF%(R4)[b]	SF%(R5)[b]
			N12-H13$\cdots$O29			
GC	2.935	5.6	10.5	4.6	50.2	21.3
GC+	2.678	3.0	6.2	3.9	75.5	12.4
GC-	2.706	3.0	6.4	3.9	72.5	12.4
			N9-H10$\cdots$N27			
GC	2.947	6.9	44.5	43.8	42.3	33.0
GC+	2.844	5.1	35.5	56.9	57.1	37.9
GC-	2.857	5.2	36.0	55.5	55.7	37.6
			N24-H26$\cdots$O8			
GC	2.798	5.4	14.9	66.4	2.7	6.8
GC+	2.984	8.0	22.2	47.8	3.8	10.4
GC-	3.017	8.8	22.7	45.4	1.6	8.3

[a] Distances in Å;
[b] SF%(RX), with X=1-5: source function contributions from atoms belonging to ring RX (see Figure 3.2 for rings labels);

As expected, both rings R3 and R4 play an important role in the definition of the electron density at the BCP of the central hydrogen bonds in both AT and GC (namely N27-H28$\cdots$N11 and N9-H10$\cdots$N27, respectively). The effect of elongation and shortening of these bonds is also reflected symmetrically in the SF%(R3) and SF%(R4) values, where they increase when the donor-acceptor distance becomes short in the ionized form and *vice versa* when it elongates. It is worth noting that rings R2 and R5 also yield important contributions for these bonds (for instance 37.8 % and 38.0 % in the AT form, respectively), denoting the peculiar delocalized nature of the central atoms. Ring R1 plays a minor but not negligible role. On the contrary, the external N-H$\cdots$O bonds are mainly determined by the source from R3 or R4 ring. N8-H10$\cdots$O26 in AT and N12-H13$\cdots$O29 in GC belong to R4 ring and they are dominated by its source contribution, while SF%(R3) dominates the N24-H26$\cdots$O8 bond in GC. R5 or R2 sources are the second dominant ones, depending if the major contribution comes from R4 or R3 ring. The very weak C-H$\cdots$O in AT has a particular source distribution and shows a very delocalized situation. The ring to which it belongs, R3, contributes with a large negative source (SF%(R3) = -188.1 %) which is overbalanced by large positive contributions from the other rings, in the order R2 > R5 > R1 > R4. It is worth to note that the sum of the all ring contributions to a BCP is not equal to 100 %, because the contributions from atoms shared by two rings are counted twice and the hydrogen atoms not involved in the interaction are not taken into account.

3.3.3 Source Function Partial Reconstructed Electron Density

Source Function partial reconstructed maps are useful tools, as already seen in Chapter 2, to get interesting insights. These maps have the advantage with respect to the classical SF

analysis to give a wider and easy-to-grasp picture about the atomic contributions in the definition of the density in a plane or in the space. In this study, SF partial reconstructed maps are able to highlight details of the localized/delocalized nature of the hydrogen bonds in the DNA base pairs.

Figure 3.5 depicts the partial reconstructed densities for the AT and GC pairs in the least-square plane of the bases.

The density reconstruction was performed considering four subsets of atoms: all the atoms ($\rho\{\text{all}\}$), R3 ring ($\rho\{\text{R3}\}$), R4 ring ($\rho\{\text{R4}\}$), R3 and R4 together ($\rho\{\text{R3+R4}\}$).

The maps obtained using the full set of atoms (Figures 3.5(a) and 3.5(e)) are indistinguishable, except for very small but negligible differences due to numerical error, from the primary ρ map directly obtained from the wavefunctions.

Figure 3.5 highlights three important points. The first consideration concerns the influence of the R3 ring sources to the electron density in the region of the external hydrogen bond in ring R4 and *vice versa* (second and fourth rows, Figure 3.5).

In the AT $\rho\{\text{R3}\}$ map (Figure 3.5(c)), the positive source yielded by the atoms of the ring R3 is largely delocalized and it is able to reach the region of the N8-H10$\cdots$O26 bond (in the bottom part of the map). One may be initially surprised to see that in the AT partial reconstruction (Figure 3.5) the atoms belonging to the ring R3 are able to partially reconstruct the density located along the central N-H$\cdots$N bond, while they can not do the same with the external hydrogen bond involved in the ring. This fact can be easily explained if one remember that this bond is the weak C12-H13$\cdots$O30 and that is extremely delocalized in nature. GC $\rho\{\text{R3}\}$ map (Figure 3.5(g)) shows more or less the same behaviour of the AT $\rho\{\text{R3}\}$ except for the partial reconstruction of the external hydrogen bond belonging to the R4 ring which is evident here.

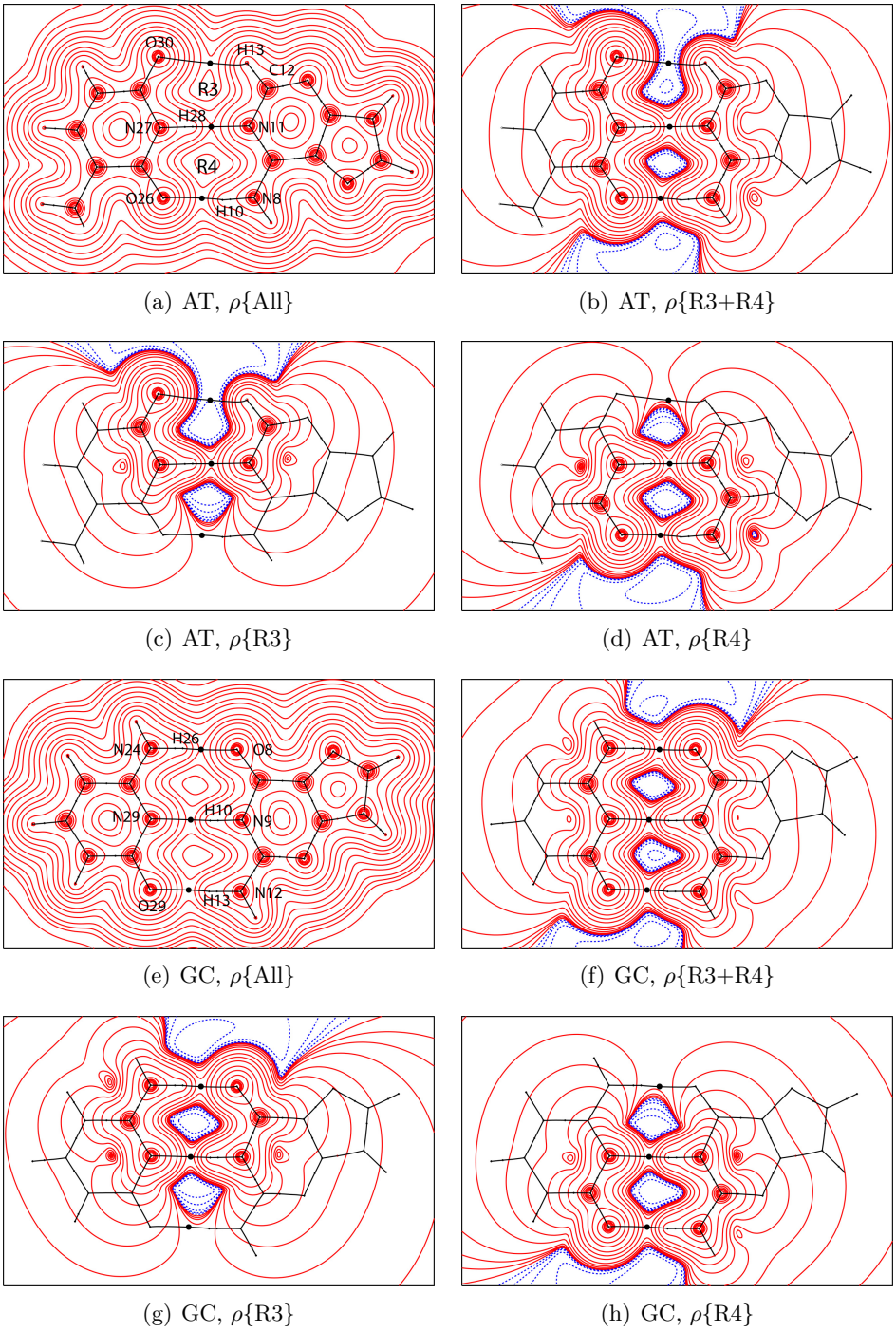
The second evidence is that in the $\rho\{\text{R3+R4}\}$ maps (Figures 3.5(b) and 3.5(f)) there are two negative electron density holes in the middle of the R3 and R4 rings. These regions become positive (as in the AT and GC $\rho\{\text{all}\}$ map) only when the contributions from the other rings (R1, R2 and R5) are taken into account. These negative regions are also present in the maps reconstructed using the R3 or R4 ring atoms.

The third result is the evident reductions of the electron density along the hydrogen bonds when only the contributions from R3 and R4 are considered with the respect to the total density.

Once again, these maps highlight the largely delocalized nature of the hydrogen bonds in DNA pairs, revealing that distant atoms play an important role not only in the reconstruction of the density at the BCP (as seen before) but in all the intermolecular space.

Similar results can be seen in Figures 3.6 and 3.7 for the ionized complexes of AT and GC base pairs, though geometry changes, due to electron uptake or detachment, lead to visible variations.

Another subset, composed by the atoms involved in the central N-H$\cdots$N hydrogen bond in AT (N27, H28 and N11), is considered and it is reported in Figure 3.8 in the neutral and ionized complexes. The three maps in Figure 3.8 are qualitatively similar, yet



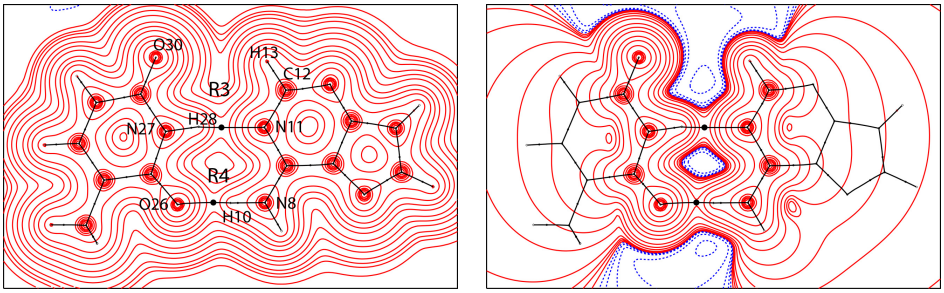
(a) AT, $\rho\{\text{All}\}$ (b) AT, $\rho\{\text{R3+R4}\}$

(c) AT, $\rho\{\text{R3}\}$ (d) AT, $\rho\{\text{R4}\}$

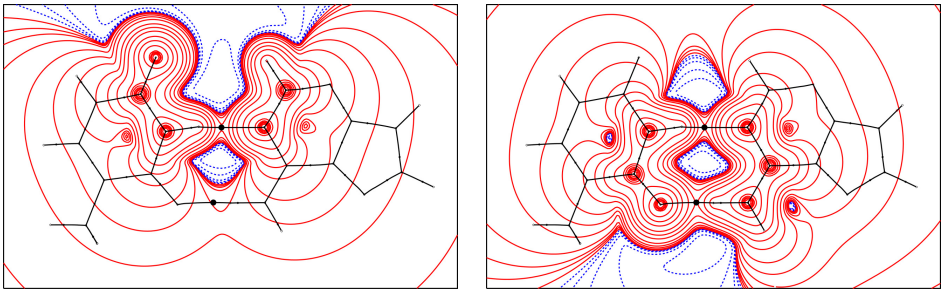
(e) GC, $\rho\{\text{All}\}$ (f) GC, $\rho\{\text{R3+R4}\}$

(g) GC, $\rho\{\text{R3}\}$ (h) GC, $\rho\{\text{R4}\}$

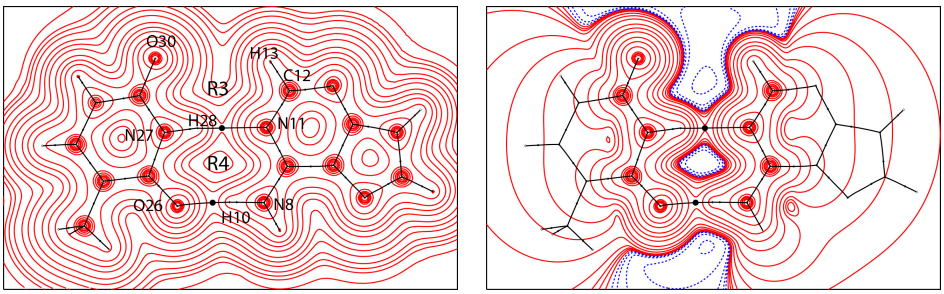
Figure 3.5: adenine:thymine (AT) and guanine:cytosine (GC) neutral complexes: source function partial reconstructed densities in the least-square molecular plane for four atomic subsets: {All}: all the atoms but the hydrogen atoms; {R3+R4}: atoms belonging to rings R3 and R4; {R3}: atoms belonging to ring R3; {R4}: atoms belonging to ring R4. Bond paths are shown and the BCPs are denoted as full black dots. Solid red/dotted blue lines indicate positive/negative contour values. Contours are drawn at interval of $\pm(2,4,8) \cdot 10^{-n}$, $-3 \leq n \leq 0$ *a.u.*. The orientation of the dimer is rotated 180° along the horizontal central axis with respect to Figure 3.2.



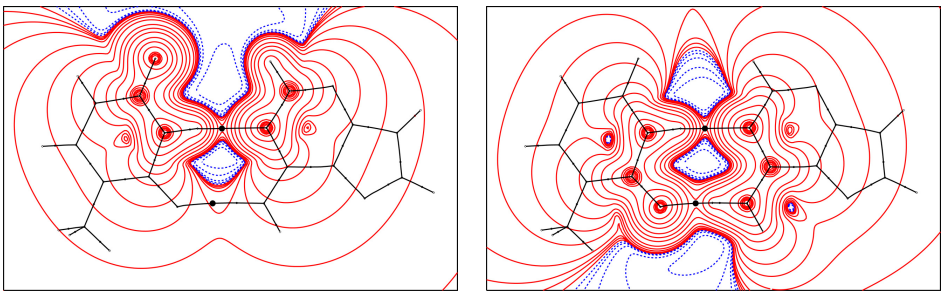
(a) AT+, $\rho\{\text{All}\}$ (b) AT+, $\rho\{\text{R3+R4}\}$



(c) AT+, $\rho\{\text{R3}\}$ (d) AT+, $\rho\{\text{R4}\}$

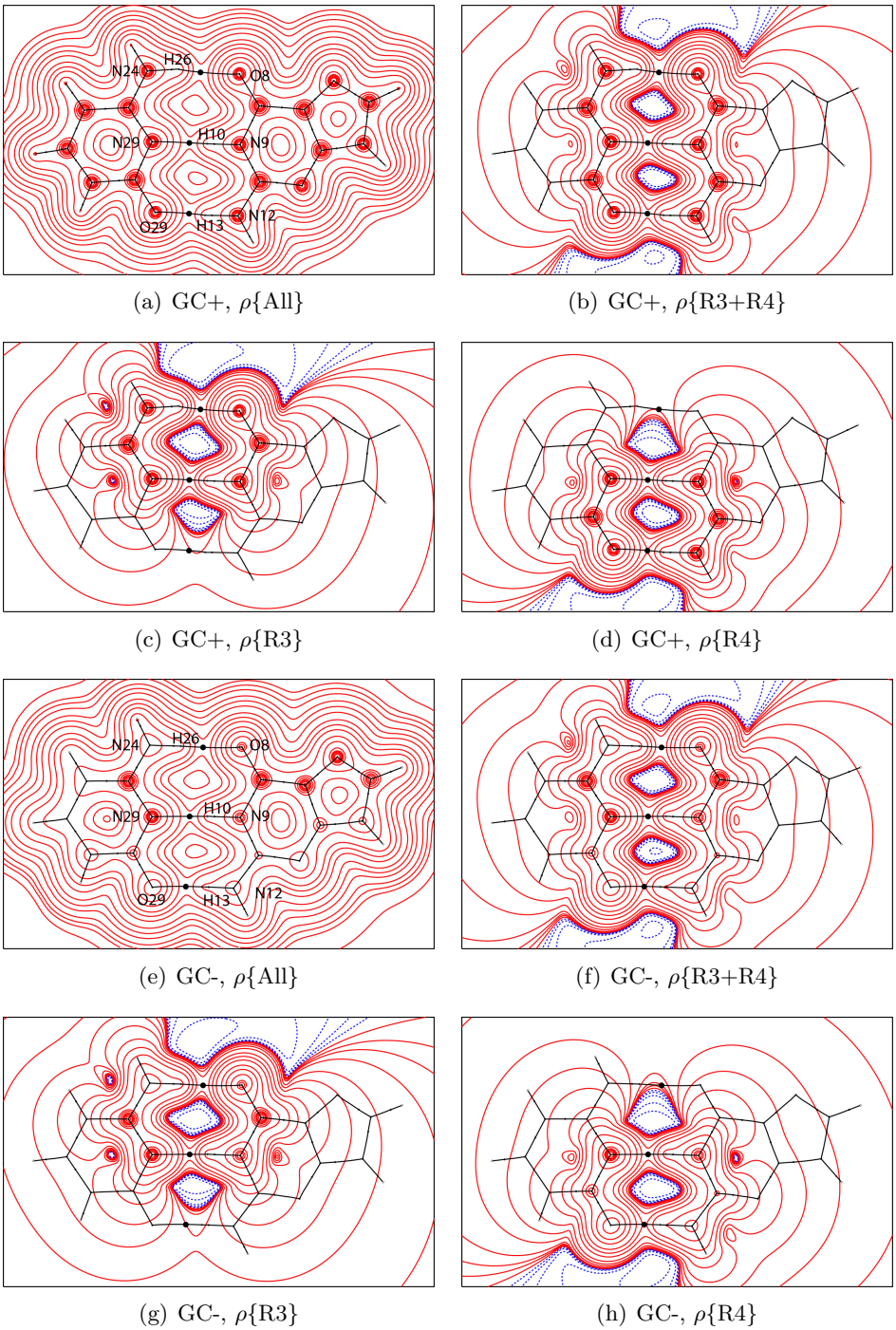


(e) AT-, $\rho\{\text{All}\}$ (f) AT-, $\rho\{\text{R3+R4}\}$



(g) AT-, $\rho\{\text{R3}\}$ (h) AT-, $\rho\{\text{R4}\}$

Figure 3.6: adenine:thymine cationic (AT+) and anionic (AT-) ionized complexes : source function partial reconstructed densities in the least-square molecular plane for four atomic subsets: {All}: all the atoms but the hydrogen atoms; {R3+R4}: atoms belonging to rings R3 and R4; {R3}: atoms belonging to ring R3; {R4}: atoms belonging to ring R4. Bond paths are shown and the BCPs are denoted as full black dots. Color code and isovalue contours are the same reported in Figure 3.5. The orientation of the dimer is rotated 180° along the horizontal central axis with respect to Figure 3.2.



(a) GC+, ρ{All}

(b) GC+, ρ{R3+R4}

(c) GC+, ρ{R3}

(d) GC+, ρ{R4}

(e) GC-, ρ{All}

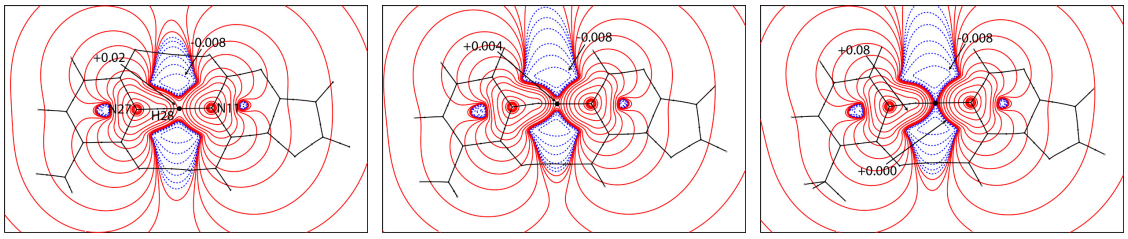
(f) GC-, ρ{R3+R4}

(g) GC-, ρ{R3}

(h) GC-, ρ{R4}

Figure 3.7: guanine:cytosine cationic (GC+) and anionic (GC-) ionized complexes: source function partial reconstructed densities in the least-square molecular plane for four atomic subsets: {All}: all the atoms but the hydrogen atoms; {R3+R4}: atoms belonging to rings R3 and R4; {R3}: atoms belonging to ring R3; {R4}: atoms belonging to ring R4. Bond paths are shown and the BCPs are denoted as full black dots. Color code and isovalue contours are the same reported in Figure 3.5. The orientation of the dimer is the same as in Figure 3.2.

they reveal the bond weakening on going from the neutral AT to the positively charged AT+ and to the negatively charged AT- complex. The two negative regions, perpendicular to the central bond and present in the ring R3 and R4 in AT, approach each other in AT+ and almost merge together in AT-. If one compares these maps with those reconstructed using the R3+R4 subset of atoms (Figures 3.5 to 3.7), it is evident that also the other atoms of the two rings have a relevant role in the reconstruction of the density along this bond, providing further evidences for the quite delocalized nature of such hydrogen bond interaction.



(a) AT, ρ{N27+H28+N11} (b) AT+, ρ{N27+H28+N11} (c) AT-, ρ{N27+H28+N11}

Figure 3.8: adenine:thymine neutral (AT) and charged (AT+ and AT-) complex: source function partial reconstructed densities in the least-square molecular plane for the triad of atoms involved in the central N-H$\cdots$N (N27, H28 and N11, {N27+H28+N11}). Bond path are shown and the BCP are denoted as full black dots. Color code and isovalue contours are the same reported in Figure 3.5. The orientation of the dimer is rotated 180 deg along the horizontal central axis with respect to Figure 3.2.

3.4 Conclusions

The DNA is the most important biopolymer because of its importance in the replication and transmission of the genetic code of all the living beings. For this reason, it is not surprising that tons of experimental and theoretical studies have been carried out since its discovery. Nevertheless, the present chapter and the work presented in Reference [106] can be considered as an unique study aimed to shed light on three important aspects of the DNA base interactions. The first point is to understand which atoms are responsible for most of the binding through the hydrogen bond dimers. In other words, we tried to understand the very intimate atomic nature of the interactions between the purine and pyridine bases. The second relevant point we focused on was to reveal the presence of possible dominant atoms, rings or monomer that drive the binding process. As a last point, we aimed to understand the effect of ionization on the atomic or group contributions. To pursuit this goals, we used the source function topological descriptor. The source function analysis provides a unique model-free decomposition of the total electron density in the system as a sum of atomic contributions. The classic "ball-and-stick" source function representation is useful to get insight on the co-operation of the atoms to reconstruct the electron density in the hydrogen bond BCP, which is taken as the most representative point of the interaction. On the other hand, the new source function partial reconstructed density

maps are able to give a more delocalized picture of the role played by the atom and they are useful to understand features such as the electron density localization/delocalization. The source function also goes beyond the picture of the electron density as a local scalar function and express it as an effect of a cause, the electron density Laplacian, distributed in all the space. In other words, the source function descriptor provides a way to discuss electron delocalization both from theoretical and experimental electron density distributions.[21,63,65,156] Applied to the study of the hydrogen bond interactions in Watson-Crick base pairs, the source function has revealed that the intermolecular interactions have a large delocalized nature and that distant atoms usually play a relevant role in the strengthening of these bonds. Moreover, the SF has shown that the ionization process can have little geometrical net effects on some hydrogen bonds that hide subtle but significant interplays of sources and sinks behind the scenes. The contribution from the hydrogen atom directly involved in the bond is extremely sensitive to the donor-acceptor distance variations and can be used as a detector for the variation in the strength of the bond. On the other hand, the contribution from the triad donor-hydrogen-acceptor can highlight localization/delocalization effects as a function of the hydrogen bond length. Evolution may have fine tuned the atomic composition of the bases to minimize the possible undesired effects related to ionization process on the hydrogen bond stability. Both the analysis of the BCP source contributions and the partial reconstructed density maps have revealed that purine and pyrimidine bases play a concerted role and equally determine the strength of the hydrogen bonds, independently on the nature (neutral or singly ionized) of the complex considered. These outcomes, besides revealing this fine cooperative effects among the atoms, confirm once more the intrinsic potentiality of the source function descriptor to provide further precious insights on the chemical bonding or on the spin transmission mechanism[68,69], as shown in the previous chapter.

Chapter 4

Full Population Analysis through the Source Function Descriptor

As briefly anticipated in Section 1.10, a *full population analysis* can be performed through the use of the source function descriptor and QTAIM partitioning.

In this chapter, the latest breakthroughs concerning the *full population analysis* based on the source function tool is shown. Since this project is still in the early stage and further developments of the program and of the method used are on the road, I briefly present what it has been done since this moment. In the first part, the main advantages and drawbacks are presented. Then, in the second part I describe the approaches used to tackle the critical features and to improve them in the case of gas-phase molecules. The application of this analysis to solid-state is going to be implemented only once it has been proved reliable on isolated molecules. Finally, few examples of application are presented.

4.1 Properties of the Source Function Full Population Matrix

The *full population matrix* obtained using the source function descriptor and the Bader's definition of atom has some peculiar advantages with respect to other analyses.

First, it is a electron density-based method and the only functions needed are the electron density (ED) and its Laplacian. This fact implies that this population matrix is amenable to experimental determination, provided that accurate enough data are available.

The second advantage is that QTAIM atoms are quantum mechanical objects. As already seen in Chapter 1, this allows to define a molecular property as a sum of atomic ones and all the quantum mechanics laws are still valid for the QTAIM atoms. Another advantage associated to QTAIM is that it divides the space in an exhaustive way, that is each point of the space is assigned to a nuclear (or non-nuclear in some case) attractor.

The source function population analysis retrieves not-symmetric off-diagonal terms in the matrix. As already stated in the introduction chapter, this is extremely important to

maintain the asymmetry of the interaction when two different atoms (different elements or different chemical environment) interact. This property is not found, for instance, in the simple Mulliken's analysis.

As a last advantage, it is worth to remind that the source function descriptor is able to reconstruct the electron density in any point within a *cause-effect* relation. This unique feature allows one to see the matrix elements directly in term of the ability of one basin to reconstruct the density within another one.

4.2 Numerical Drawbacks

Even though the SF population analysis has all of the advantages showed in the previous section, it also has some important numerical drawbacks that has limited its application since now.

First, the computational cost of this analysis is extremely high. The reasons of this has to be looked for in its definition and they are evident analysing Equation 1.36 in Section 1.10. For each atom, the calculation of the source function contributions at one reference point (RP) $\boldsymbol{r}$ requires the evaluation of the electron density Laplacian values and of the distances between the chosen RP and the grid points within the basin $\boldsymbol{r'}$ used for the integration. Assuming that the number of integration points within a basin are equal to K, the software has to calculate K times the Laplacian value and the distance $|\boldsymbol{r} - \boldsymbol{r'}|$ for assessing the contribution of this atom.

To obtain a population matrix, one needs to reconstruct the electron density through the source function on a large three-dimensional set of points Q. For each of these Q points, the software needs to evaluate K times per atom the distance $|\boldsymbol{r} - \boldsymbol{r'}|$. Instead, the Laplacian values have to be evaluated only K times since they depend on $\boldsymbol{r}$ and not on $\boldsymbol{r'}$. It is clear that for the calculation of all the source contributions, the code needs to evaluate $K \cdot Q$ times the point to point distance. This procedure has to be carried out for each integrated atom. If the studied system is composed by N atoms, the total number of operations performed is approximatively $N \cdot K \cdot Q$ for the evaluation of $1/|\boldsymbol{r} - \boldsymbol{r'}|$ and $N \cdot K$ for the Laplacian values. That said, it is clear that the running time and the memory request dramatically explode with the increase in the number of atoms and the accuracy (resolution of the integration grids). This consideration is even more tangible if one considers that for a mid-high accuracy calculation, hundreds of thousands or even millions of points are required for each integration.

The second drawback is the presence of numerical errors associated with the grid integrations. It is well known that the source function descriptor is able to reconstruct the electron density in a reference point if this is greater or equal to $10^{-4}/10^{-5}$ *a.u.*[63,71] Below these thresholds, the electron density reconstruction becomes problematic, particularly for those points far from the nuclei. A small improvement can be obtained increasing the number of points (K) used for the calculation of the source function contributions, but this approach leads to an increased computational time. As a consequence of this,

one has to be careful to not include too large numerical errors in the integration of the source function contribution within the basins. On the other hand, to fulfil the property of integral over $\boldsymbol{r'}$ in Equation 1.36 it is important to extend the integration as far as one can in the three-dimensional space.

The third issue concerns the instability of the single $M(\Omega_X, \Omega_Y)$ matrix terms. The main responsible for this large oscillation is the choice of the electron density cutoff for the integral of Equation 1.36. As it is shown in the following sections, different ED cutoffs predict large variations in the single $M(\Omega_X, \Omega_Y)$ terms, while no issues are evident for the total reconstructed populations. Moreover, it is shown that also the application of some *a posteriori* corrections leads to a large oscillation of these terms. This instability is particularly delicate since it can undermine the reliability of the obtained numerical results.

4.3 Working on Weaknesses

In this section, I discuss about some solutions we have proposed to overcome the drawbacks shown before in the case of gas-phase molecules. It is important to underline that this work is still in progress and further analyses have to be carried out to assure the real potentiality of this method and the degree of improvements (both related to computational time and results accuracy) that our solutions introduce. Some works will be hopefully published in the next months with more detailed discussion and examples. However, I believe that this sections will be useful to the Readers to understand the way we have chosen to follow and the breakthroughs we have obtained up to now.

4.3.1 Definition of the Grids

The definition of the *full population matrix* requires, as already discussed, the evaluation of a six-dimensional integral, where the first variable, $\boldsymbol{r'}$, runs over Ω' and the second over the whole space. From a numerical point of view, this implies the definition of two grids of points: the first one, where the electron density Laplacian values are evaluated over the points $\boldsymbol{r'}$ and the second one where the ρ is reconstructed at the points $\boldsymbol{r}$ using the source function descriptor. For clarity, the former grid is called *primary* grid hereinafter, while the second one is defined *secondary* grid.

PROMEGA and PROAIMV codes belonging to the AIMPAC software package[92] use spherical polar *primary* grids centred on the atomic nuclei and they integrate the atomic properties through a Gauss-Legendre algorithm for the angular part and a Gaussian quadrature for the radial part.[168] This grid is the same used for the usual evaluation of integral properties in the basins Ω. The reason of this choice resides in the spherical-like distribution of the electron density and of its Laplacian in the isolated atom. In a molecule, the density of an atom undergoes a more or less significant departure from the spherical symmetry but, especially close to the nucleus where both the density and its

Laplacian vary very rapidly, this is a very good approximation, so that the choice of a spherical grid is much better than the use of a cubic grid. This *primary* grid is unique for each integrated atom since it is defined in term of relative coordinates with respect to the nuclei positions and because the angular radii of the grid are unique as well. From a computational point of view, the grid is generated at the beginning of the integration of the atom and at the end of the procedure the allocated memory can be released.

On the contrary, the definition of the *secondary* grid is less trivial. At variance with the *primary* one, this grid has to be defined in term of absolute coordinates and it has to be known since the beginning of the atomic integration because the evaluation of the source function contributions are performed during the atomic integrations. These two requirements are necessary to guarantee the uniform sampling of the source function contributions and the correct reconstruction (within the numerical errors) of the density in any point $\boldsymbol{r}$ in the space. From a computational viewpoint, this grid has to be held in memory until the end of the calculation.

Several points distributions can be chosen as *secondary* grid. The first idea is to use the patchwork of all the spherical polar grids used to integrate the Laplacian values. The advantage of this method is that the *primary* and *secondary* grids coincide. Other spherical-centred distributions like the Lebedev-Laikov[169] or the icosahedra grids[170] can be used. However, one drawback related to all these choices is the presence of small but not negligible regions of the space where two or more spherical grids overlap to each other. This fact introduces a bias in the final reconstruction of the electron density and in the population, since some regions are overweighted.

Another possibility is to partition the three-dimensional space using regular Voronoi polyhedra (VPs).[171] Voronoi solids allow a exhaustive partitioning of the space with no overlaps or empty spaces and the grid can be easily defined using an absolute reference system. Following this direction, the simplest Voronoi grid can be built using cubes as VP and stacking them in a regular pattern.

In our implementation of the *secondary* grid, we decided to start with this simple situation, aware of the possibility to extend it to any VPs though. In the modified version of the AIMPAC software, we introduced a subroutine for the generation of the *secondary* grid knowing the centre of the grid (C_{XYZ}), the resolution of the cubes (*i.e.* the length of their edges, *res*) and the number of points along the three directions X, Y and Z (N_X, N_Y and N_Z).

4.3.2 A Dynamical-Resolution Grid

The number of points and the resolution of the cube are the parameters that defines the spatial extension and the accuracy of the reconstructed density and consequently the number of points to evaluate. On one hand, it is important to have a high resolution (small step) and a large number of points to properly describe all the features of the electron density, but this implies a large computational cost and sometimes it is not feasible. On

the other hand, the use of a coarser grid reduces the number of points where the SF has to be reconstructed, but at the same time it introduces a major risk to badly describe regions where the electron density rapidly changes. A compromise is in principle the best possible option. The ultra-fine grid is very useful near the nuclei, where the $\rho(\boldsymbol{r})$ gradient is very high, while a coarse grid is sufficient far from the nuclei, where the electron density varies smoothly.

To tackle this problem, we have implemented an algorithm to build a grid with a variable resolution, calculated according to the variation of the electron density in the space. This algorithm works schematically in this way:

1. the starting grid with a fixed resolution is generated using the input parameters (centre, resolution and number of points);
2. a point $\boldsymbol{r}$ on the grid is selected as *parent* point;
3. the value of the *primary* electron density $\rho(\boldsymbol{r})$ is calculated using the wavefunction file produced by Gaussian package[90];
4. eight new *child* points $\boldsymbol{r'}$ are generated from the original one dividing the *parent* cube in eight *child* cubes.
5. the $\rho(\boldsymbol{r'})$ of the *child* points are evaluated and integrated using the Riemann sum;
6. the convergence criterion is evaluated. If it is satisfied, then the position, the grid resolution and the electron density value of the *parent* point are stored in the memory and the cycle restarts from step #2;
7. if the convergence criterion is not satisfied, the *child* points become the new *parent* points and the iteration restarts from step #4. Steps 4-7 are repeated until all the original child points satisfy the convergence criterion, then the algorithm starts again from the step #2.

The flow chart reported in Table 4.1 describes the same algorithm in a more schematic way.

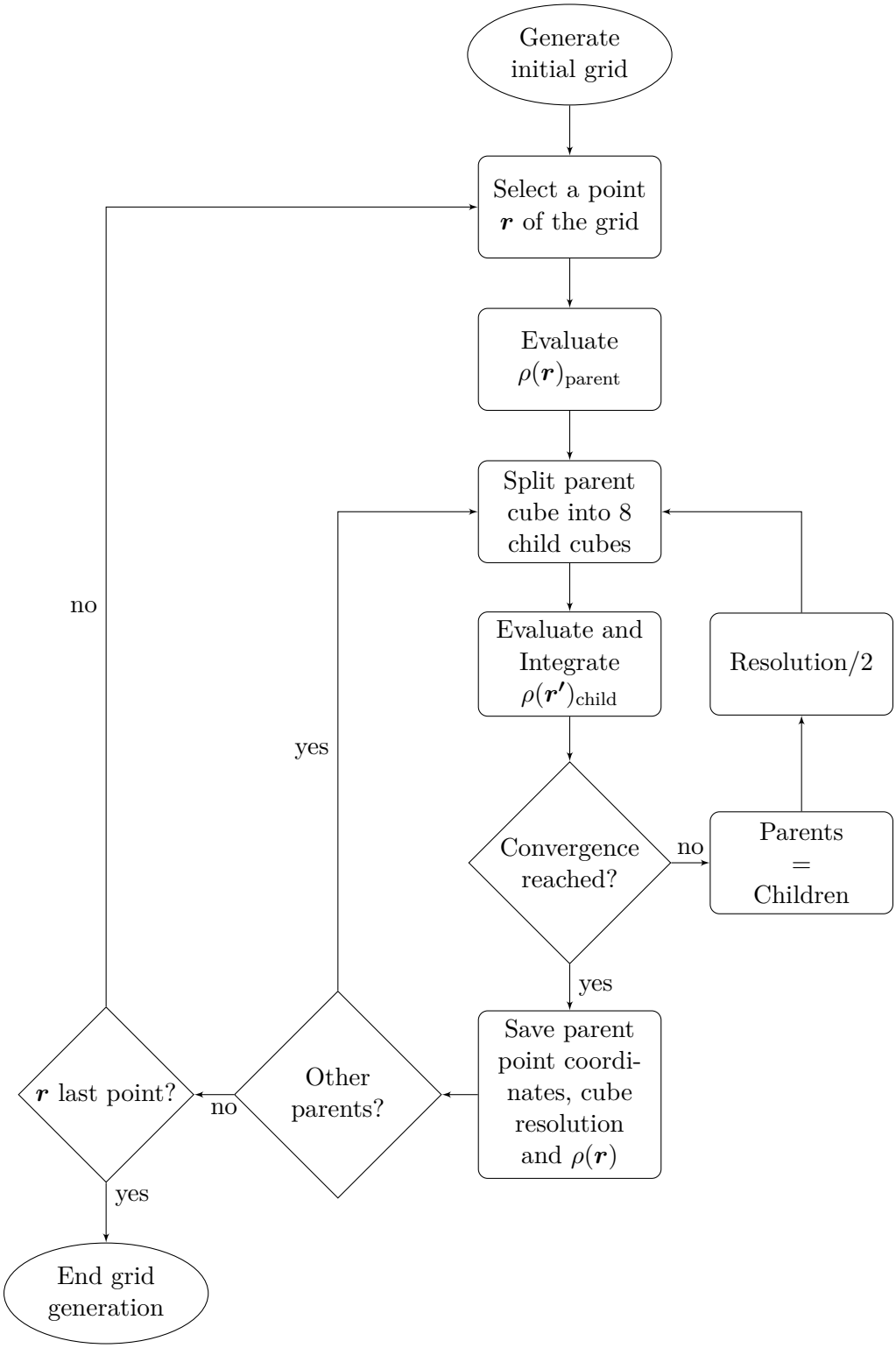
The convergence criterion used to assume a full convergence of the iterative algorithm is expressed in Equation 4.1:

$$\frac{\left|\rho(\boldsymbol{r})_p - \overline{\rho(\boldsymbol{r'})_c}\right|}{\rho(\boldsymbol{r})_p} \cdot 100 \leq \varepsilon\left[\rho(\boldsymbol{r})_p\right] \tag{4.1}$$

where $\rho(\boldsymbol{r})_p$ and $\overline{\rho(\boldsymbol{r'})_c}$ indicates the electron density of the *parent* and the average value of the eight *child* points, respectively, while $\varepsilon\left[\rho(\boldsymbol{r}_m)\right]$ is the percentage error used as convergence criterion.

If the percentage variation between the electron density values evaluated in one point and the one calculated for the eight sub-cubes created splitting the original cube differs less than the selected ε value, then the algorithm is considered to be at convergence and the *parent* point is saved in the memory. On the contrary, the procedure continue until all the points have reached the convergence.

Table 4.1: flow chart of the dynamical-resolution grid generation subroutine.



The ε is defined as a function of the ED value of the original point considered. It is then clear that one can decide to select different criteria according to the interval of the electron density considered. Strict criteria are suggested for large electron density regions, while looser values can be considered for regions where the ED is low.

Usually, if one starts with coarse grids (for instance 0.4 or 0.2 *a.u.* as initial resolution) and imposes strict criteria for the high density regions (for instance 0.5% for $\rho(\boldsymbol{r})$>1 *a.u.*), the algorithm maintains the original resolution for regions very far from the nuclei while it progressively increases the accuracy (decreases the value of resolution) up to values lower than 0.01 *a.u.* in the regions close to the nuclei. It is then clear that through this stratagem one is able to properly reconstruct critical areas of the density without using an exaggerate number of points.

Of course, criterion expressed in Equation 4.1 is not the only possible choice. Other parameters, such as the gradient of the electron density in the evaluated point, can be considered as a possible feature to predict resolution.

To accelerate the calculation it is possible to exploit the molecular symmetry of the molecule and evaluate the dynamical grid only in the symmetry-independent part of the space.

Once the dynamical-resolution grid has been evaluated, each point has to be uniquely assigned to a basin. To do so, we implemented the *octal tree algorithm* (OT) proposed by Malcolm and Popelier[172] because of its robustness and high-precision. Since this method requires the evaluation of the gradient paths (GPs) for each point and in an iterative way, the computational cost is quite high.[172] We have chosen the OT method to be confident with the obtained points assignation and populations. Now that we are aware of the reliability of our population matrices, faster algorithms, like the *near-grid*[173] and the *Yu-Trinkle*[174] methods, are going to be adapted and implemented in our code to speed-up the calculation and save computational time.

4.3.3 A Posteriori Treatment of Source Function Contributions

As already anticipated earlier, the source function descriptor is affected by numerical errors due to the finite accuracy of the integration process. This error is almost negligible when the reconstructed density in the point $\boldsymbol{r}$ is large, while it becomes relevant when the $\rho(\boldsymbol{r})$ $< 10^{-3}$ - 10^{-5} *a.u.*. For instance, in the evaluation of the population matrix of the water molecule it resulted that the average percentage errors for the intervals (expressed in *a.u.*) $\rho(\boldsymbol{r}) \leq 10^{-7}$, $10^{-7} < \rho(\boldsymbol{r}) \leq 10^{-5}$, $10^{-5} < \rho(\boldsymbol{r}) \leq 10^{-3}$ and $10^{-3} < \rho(\boldsymbol{r}) \leq 10^{0}$ are approximatively equal to 36000%, 70%, 1% and 0.2%. It follows that those points where the reconstruction is not satisfactory can introduce important biases in the final results if they are integrated without being corrected in some way.

Three different approaches have been explored to treat the data *a posteriori*. All of these methods rely on the evaluation of the reconstruction error and on its comparison with a threshold parameter, as expressed in Equation 4.2.

$$\frac{|S(\boldsymbol{r}) - \rho(\boldsymbol{r})|}{\rho(\boldsymbol{r})} \cdot 100 > ERR\,[\rho(\boldsymbol{r})] \tag{4.2}$$

where $S(\boldsymbol{r})$ is the reconstructed density obtained as the sum of all the atomic source function contributions, $\rho(\boldsymbol{r})$ is the value of the primary density obtained from the wavefunction and ERR is the ED-dependent cutoff value considered for the *a posteriori* correction. As for the definition of the dynamical grid, the user can decide the degree of accuracy as a function of the electron density value.

These proposed corrections only apply if the condition reported in Equation 4.2 is satisfied. The three approaches are the following:

1. *skip*: the point $\boldsymbol{r}$ is not taken into consideration in the integration;
2. *assign*: the total SF contribution is assigned to the basin Ω to which the point $\boldsymbol{r}$ belongs. This means that $S(\boldsymbol{r}, \Omega) = \rho(\boldsymbol{r})$ and $S(\boldsymbol{r}, \Omega') = 0$;
3. *scale*: the total reconstructed density S is scaled to the real value using a scaling factor $k = \rho(\boldsymbol{r})/S(\boldsymbol{r})$. All the single atomic contributions are linearly scaled by this factor k, *i.e.* $S'(\boldsymbol{r}, \Omega) = k \cdot S(\boldsymbol{r}, \Omega)$.

Each of them is valid for physical reasons but also shows some drawbacks.

The *skip* method is the only one that do not correct the SF values but exclude them from the integration. This decision is based on the fact that one does not know anything about the error of the SF descriptor and how this is distributed among the single contributions, so it is not possible to correct the contributions in the right way. Since the largest errors occur for small density values, it mainly affect regions far from the nuclei if the chosen ERR threshold is reasonably large. The removal of these points affects both the total number of integrated electrons and the distribution of the source function terms.

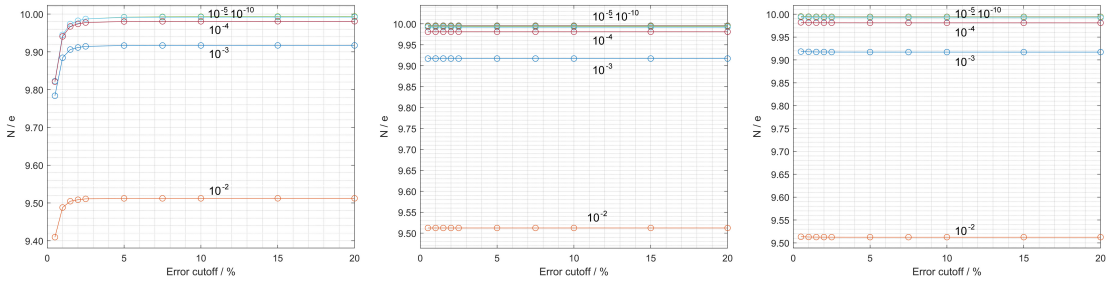
The *assign* correction relies on the assumption that when a point is far from the nuclei of the molecule, its density is mainly reconstructed by the basin it belongs to. Looking at Equation 1.27, this is reasonably true if one figure out that the $|\boldsymbol{r} - \boldsymbol{r'}|$ term is, on average, smaller for the points belonging to the same basin than for the others. This lead to a Green's function less effective in dropping the SF values. On this assumption, the contributions of the other basins are considered negligible. In this case, the assignation of the whole SF contribution to one basin only change the entity of the atomic contributions but it does not affect the total number of electrons integrated nor the reconstruction errors for the corrected points.

On the contrary, the *scale* correction is based on the idea that the reconstruction error in one point is the same for all the atomic source function contributions, being it based on the same intrinsic limitation of the numerical integration. The single terms are then scaled according to the same scale factor k so that their sum is equal to the total density in the corrected point. This method implies a small variation of the total number of electrons with respect to the value obtained from the non-corrected integration. On the other hand, the number of electrons tends to the real value, *i.e.* the number of electrons obtained

integrating the *primary* ρ, when the ERR value approaches to zero. The single atomic contributions are affected too, as in the case of the previous two approaches.

To test the efficiency and robustness of these three proposed corrections, we evaluated the full population matrices of a water molecule in its optimized geometry at the HF/6-311++G(2d,2p) level of theory in gas phase applying all the proposed corrections. We verified the effect of changing the ERR cutoff value on the total population N and on the population matrix elements $M(\Omega_X, \Omega_Y)$ using different ED cutoff levels. We tested these results on a very accurate grid, where the initial step was set to be 0.4 *a.u.* and the ε parameter was equal to 0.2%. The sizes of the grid were chosen to be sure that the $\rho = 10^{-10}$ isosurface was included within the boundaries of the box.

Figure 4.1 depicts the effect of the three corrections on the total number of electrons when the ERR value is changed.



(a) *skip* (b) *assign* (c) *scale*

Figure 4.1: total reconstructed population (N) trend with respect to the reconstruction error cutoffs for the three corrections proposed (*skip*, *assign* and *scale*) at different electron density cutoffs.

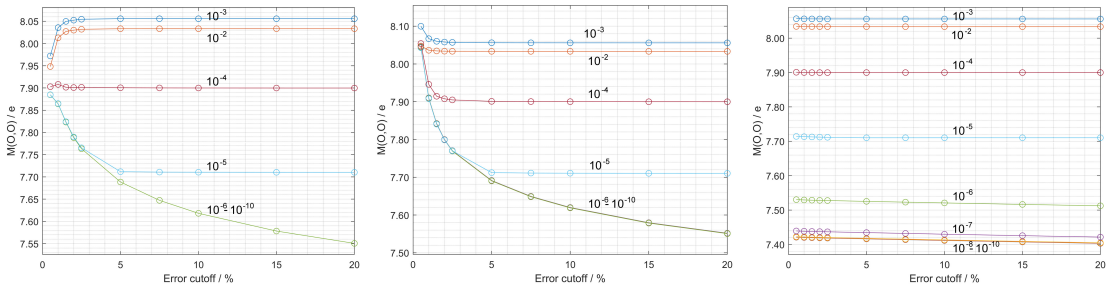
As expected, the total number of electrons is extremely sensitive to the ρ cutoff: the lower is the ρ value, the closer to 10 is the total population. The integrated values are equal to the real value within 0.02 e for ED cutoff equal or lower than 10^{-4} *a.u.* when *assign* or *scale* corrections are applied, confirming that the majority of the total population is within this limit. Usually, the quality of the integration can be measured looking at how far is the final value from the real value. Divergences up to 1% can be considered in this case acceptable for a good quality integration.

The behaviour of the curves is maintained similar or equal independently on the ERR cutoff. In the case of the *assign* correction, no effect on the total population is present since, as it was explained before, there are no variations on this parameter. The *scale* correction has a similar behaviour, although very small variations (not appreciable from the graphs) are present. For both methods, the same amount of electrons is retrieved if the same ED cutoff is considered. Finally, the *skip* correction shows a plateaux behaviour, similar to that found for the previous corrections, for large values of ERR cutoff, while the value drops by 0.1-0.2 e when $ERR < 5\%$. These different trends can be explained considering two aspects related to the error distribution. First, as already said before, the points affected by the larger errors are the ones where the electron density is low. These

points are already removed from the integration when the ED cutoff is applied, so the number of corrected points decreases. For instance, the points affected by an error greater than 5% are the 84.3% of the totality when the ED cutoff is equal to 10^{-10} *a.u.*, while the same value drops to 54.2% and 1.1% when $\rho = 10^{-7}$ *a.u.* and 10^{-5} *a.u.*, respectively. The second consideration is related to the relative weight of the removed electrons. Moving from 10^{-10} to 10^{-5}, the number of electrons lost is equal to 0.0023. It is then evident that this fraction of the density is not relevant for the total reconstruction of the electron density and then the effect is low. On the contrary, when the ERR is lower than 5%, also the high density points start to be removed from the integration. These points are not excluded by the application of the ED cutoff and they are responsible for a large number of electrons. For instance, the number of points affected by errors greater than 1.5% for the three density cutoff evaluated before are 89.7% ($\rho(\boldsymbol{r})$=10^{-10} *a.u.*), 70.0% ($\rho(\boldsymbol{r})$=10^{-7} *a.u.*) and 30.8% ($\rho(\boldsymbol{r})$=10^{-5} *a.u.*). Excluding these points from the integration process leads to a large loss of electrons and a consequent drop, independently from the ED cutoff, in the curve when $ERR < 5\%$.

Except for this last case related to the *skip* approach, the total population results almost unaffected by the variation of the error cutoff.

A drastically different situation occurs when the single matrix contributions are considered. Figure 4.2 shows the *inner self-contribution* $M(O,O)$ for the oxygen atom.



(a) *skip* (b) *assign* (c) *scale*

Figure 4.2: inner self-contribution M(O,O) trend with respect to the reconstruction error cutoffs for the three corrections proposed (*skip*, *assign* and *scale*) at different electron density cutoffs.

From a first look, it is clear that this term is extremely sensitive to the selected correction method. All the three corrections retrieve an initial increase of the contribution moving from ρ cutoff of 10^{-2} *a.u.* to 10^{-3} *a.u.*, followed by monotonic shift of the curves toward lower values. The total variation between the highest and the lowest density cutoff ranges from 0.07 e when the considered error is 0.5% to 0.5 e for an error equal to 20%.

The *skip* and *assign* approaches show a similar general behaviour for $ERR > 5\%$ values while they disagree in the $ERR < 5\%$ region for the high-density cutoffs ($\rho \geq 10^{-4}$ *a.u.*). It is interesting noting that both methods do not retrieve any convergence on the M(O,O) contribution with ED cutoff lower than 10^{-6} *a.u.* up to 20% errors.

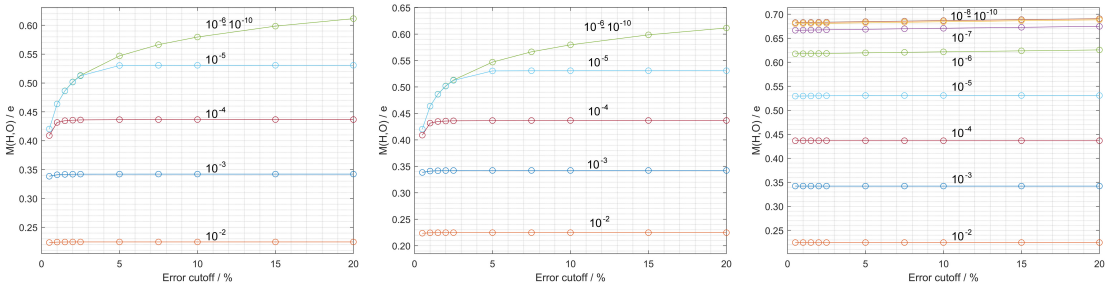
The *scale* correction shows the same average $M(O,O)$ *vs* ρ trend found for the other

two approaches. On the contrary, the behaviour with respect to the variation of the ERR cutoff is definitely more stable and the variations are limited within ± 0.02 e for each density cutoff. The drop in the number of electrons related to the different ED cutoffs is ≈ 0.6 e between the highest and the lowest cutoff. These results suggest that the *scale* correction is the most stable with respect to the change on the error cutoff.

Apart these differences, all the three corrections retrieve more or less the same $M(O,O)$ contribution when convergence is reached, that is when the contribution remains stable when the ERR parameter changes.

The explanation of the peculiar behaviour of the *scale* correction is related to its definition. While *skip* and *assign* are strong "step corrections", *i.e.* they both apply a correction whose nature is completely independent on the degree of error, the *scale* approach is a smoother function and the applied correction is proportional to the entity of the error. When the reconstruction error is low (usually for high electron density values), the *scale* effect is weak and almost negligible, while when the error is large (low electron densitiy values) the local effect is strong but not so relevant because of the small total integrated population associated to the low ED regions. For this reason the effect produced by *scale* correction does not change drastically between low and high ERR values.

Similar differences between *skip/assign* and *scale* approaches are found for the inter-atomic $M(H,O)$ contribution, which is equal to $M(H',O)$ for symmetry reasons (Figure 4.3).



(a) *skip* (b) *assign* (c) *scale*

Figure 4.3: outer interatomic contribution M(H,O) trend with respect to the reconstruction error cutoffs for the three corrections proposed (*skip*, *assign* and *scale*) at different electron density cutoffs.

The former two methods are stable against the variation of the error cutoff when $ERR > 5\%$, while they retrieve the same behaviour observed in the $M(O,O)$ plots (Figure 4.2) for smaller errors and low density cutoffs. On the contrary, the *scale* correction is robust and almost stable along all the considered intervals. The single contributions decrease in absolute value when the cutoff on ρ becomes more tight. The total variations of the contributions associated with different ED cutoffs are in the range of 0.45 e for all the methods employed.

All the other intra- and inter-atomic contributions not shown here retrieve the same features already explained here. It is clear that the single atomic contributions are very

sensitive with respect to the reconstruction error correction that one applies. On the other hand, this process is necessary to try to limit the introduction of biases in the final population matrix. All three proposed approaches are based on reasonable considerations and then are equally valid. However, the shown behaviours lead us to believe the *scale* correction as the most robust and reliable method for all the error cutoffs considered.

A different reasoning must be applied to the electron density cutoffs, since it has been shown that its choice extremely affects the single values independently on the correction applied. In the next section, some aspects concerning the choice of the ED cutoff is discussed.

4.3.4 Electron Density: Where to Cut?

The strong dependence of the reconstruction errors with respect to the electron density values has already been largely discussed in the previous sections, together with some methodologies we proposed to reduce and correct these errors.

Another important detail to discuss is the choice of the electron density cutoff. On the one hand, Equation 1.36 is mathematically valid when the integrations are extended to infinity and then it is important to use the lowest electron density cutoff possible. On the other hand, it has been discussed that the source function descriptor suffers large numerical errors when the reconstructed density is lower than 10^{-3}-10^{-5} *a.u.* and *a posteriori* corrections are required. High ED cutoffs are in this case better in order to be sure that negligible biases are introduced in the final population matrix. Moreover, from a computational point of view high ρ cutoffs are associated to a smaller amount of points and then to a faster calculation.

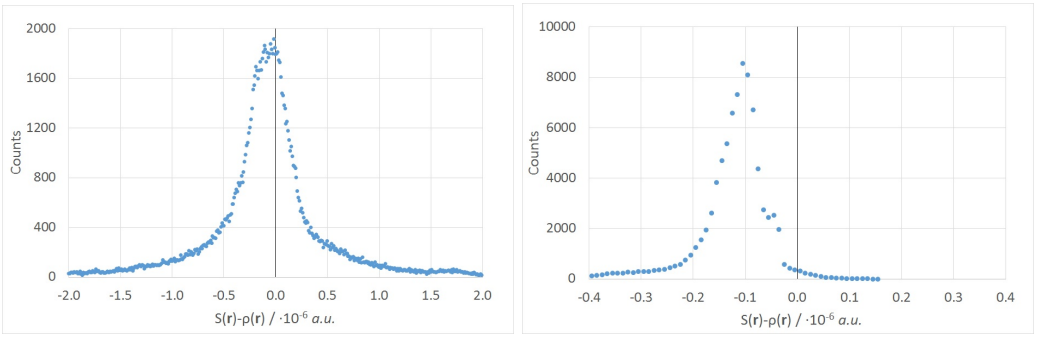
The questions that may arise are the following: what criteria should one use to decide the cutoff to apply on the electron density value? Can one be confident with the cutoff he/she chooses?

It has to be said clearly that no unambiguous criterion can be used to make this decision.

A possible answer to the questions above may be related to the issue of the numerical reconstruction error of the source function descriptor. The *a posteriori* corrections introduced in the previous section allow one to slightly extend the range of applicability of the source function. However, these corrections (in particular the *scale* method) can be considered reliable if the set of data on which they are applied follows a statistical distribution of errors. If the last requirement is not satisfied, then biased values are introduced in the population matrix even if corrections are applied.

The absolute reconstruction error distributions $S(\boldsymbol{r}) - \rho(\boldsymbol{r})$ evaluated in the water molecule are shown in Figure 4.4

Figure 4.4 shows two error distributions: the first one is associated with those points that have $\rho(\boldsymbol{r})>10^{-5}$ *a.u.* (Figure 4.4(a)), while the second one for those points with ED lower than 10^{-5} *a.u.* (Figure 4.4(b)). It is evident that the former set of data presents an



(a) $\rho(\boldsymbol{r}) > 10^{-5}$ (b) $\rho(\boldsymbol{r}) < 10^{-5}$

Figure 4.4: distribution of the reconstruction error in the water population matrix for points with electrons density $\rho(\boldsymbol{r})>10^{-5}$ *a.u.* (a) and $\rho(\boldsymbol{r})<10^{-5}$ *a.u.* (b).

almost symmetric distribution of errors centred on the zero, while the latter set is shifted toward negative values. This is a proof of the fact that the source function introduces non-statistical errors below a certain value of the electron density. The effect of the *scale* correction on the error distribution is to reduce its width and not to shift it with respect to the zero. For this reason, the *scale* correction and all the other *a posteriori* approaches do not reduce the intrinsic bias of the considered points.

To avoid any non-statistical bias, we believe that the error distribution may be considered as a possible discriminating factor for the choice of the electron density cutoff. In this example, $\rho(\boldsymbol{r})\geq 10^{-5}$ *a.u.* is the minimum cutoff that can be used to limit the introduction of biases in the population matrix.

4.4 Numerical Examples

In this section a couple of calculated population matrices are presented to show the potentiality of this method. A full and detailed analysis of the ability to extract possible chemical information from these matrices has not been carried out yet and will be presented in future works. Although this part is for sure the most interesting one, we have first preferred to carefully understand the behaviour of this novel population analysis in order to be aware of its reliability.

The two examples here prosed are the following: in the first section, the population matrix of the water molecule is shown and the effects of symmetric and asymmetric O-H bonds stretching have been analysed; in the second section, preliminary results on the full matrices of the triad ethane, ethene and ethyne are presented and shortly discussed.

4.4.1 Water Molecule and O-H Symmetric and Asymmetric Stretching

The full population matrix of the water molecule in its singlet state was evaluated at the HF/6-311G++(2d,2p) level of theory and in its fully relaxed geometry. The *primary* grid was built using a set of 144x96 angular points and 200 radial points for each atom. The

initial resolution of the cubic *secondary* grid was set equal to 0.4 bohr and a dynamical-resolution parameter ε=0.5% was used for all the electron density range. *Scale a posteriori* correction was applied using a reconstruction error cutoff (ERR) dependent on the electron density values: 10^{-10} *a.u.* $\leq\rho(\boldsymbol{r})< 10^{-7}$ *a.u.*: 5.0 %; 10^{-7} *a.u.* $\leq \rho(\boldsymbol{r})< 10^{-5}$ *a.u.*: 2.5 %; 10^{-5} *a.u.* $\leq\rho(\boldsymbol{r})< 10^{-3}$ *a.u.*: 1.0 %; 10^{-3} *a.u.* $\leq\rho(\boldsymbol{r}) < 10^{0}$ *a.u.*: 0.2 %; $\rho(\boldsymbol{r})>10^{0}$ *a.u.*: 0.05 %.

The final integration on the *secondary* grid was cut at $\rho(\boldsymbol{r})=10^{-5}$ *a.u.* using as criterion the statistical error distribution.

Numerical results of this integration are reported in Table 4.2.

Table 4.2: full population matrix elements $M(\Omega,\Omega')$ of the water molecule evaluated using as cutoff values: ERR=0.5% (*scale* correction) and $\rho(\boldsymbol{r})\geq10^{-5}$ *a.u.*. The geometry was optimized at HF/6-311++G(2d,2p) level of theory. Contributions were scaled to normalise the total population to the real number of electrons.

Ω	$M(\Omega,H)$	$M(\Omega,O)$	$M(\Omega,H')$	$N(\Omega)$
$M(H,\Omega')$	-0.255	0.528	0.096	0.369
$M(O,\Omega')$	0.769	7.723	0.769	9.262
$M(H',\Omega')$	0.096	0.528	-0.255	0.369
$M(\Omega')$	0.610	8.779	0.610	10.00

The numerical values reported in Table 4.2 have been obtained evaluating the source function contributions in 1 085 038 points of the secondary grid. Another more accurate grid has been used to test the accuracy of the chosen one using more strict ε values, resulting in 3 045 232 total points. Although this more accurate grid, the larger variation obtained interests the $M(H,H)$ population and it is as high as 0.003 e.

The final matrix contributions were normalised to obtained a total number of reconstructed electrons equal to the real one (10 for the water molecule). The actual number of integrated electrons was as high as 10.011 e, slightly larger than the expected result. The fact that this value is higher than the real one has no physical meaning but it is only due to the finite accuracy of the numerical integration. However, since the error of the reconstructed value is $\approx$ 0.1%, the accuracy of the integration is good enough to trust the numerical values obtained. The effect on single matrix elements was quite low, with the largest difference on the $M(O,O)$ contribution that moved from 7.732 e to 7.723 e, with a net variation of -0.009 e. We decided to use this aesthetic procedure to have a direct correspondence between the reported values and the real number of electrons.

The precision of the SF integration can be also analysed comparing the reconstructed population for each basin and the values obtained through the *primary* integration. The population of the H atoms are equal to 0.369 e for the source function reconstructed density and 0.373 e for the spherical polar grid integration. The difference between the two values is very low both in absolute value (0.004 e) and in percentage ($\approx$ 1%). For the oxygen atom the situation is similar, with the corresponding values being equal to 9.262 e and 9.254 e, respectively. In this case, the absolute error is slightly larger (0.008 e) than in the previous case, but the percentage error is definitely lower (<0.1%). For both atoms, the accuracy of the reconstructed total population is good enough.

Analysing the single population terms, it results immediately evident one of the properties belonging to the source function population matrix, *i.e.* the asymmetry of the interatomic contributions. The total interatomic caused population (ICP) between the oxygen and the hydrogen atoms, evaluated as $ICP(O \leftrightarrow H) = M(O,H) + M(H,O)$ is equal to 1.297 e, but the two atoms do not act in the same way. The electron population caused by the oxygen atom into the H basin is equal to 0.528 e ($\approx 40\%$), while that determined by the hydrogen atom into the O basin is equal to 0.769 e ($\approx 60\%$). This asymmetry is related to the different nature of the two atoms involved in the interaction and can give important information. On the contrary, the caused population by the two H atoms is only 0.192 e, where the two terms $M(H,H')$ and $M(H',H)$ are equal for symmetry.

The H atom is the major contributor to the caused population with the Oxygen despite it formally carries a lower number of electrons. This fact is far more interesting if one looks at the self-contribution $M(H,H)$ where a negative term (-0.255 e) appears. These two numbers suggest that the H atom tends to subtract electron density from itself to distribute it to the Oxygen and in small part to the other Hydrogen. That said, it is now clear that the major contributor to the reconstruction of the 0.369 e within the H basins is the Oxygen atom, which contributes with 0.528 e and it is then able to overcome the negative H self-contribution. The population on the O atom can be divided into a 7.723 e (83.4%) coming from the internal self-contribution and an external 0.769 e (8.3%) terms coming from each H atom.

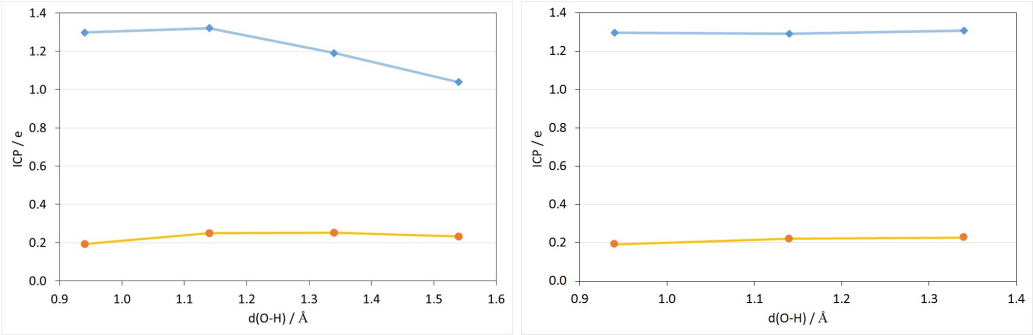
The values reported in the population matrix clearly indicates that there is a large exchange of population between the atoms. The total interatomic caused population amounts to 2.787 e, which is $\approx 27.9\%$ of the total number of electrons of the system.

A last thing to be noted in the matrix is the remarkable difference between the last column, which represents the populations of the basins and the last row, which on the contrary describes the electron population caused by each atom in the remaining one, included itself (*i.e.* the "source ability"). These two different terms can be related to the ability of the source function descriptor to highlight the reconstruction of the electron density in any point within a *cause-effect* relationship. In this case, the *causes* are the "source abilities" of single atoms $M(\Omega')$, which depend from the electron density distribution within the basin, while the *effect* is the atomic populations $N(\Omega)$, which includes also the Ω atom source ability on its own basin.

A full comparison of these results with those obtained from other population analyses will be discussed in future works.

The effects of the O-H symmetric and asymmetric stretching on the population matrix were analysed using the same parameters of the previous integration. The length of the O-H bond was fixed to 1.14 and 1.34 Å while the angle H-O-H was let free to refine during the geometry partial optimization. An additional O-H bond length of 1.54 Å was added for the symmetric stretching. The optimized H-O-H angle undergoes small variation and it ranges from 100.03° to 91.25° for the symmetric streching and from 102.54° to 99.20° for the asymmetric one.

Figure 4.5 shows the effect of the bond symmetric (Figure 4.5(a)) and asymmetric (Figure 4.5(b)) stretching on the ICP(O↔H) and ICP(H↔H) terms.



(a) asymmetric stretching (b) symmetric stretching

Figure 4.5: interatomic caused populations (ICPs) for the symmetric (a) and asymmetric (b) stretching of O-H bonds in the water molecule. ICPs(O↔H) are shown as blu diamonds, while ICPs(H↔H) as orange circles. Solid lines serve as a guide for the eyes.

In the symmetric stretching, the ICP between the O and H atoms decreases, except for the first point, when the bond is stretched. This fact is reasonable from a chemical point of view, since distant atoms "talk" to each other in a less efficient way than close ones. Other population analyses, like the Mulliken's, are able to recover the same trend but the interpretation of this phenomenon is totally different. The Mulliken's analysis describes the ICP in terms of overlap population, while the SF descriptor sees it as reciprocal exchange of electrons within a *cause-effect* relation. The ICP(H↔H) does not undergo significant variation in the tested range, even if the H-H distance slightly shorts.

The same trend seen for the caused populations can be highlighted looking at the total inner and outer populations. Inner populations tend to increase with the elongation of the O-H bonds, ranging from ≈ 7.2 e in the relaxed water molecule to ≈ 7.7 e in the one with O-H distance equal to 1.54 Å. On the contrary, the opposite trend occurs for the total outer contribution, where it decreases by $\approx$ 0.4-0.5 e. This trend is chemical reasonable since the bonds become more ionic and the atoms tend to be more isolated from each other.

Table 4.3 shows the population matrices for the three symmetric stretched molecules.

The same trend is confirmed by the integrated populations of the atoms, which approach to the elemental number of electrons (8 e for the Oxygen and 1 e for the Hydrogens) with the increasing of the O-H bond length.

The asymmetric stretching (Figure 4.5(b)) reveals an intermediate situation between the fully relaxed water molecule and the symmetric-stretched one. In particular, the population matrix becomes asymmetric because of the molecular symmetry breaking in the molecule, as it can be seen in Table 4.4. The matrix elements associated to H' and O-H' are quite similar to those found in the matrix 4.2 since this bond is not affected by the stretching. On the contrary, the contributions related to H and O-H follow similar trends observed in the symmetric stretching (Table 4.3).

Table 4.3: full population matrices of the water molecule with O-H symmetrically stretched to 1.14, 1.34 and 1.54 Å. Cutoff values: ε=0.5%, ERR=0.5% (*scale* correction), $\rho(\boldsymbol{r})\geq 10^{-5}$. The H-O-H' angle was optimized at HF/6-311++G(2d,2p) level of theory. Contributions were scaled to normalise the total population to the real number of electrons.

	Ω	$M(\Omega,H)$	$M(\Omega,O)$	$M(\Omega,H')$	$N(\Omega)$
O-H = 1.14 Å	$M(H,\Omega')$	-0.194	0.559	0.125	0.490
	$M(O,\Omega')$	0.762	7.495	0.762	9.020
	$M(H',\Omega')$	0.125	0.559	-0.194	0.490
	$M(\Omega')$	0.693	8.613	0.693	10.00
O-H = 1.34 Å	$M(H,\Omega')$	-0.071	0.517	0.128	0.572
	$M(O,\Omega')$	0.674	7.507	0.674	8.856
	$M(H',\Omega')$	0.126	0.517	-0.071	0.572
	$M(\Omega')$	0.730	8.541	0.729	10.00
O-H = 1.54 Å	$M(H,\Omega')$	0.061	0.460	0.117	0.638
	$M(O,\Omega')$	0.579	7.567	0.579	8.725
	$M(H',\Omega')$	0.117	0.460	0.061	0.638
	$M(\Omega')$	0.756	8.487	0.756	10.00

Table 4.4: full population matrices of the water molecule with O-H asymmetrically stretched to 1.14 and 1.34 Å. Cutoff values: ε=0.5%, ERR=0.5% (*scale* correction), $\rho(\boldsymbol{r})\geq 10^{-5}$. The H-O-H' angle was optimized at HF/6-311++G(2d,2p) level of theory. Contributions were scaled to normalise the total population to the real number of electrons.

	Ω	$M(\Omega,H)$	$M(\Omega,O)$	$M(\Omega,H')$	$N(\Omega)$
O-H = 1.14 Å	$M(H,\Omega')$	-0.192	0.544	0.132	0.484
	$M(O,\Omega')$	0.775	7.613	0.756	9.143
	$M(H',\Omega')$	0.089	0.538	-0.254	0.373
	$M(\Omega')$	0.672	8.694	0.634	10.00
O-H = 1.34 Å	$M(H,\Omega')$	-0.074	0.488	0.150	0.564
	$M(O,\Omega')$	0.696	7.614	0.753	9.063
	$M(H',\Omega')$	0.077	0.555	-0.258	0.374
	$M(\Omega')$	0.699	8.657	0.644	10.00

4.4.2 The Ethane/Ethene/Ethyne Series

The full population matrices of the serie ethane, ethene and ethyne were evaluated at the B3LYP/6-311G++(2d,2p) level of theory. The initial resolution of the grid was set equal to 0.4 bohr, while the grid resolution cutoff ε was taken in the range 0.2%-2.5% according to the electron density values and similarly to that adopted for the water molecule integration above. The points were integrated only if $\rho(\boldsymbol{r})>10^{-5}$ *a.u.* and the *scale* correction with cutoff equal to 0.5% was applied. The same *primary* grid as for the water molecule described above (144x96 angular and 200 radial points) was taken.

Tables 4.5, 4.6 and 4.7 report the full population matrices for the ethane, ethene and ethyne molecules. Numbers highlighted in bold refer to the inner and outer contributions of the two C atoms.

Table 4.5: full population matrix of the singlet ground state of the ethane molecule (C_2H_6) using as cutoff values: ERR=0.5% (*scale* correction), $\rho(\boldsymbol{r})\geq10^{-5}$ *a.u.*. The geometry was optimized at B3LYP/6-311++G(2d,2p) level of theory. Contributions were scaled to normalise the total population to the real number of electrons. Real integrated populations are reported in parenthesis in the last column.

Ω	H1	H2	H3	C1	C1'	H3'	H2'	H1'	N(Ω)
H1	0.126	0.189	0.188	0.233	0.029	0.060	0.060	0.126	1.012 (1.013)
H2	0.189	0.124	0.189	0.232	0.029	0.060	0.127	0.060	1.010 (1.013)
H3	0.189	0.189	0.125	0.233	0.029	0.126	0.060	0.060	1.010 (1.013)
C1	0.547	0.548	0.548	**3.532**	**0.254**	0.180	0.180	0.180	5.968 (5.961)
C1'	0.180	0.180	0.180	**0.254**	**3.532**	0.548	0.548	0.547	5.968 (5.961)
H3'	0.060	0.060	0.126	0.029	0.233	0.125	0.189	0.189	1.010 (1.013)
H2'	0.060	0.127	0.060	0.029	0.232	0.189	0.124	0.189	1.010 (1.013)
H1'	0.126	0.060	0.060	0.029	0.233	0.188	0.189	0.126	1.012 (1.013)
$M(\Omega')$	1.477	1.475	1.476	4.572	4.572	1.476	1.475	1.477	18.000

Table 4.6: full population matrix of the singlet ground state of the ethene molecule (C_2H_4) using as cutoff values: ERR=0.5% (*scale* correction), $\rho(\boldsymbol{r})\geq10^{-5}$ *a.u.*. The geometry was optimized at B3LYP/6-311++G(2d,2p) level of theory. Contributions were scaled to normalise the total population to the real number of electrons. Real integrated populations are reported in parenthesis in the last column.

Ω	H1	H2	C1	C1'	H2'	H1'	N(Ω)
H1	-0.008	0.243	0.258	0.257	0.063	0.169	0.983 (0.983)
H2	0.244	-0.009	0.258	0.258	0.169	0.063	0.983 (0.983)
C1	0.680	0.680	**3.255**	**0.765**	0.327	0.327	6.035 (6.035)
C1'	0.327	0.327	**0.766**	**3.255**	0.680	0.680	6.035 (6.035)
H2'	0.063	0.169	0.258	0.258	-0.009	0.244	0.983 (0.983)
H1'	0.169	0.063	0.257	0.258	0.243	-0.007	0.983 (0.983)
$M(\Omega')$	1.475	1.473	5.052	5.052	1.473	1.476	16.000

The accuracy of the integration may be judged by comparing the reconstructed population of the basins as a sum of the source function contributions and the real integrated values reported in parentheses. For all the molecules, the maximum absolute difference

Table 4.7: full population matrix of the singlet ground state of the ethyne molecule (C_2H_2) using as cutoff values: ERR=0.5% (*scale* correction), $\rho(\boldsymbol{r})\geq 10^{-5}$ *a.u.*. The geometry was optimized at B3LYP/6-311++G(2d,2p) level of theory. Contributions were scaled to normalise the total population to the real number of electrons. Real integrated populations are reported in parenthesis in the last column.

Ω	H1	C1	C1'	H1'	N(Ω)
H1	-0.141	0.258	0.576	0.157	0.850 (0.856)
C1	0.819	**2.987**	**1.783**	0.562	6.150 (6.144)
C1'	0.562	**1.783**	**2.987**	0.819	6.150 (6.144)
H1'	0.157	0.576	0.258	-0.141	0.850 (0.856)
$M(\Omega')$	1.397	5.603	5.603	1.397	14.000

between the two populations is equal to 0.007 e for the C atoms in the ethane molecule. Moreover, the total number of electrons not integrated are equal to 0.004 (ethane), 0.015 (ethene) and 0.016 (ethyne), values that in percentage are equal or lower to 0.11% with respect to the total molecular populations.

Looking at the populations associated to the C atoms (highlighted in bold in the matrices), it is evident that with the increase of the formal bond order from single (ethane) to triple (ethyne), the self-contributions of the carbon atoms decrease from 3.532 e to 2.987 e. On the contrary, the interatomic caused populations between C increase from 0.508 e to 3.566 e. Both trends suggest that the electrons sharing between the two atoms increases and the population matrices are able to recover the double and triple nature of the C-C bond in ethene and ethyne, although the variation of the ICP(C↔C) is not linearly proportional as the formal bond orders. As already said before, further analyses have to be done to try to associate the ICPs and the bond orders in a numerical way.

Table 4.8 reports all the possible ICPs between couples of atoms and the total inner (N_i) and outer (N_o) contributions.

Table 4.8: inner (N_i) and outer (N_o) contributions, together with all the possible interatomic caused population (ICPs) of the series ethane (C_2H_6), ethene (C_2H_4) and ethyne (C_2H_2). All the reported IEPs refer to single X-Y interaction.

			ICP					
System	N_i	N_o	C↔C	C↔H	C↔H'	H↔H'$_{trans}$	H↔H'$_{cis}$	H↔H
C_2H_6	7.814	10.186	0.508	0.780	0.209	0.252	0.120	0.377
C_2H_4	6.477	9.523	1.531	0.938	0.585	0.338	0.125	0.487
C_2H_2	5.693	8.307	3.566	1.076	1.138	0.314	//	//

Other than the increase of the C↔C ICPs, Table 4.8 also shows that the total amount of electrons exchanged by the C and the bonded H atoms increase by a ≈ 0.27 e for each hydrogen atom removed.

A larger increase is observed in the caused population between the Carbons and the non-bonded H' atoms, where the amount of the electrons shared is $\approx$ 2.8 and 5.4 times larger in ethene (0.585 e) and ethyne (1.138 e) with respect to ethane (0.209 e). This is

another evidence of the increased electron delocalization and shows that not only the C atoms are involved in the delocalization but also the hydrogen atoms.

Another observed feature from Table 4.8 is that *trans* H atoms bonded to different carbon atoms exchange more electrons than those in a *cis* or *staggered* conformation. For instance, in the ethene molecule the ICP between H1 and H1' (*trans*) is equal to 0.338 e, while the same value between H1 and H2' is only 0.125 e. Similar but less marked differences are found for the ethane molecule. This effect may be related in some way to the electron delocalization of the systems, but no conclusions can be drawn before further investigations.

4.5 Conclusions

The analysis of the population distribution among the atoms is an important descriptor to gain insights on localization/delocalization effects and on chemical bonding. Several methods have been proposed over the years to obtain population matrices based on different atomic partitioning. All these methods present as obvious some advantages and disadvantages. In particular, methods based on the electron density scalar field has the great advantage to be amenable of both theoretical and experimental determination.

The main advantage of the source function population analysis over all the other methods is that it allows one to directly retrieve the number of caused (or reconstructed) electrons between the atoms within a *cause-effect* relationship. Moreover, it allows to discriminate between the effects of two different atoms maintaining the asymmetry of the interaction. Although its large potentiality, it has been demonstrated that severe drawbacks, such as the large computational cost, the numerical instability and the reconstruction errors, are present.

In this chapter, the latest breakthroughs on the full population analysis performed through the integration of the source function descriptor have been shown. To face the problem of the computational cost, we implemented a simple algorithm that determines the resolution of the *secondary* grid, *i.e.* the set of points where the source function contributions are evaluated, in a dynamical way based on the local variation of the electron density distributions. Through this approach it is possible to treat each region of the molecule with the proper accuracy without using an excess of points where not needed.

To face the source function error in the reconstructed density we have introduced and tested three different corrections: *skip*, *assign* and *scale*. All these three approaches are applied to those points where the electron density reconstructed error is larger than a set threshold. The former two approaches are extremely sensitive to the selected error cutoff, with large non-physical variation for very strict thresholds. On the contrary, the *scale* correction is particular stable in the tested range of error cutoffs for all the electron density thresholds considered.

Concerning the numerical instability, we have found out that a cutoff on the electron density is needed in order to remove biases related to non statistical error distributions

that can not be corrected by the three methods proposed.

The potentiality of this novel population analysis has been shown in a couple of simple cases. The population matrix of the water molecule reveals how the hydrogen atoms preferentially determine electron population on O rather than on themselves, as they even act as a sink for their own basin. The symmetrical and asymmetrical elongations of the O-H bonds in this molecule highlight the ability of this method to follow the chemical sense. Interatomic caused populations between the atoms decrease, while the self-contributions of the atoms increase as expected from the decreased covalent character of the bonds.

The ethane-ethene-ethyne series shows that the formation of single, double and triple bonds are cleary visible from the values of the population matrices. In particular, the interatomic caused population between the two C atoms increases accordingly to formal bond order of the C-C bond showing an increase of the electron delocalization. Morever, the increment of the bond strength is associated to larger electrons sharing also between the C atoms and their non-bonded Hydrogens.

Although further tests and investigations on the performance of this analysis have to be carried out, we believe that some steps towards the possibility to apply this method for more interesting situations have been done.

Chapter 5

Intermolecular Recognition of the Antimalarial Drug Chloroquine[i]

Chloroquine (CQ) is an antimalarial drug belonging to the 4-aminoquinoline (4-AQ) antiplasmodials. It was largely used in the past until the *Plasmodium falciparum*, the most lethal malaria protozoan, developed a strong specific resistance against this class of compounds. CQ singly and doubly charged forms are supposed to interfere with the detoxification process of the free heme in the acidic digestive vacuole (DV) released by the digestion of the hemoglobin.[176] It is believed that aminoquinoline drugs act as bio-crystallization inhibitors of hemozoin (or β-hematin), the heme crystal.[177–180] In the presence of 4-AQ drugs, the digestive vacuole releases free heme molecule in the cytosol, where it increases the cellular oxidative stress because the Fe atom is able to catalyse Fenton type reactions.[181,182]

It is widely accepted that the 4-aminoquinoline drugs activity is due to their interaction with the monomeric heme. There is also a large consensus, demonstrated by UV[183], EXAFS[184], NMR[178], MM[178,185] and DFT[186] outcomes, on the relevant role of $\pi\cdots\pi$ stacking interactions between the quinoline ring of the drug and the protoporphyrin pyrrole subunits in the stabilization of the CQ-heme adduct.

Another possible mechanism of action proposed in literature involves a direct Fe-$N_{quinoline}$ coordinative bond. This way of action is supported by solid-state NMR[179] and DFT-EXAFS[187] findings. These two heme:CQ recognition modes ($\pi\cdots\pi$ stacking and Fe–N coordinative bond) are usually believed as mutually exclusive.[188] However, no conclusive evidence has been found yet.

The main purpose of this work is to get further details on the mechanism of action of chloroquine.

Due to the well-known difficulties to understand intermolecular recognition in solution, we analysed the solid-state structure of the dihydrogen phosphate dihydrate salt of diprotonated chloroquine, $CQH_2{}^{2+}(H_2PO_4{}^-)_2\,2\,H_2O$. Analysis of the most important packing features is carried out on the basis of the interaction energies of the molecular

[i]This chapter fully reports the results already published in references [175]

pairs and of the properties of their corresponding electron density (ED) distributions. In particular, topological descriptors have the ability to extract important information on the intermolecular interactions and can be applied in principle both on theoretical and experimental electron density distribution. The most striking structure-determining non-covalent interactions (NCI) are unravelled and correlated with features found in the heme:CQ adduct. To this end, dispersion-corrected DFT calculations were used to compute a reasonable structure for the drug:substrate complex. Understanding the key features of this intermolecular recognition is an important step toward the development of novel effective CQ-based antimalarial drugs, able to overcome the evolved resistance of the *Plasmodium falciparum* protozoan.[189,190]

5.1 Materials and Methods

5.1.1 Crystal Growth and X-ray data collection

Reagent-grade anhydrous powder of N_4-(7-chloro-4-quinolinyl)-N´,N´-diethyl-1,4-pentanediamine (chloroquine) diphosphate salt was purchased from Sigma-Aldrich and it was used without further purification. Single crystals of the dihydrate salt were grown by sol-gel diffusion within a glass tube (⌀ 12 mm) preparing 0.3 M aqueous solution of CQ and using THF as antisolvent. Agarose (1% m/v) was used as gelling agent.[191] Single-crystal X-ray diffraction (SC-XRD) data were collected both at room temperature (RT) and at low-temperature T = 103(2) K on a 3-circle Bruker Apex II CCD diffractometer equipped with an Oxford cryosystems N_2 gas blower, at a nominal source power of 50 kV x 30 mA. 100 % complete data sets, up to a maximum resolution of $sin\theta/\lambda$ = 0.6 $Å^{-1}$ (RT) and 1.0 $Å^{-1}$ (103 K), were obtained. Integration and preliminary data reduction were performed using SAINT[192], while empirical absorption correction and scaling were performed by SADABS[193] and XPREP[194] programs. The crystal structure was solved by direct methods and refined by least-squares against F^2 using the SHELX software package[195]. Hydrogen atoms were located through Fourier difference methods when bonded to heteroatoms, while they were ideally placed according to geometrical factors when bonded to C atoms. Relevant refinement details and agreement statistics can be found in Table 5.1. CCDC 1471834 (103 K) and 1494003 (RT) contain the crystallographic data, provided free of charge by the Cambridge crystallographic data centre.

5.1.2 Solid State DFT Simulations

Periodic DFT (P-DFT) optimization of the experimental structure was performed within a linear combination of gaussian-type function (LCGTF) approach, as implemented in the CRYSTAL14 code[196,197]. Various basis sets and empirically dispersion-corrected[36] Hamiltonians were tested. A 6-31G(d) double-ζ basis set previously optimized for solid-state calculations[23,198,199] was selected in tandem with the hybrid B3LYP[200,201] functional, as this combination ensured the best agreement in terms of root mean squared deviations

Table 5.1: crystallographic details and refinement statistics of the chloroquine diphosphate dihydrate salt ($C_{18}H_{28}N_3Cl^{2+}(H_2PO_4^-)_2 \cdot 2\,H_2O$)[a]

T/K	RT	103(2)
a/Å[b]	9.8350(2)	9.7212(1)
b/Å[b]	16.8654(3)	16.7733(2)
c/Å[b]	15.7859(3)	15.6966(2)
β/deg[b]	105.750(1)	105.1788(2)
V/Å^{3}[b]	2520.12(8)	2470.14(5)
Dx/g cm^{-3}	1.455	1.484
λ/Å, μ/mm^{-1}	0.71073, 0.335	0.71073, 0.342
crystal size/mm^3	0.725, 0.620, 0.425	0.725, 0.620, 0.425
tot. reflns., unique, unique>2σ(I)	45524, 8218, 7164	236057, 20697, 18642
completeness	98.6%	100.0%
$(sin\theta/\lambda)_{max}$/Å^{-1}	0.72	1.0
R_{int}	0.0155	0.0277
Refinement (Shelx)		
ref. parameters, data/parameter	427, 18.04	381, 54.32
R_F^2, wR_F^2, goodness-of-fit (all data)	0.0438, 0.1138, 1.056	0.0427, 0.1294, 1.154
$\Delta\rho_{max/min}/e$ Å^{-2}[c]	+0.91,–0.26	+2.00, –0.89

[a] Monoclinic $P2_1/c$ (Space Group 14), F_{000}=1168 e, M_W=551.89 amu, Z=4 and Z'=1. Z: space group multiplicity; Z': number of formula units per asymmetric unit.

[b] Lattice constant at 103 K were obtained by the least-squares fitting of the crystal orientation matrix against 29733 intense reflections integrated among 4.8 °$\leq 2\theta \leq$ 107.8 °. At RT, the same quantities are 9193 and 4.4 °$\leq 2\theta \leq$ 60.9 °.

[c] High Fourier residuals at RT and T = 103(2) K are partly due to disorder affecting water molecules.

(RMSD) of the relaxed structure with respect to that experimentally observed at 103 K (see Table 5.2). Cell parameters and crystal symmetries were kept fixed during the whole optimization process to those provided by the experiment at 103 K (see Table 5.1).

The starting electron population of the basic N and O centres were initially modified to assign formal charges of +2 to diprotonated chloroquine (CQH_2^{2+}) and –1 to each hydrogen phosphate ion ($H_2PO_4^-$). The tolerances of numerical approximation in evaluating the Coulomb and the exchange series were set to standard values.[196,197] A 40 % mixing of the Fock matrices and an eigenvalue level shift of 0.4 Hartree were applied to accelerate convergence.[196,197] The reciprocal space was sampled according to a regular sublattice defined by 2 points on each axis in the irreducible Brillouin zone. To optimize the nuclear positions, the minima on the potential energy surface were located through a modified form conjugate gradient algorithm proposed by Schlegel.[202] All the gradients were computed analytically. Default thresholds on the minimum force, the root-mean square (RMS) force, the minimum atomic displacement and the RMS atomic displacement were employed for the geometry optimization.[196,197]

Input orientations of the water molecules were obtained through a partial optimization step, where only the orientation of the hydrogen atoms was optimized. Then, all the molecules in the asymmetric unit were left free to relax without constraints. The lattice

Table 5.2: Root mean square deviations (RMSDs, in fractional coordinates) of the periodic DFT optimized structures at different level of theory with respect to the experimental geometry and torsional angles C5-C4A-C4-N9 (in degree). Dispersion correction proposed by Grimme[36] was used for all these calculations. The atom labels are the same of Figure 5.1.

Method	RMSD	tors(C5-C4A-C4-N9)/ °
Experimental	//	-0.040
B3LYP/6-31G	0.014	-0.054
BP86/6-31G	0.019	-1.550
BP86/6-311G(p,d)	0.021	-3.275
B3LYP/pob-TZVP	0.028	-1.752
PBE/6-314G	0.018	-4.790

parameters were not refined to have interaction energies and geometries as much as possible comparable to the experimental structure. Moreover, since the collected data suffer from diffuse scattering and disorder issues, all the analyses here reported were performed on the DFT optimized structure.

Topological analysis of the periodic electron density, $\rho(\boldsymbol{r})$, was performed through the TOPOND[203] module of CRYSTAL14, according to the quantum theory of atoms in molecules (QTAIM)[4,6–8].

5.1.3 Intermolecular Interactions

Intermolecular interaction energies were computed from the optimized periodic structure, considering all the symmetry-independent molecular pairs with centre-of-mass distance less than 20 Å. Single-point DFT calculations were carried out on each pair extracted from the crystal at the Grimme-corrected[36] B3LYP/pob-TZVP[91] level of theory using the Gaussian09 program package[89]. All the pair energies were corrected for basis set superposition error (BSSE) by the counterpoise method[166] and for relaxation energy.[204–207]. Energies of the intermolecular interactions were also evaluated using Spackman's experimental charge density approach (ECDA)[38,44–47] using the software PAMoC.[208] ECDA analysis provides a reasonable energy decomposition scheme into electrostatic, dispersion and Pauli repulsion terms. ECDA atomic-atomic contributions and natural bond order (NBO)[209] analyses were employed to estimate the hydrogen bonds (HBs) energies. PAMoC software was also employed to extract QTAIM descriptors from the gas-phase calculated electron density. NCIs were studied also through visual inspection of isosurfaces of the reduced density gradient (RDG) descriptor[29,31], computed by the NCImilano code.[210] The step size of the grids was set at 0.1 Å and RDG was calculated only for those molecular regions representative of intermolecular interactions, namely with $0.00 \leq \rho(\boldsymbol{r}) \leq 0.05$ *a.u.*. The $\rho(\boldsymbol{r}) \cdot sign(\lambda_2)$ quantity was plotted on RDG-isosurfaces with an isovalue of 0.4 if not otherwise specified. The recipes here employed highlight different aspects of the recognition process of CQ.

5.1.4 Chloroquine:Heme Adduct

The drug:heme adduct proposed in Reference [187] was here further optimized to improve its quality and to be able to compare it with the solid-state results on the same ground. *In vacuo* optimization at m-GGA B3LYP/pob-TZVP[91] level of theory was performed using Gaussian09 software[89]. Grimme dispersion correction was taken into account.[36] The curvature of the potential energy surface was evaluated at the stationary point to verify the full convergence of the optimization process into a stable minimum. Electron density and interaction energy analyses were performed as already described above.

5.1.5 Molecular Dynamics

The calculations on the crystalline salt were carried out at finite temperature with the Gromacs 5.0.2 package[211] using the SPC model[212] to account for the hydration water molecules. The all-atom OPLSAA force field[213] was employed throughout. Periodic boundary conditions were defined through an approximately cubic 4x2x2 supercell, while the initial positions of the water molecules were the same as in the final DFT model. All the covalent bonds were constrained using the LINCS algorithm,[214] while improper dihedrals were suitably defined to force the quinoline ring of chloroquine molecule to remain flat, in agreement with the SC-XRD geometry. Newton equations were solved through the leap-frog integrator, while the particle mesh Ewald (PME) method[215] was used to treat long-range electrostatics in conjunction with a 10 Å cutoff for non-bonded interactions. First, the system underwent an initial energy minimization step through the steepest descent algorithm until all the forces were lower than 100 $kJmol^{-1}n^{-1}m$. Then, pre-production equilibration was carried out for 500 ps with a time step of 2 fs in NVT conditions (T = 103 K), followed by a 500 ps-long run in NPT conditions (T = 103 K, p = 1 bar) under isotropic pressure coupling. A modified Berendsen thermostat[216] and the Parrinello-Rahman barostat[217] were employed to restrain thermodynamic variables to the reference values, considering chloroquine and inorganic species ($H_2PO_4^-$, H_2O) as independent coupling groups. Finally, a 20 ns-long P-MD run was carried out under NPT conditions (T = 103 K, p = 1 bar).

5.2 Crystal Structure of 103 K Chloroquine Dihydrogenophosphate Dihydrate

The crystal structure of the $CQH_2^{2+}(H_2PO_4^-)_2\,2\,H_2O$ salt was solved for the first time by Preston & Stewart (PS) in 1970[218]. In the late 80s, Karle & Karle (KK) re-determined the structure in a more accurate way.[219] The data collection here described confirms, at highest resolution and accuracy, the main structural features previously described. However, at variance with KK, it also evidences disorder involving $H_2PO_4^-$ ions and water molecules.

Fully protonated CQ crystallizes in the monoclinic $P2_1/c$ lattice (space group 14), with one CQH_2^{2+} ion, two $H_2PO_4^-$ ions and two H_2O molecules in the asymmetric unit (ASU,

Figure 5.1).

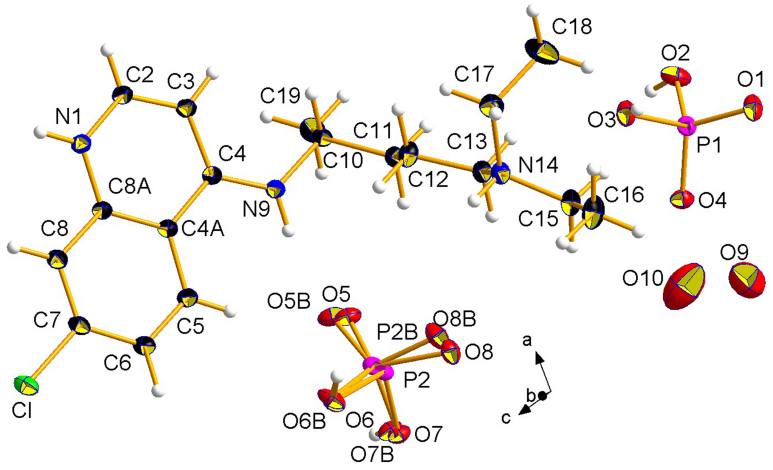
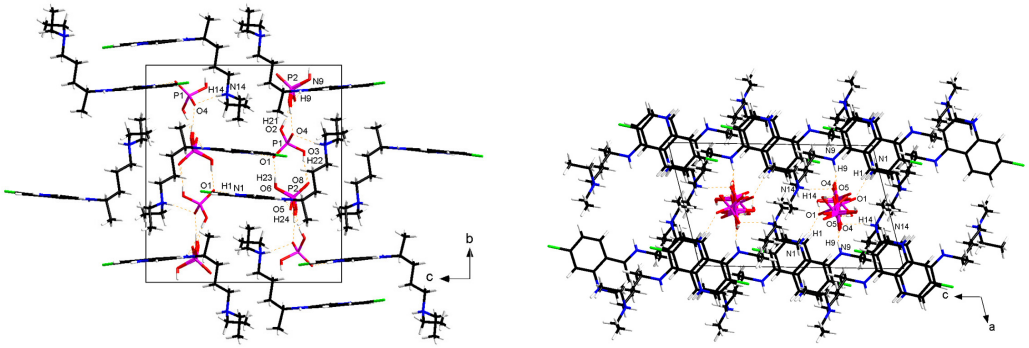


Figure 5.1: Asymmetric unit (ASU) of chloroquine diphosphate salt at T = 103 K. Thermal ellypsoids are drawn at 50 % probability level. Color code: H: white, C: black, N: blue, O: red, P: purple, Cl: green.

Infinite chains of phosphate anions are formed along the b-axis. Chloroquine molecules interact with the phosphate groups through hydrogen bonds, some of them classifiable as charge assisted hydrogen bonds (CAHBs). The hydrocarbon side chains wrap around a single phosphates pillar, setting up a sort of helical arrangement, reinforced by strong hydrogen bonds. Figure 5.2 shows the molecular disposition and Table 5.3 summarises the main interactions.



(a) a axis (b) b axis

Figure 5.2: wired-and-stick representation of the crystal packing of chloroquine diphosphate salt at 103 K. Relevant hydrogen bond interactions are shown as red dotted lines.

As expected, CQH_2^{2+} shows a neat preference for the direct interactions with the $H_2PO_4^-$ counterions (Table 5.3). Three short N-H$\cdots$O interactions are present between the charged CQ molecule and the phosphate. O-H$\cdots$O hydrogen bonds hold together the phosphate units along the infinite chains, with an average donor-acceptor distance of 2.56(3) Å. Just weak C-H$\cdots$O contacts, deemed not significant for the overall structure

Table 5.3: Geometrical descriptors (Å, °) of the N-H$\cdots$O and O-H$\cdots$O hydrogen-bonded contacts shown in Figure 5.2, as estimated from the X-ray diffraction experiment at T = 103(2) K. When meaningful, estimated standard deviations are given in parentheses.

Bond[a]	d(D$\cdots$A)	d(D-H)	d(H$\cdots$A)	∠(D-H$\cdots$A)	Symm. Op.
		$CQH_2^{2+}\cdots H_2PO_4^-$			
N1-H1$\cdots$O1	2.6630(6)	0.90(2)	1.77(2)	167(2)	X+1, Y, Z+1
N9-H9$\cdots$O5	2.898(4)	0.85(2)	2.07(1)	162(1)	X, Y, Z
N14-H14$\cdots$O4	2.7568(7)	0.84(2)	1.93(2)	165(1)	-X+1,-Y, -Z+1
		$H_2PO_4^-\cdots H_2PO_4^-$			
O2-H21$\cdots$O5	2.544(4)	0.77(2)	1.78(2)	172(2)	X, -Y+1/2, Z – 1/2
O3-H22$\cdots$O8	2.531(2)	0.70(2)	1.84(3)	167(3)	-X+1, -Y, -Z+1
O6-H23$\cdots$O1	2.59(1)	0.90(2)	1.69(2)	174(2)	-X+1, -Y, -Z+1
O7-H24$\cdots$O4	2.59(1)	0.92(2)	1.67(2)	175(2)	X, -Y+1/2, Z+1/2
		$CQH_2^{2+}\cdots H_2O$ and $H_2O\cdots H_2O$			
O9$\cdots$O8	3.191(6)	//	//	//	X, -Y+1/2, Z+1/2
O10$\cdots$O3	2.955(2)	//	//	//	-X+1, Y-1/2, -Z+1/2
O9$\cdots$O10	2.852(2)	//	//	//	X, Y, Z

[a] Geometrical parameter derived from the spherical atom model.

stability[220], are set up among $CQH_2^{2+}$and H_2O molecules.

In each unit cell, dihydrogen phosphate ions form two symmetry-related infinite chains parallel to the monoclinic b axis (Figure 5.2(b)), while the quinoline system of CQH_2^{2+} roughly lies in the orthogonal (a,c) plane (Figure 5.2(a)). Each P1 $H_2PO_4^-$ ion accepts a couple of strong CAHBs involving both the N1 and N14 formally charged amino groups from two distinct chloroquine molecules (Figure 5.2 and Table 5.3). On the contrary, the P2 ion is involved just in a weaker N-H$\cdots$O interaction with the uncharged N9 amine. It is interesting to notice that the disorder found at 103 K on the dihydrogen phosphate molecule is only localized on the P2 $H_2PO_4^-$, while the P1 phosphate group is perfectly ordered. This difference could be related to the different number and type of hydrogen bonds set up with the chloroquine molecule.

It is worth mentioning that pairs of inversion-related $H_2PO_4^-$ pillars in adjacent unit cells are bonded by a three-dimensional zig-zag motif of hydrogen-bonded bridging water molecules (Figure 5.3).

5.2.1 On the Crystal Disorder

According with KK[219], we were not able to unequivocally locate hydrogen atoms around the two water oxygen atoms, O9 and O10, neither at RT nor at 103 K. This is a sign of another kind of disorder which, in this case, intrinsically concerns the H atoms of the water molecules due to HB frustration. It should be noted that the structural determination by KK relied on a low-order ($2\theta < 45$ °) and intense ($|Fo| > 3\ \sigma Fo|$) quarter-of-sphere subset of data not corrected for absorption. Therefore, subtle features related to symmetry breaking were likely not detectable in their experiment. A disordered model refining to partial

occupations as high as 0.65(4) and 0.46(4) for the P1 and P2 $H_2PO_4^-$ ions significantly improves the agreement factors against all the measured structure factor amplitudes at room temperature. On the other hand, at T = 103(2) K, the P1 $H_2PO_4^-$ ion is completely ordered, while a residual disorder refining to partial occupancy of 0.66(1) is still present on the P2 ion. The gradual evolution toward a more ordered structure while T is lowered suggests that the disorder of $H_2PO_4^-$ ions is dynamic, at least in part, in nature. In any case, no structural phase transitions were detected between RT and ≈ 100 K.

The orientations of one half of the H atoms of the water molecules are free due to the lacking of an acceptor in their close neighbourhood, while it is not possible to unequivocally define to which oxygen most of the other H atoms are covalently bonded. Likely, all the possible motifs are present in the crystal and this translates into disorder which should be both dynamic and static as it persists down to ≈ 100 K. Besides, it is easy to see that perfect $P2_1/c$ symmetry is incompatible with the regular H-bond network sketched in Figure 5.3. Indeed, any H atom bonded to inversion-related O9 atoms along the O9···O9 direction would produce a O9–H···H–O9 steric clash. This means that the $P2_1/c$ symmetry of the crystal emerges owing to the statistical average of allegedly equivalent configurations in the H-bonded frustrated water motif.

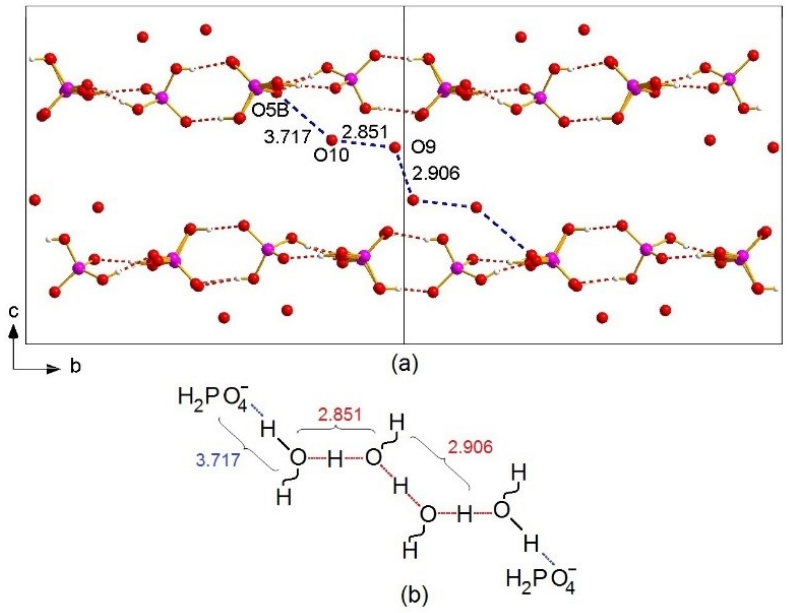


Figure 5.3: (a): $H_2PO_4^-$-water motif in the chloroquine diphosphate salt. Symmetry-independent oxygen-oxygen distances are given in Å. (b): Same as (a), outlining the frustrated H-bond motif. H-bonds are drawn in blue and covalent bonds in black. Dotted red lines mark interactions which might be equivalently either covalent or H-bonds. Tilde lines mean covalent bonds whose H∼O···H···O dihedrals are unknown.

The picture is further complicated by the fact that at 103 K a ≈ 2 e Å^{-3} large residual Fourier peak is evident roughly halfway between O9 and O10 (Table 5.1 and 5.4). At RT the same peak was also detectable in all the samples analysed (Table 5.4), even though it

was much less intense and hardly distinguishable from the Fourier noise. This might be somewhat related to the disordered hydrogen between the O9 atoms (O9–H$\cdots$O9 $\Longleftrightarrow$ O9$\cdots$H–O9); however, on the basis of its high intensity, we can not exclude that it could be also due to non-stoichiometry of the co-crystallized solvent. The maximum residual Fourier peak lies $\approx$ 2.8-2.9 Å far from the O3 atom of the P1 $H_2PO_4^-$ ion, meaning that the latter could probably stabilize a single disordered water molecule with partial occupancy in this position through a strong hydrogen bond. The fact that the coordinates of the residual slightly change depending on either the temperature or the sample considered (Table 5.4) corroborate this hypothesis.

Table 5.4: Distances, angle and intensity of the Fourier peak (FP) located between the two oxygen O9 and O10 of the water molecules in two different samples (I and II) and at different temperature. Sample I is the sample here discussed; sample II was crystallized by means of different experimental conditions (liquid-liquid diffusion technique, pH buffered at 7.0).

Samples	d(O9-FP)	d(O10-FP)	α(O9-FP-O10)	I / $e \cdot$ Å^{-3}
I - 100 K	1.520	1.375	160.16	2.00
I - RT	1.506	1.436	154.17	0.91
II - RT	1.253	1.672	168.98	0.27

Finally, it should be also noted that all the tested specimens suffered by strong diffuse scattering (DS) up to intermediate angles. DS is probably correlated with both static and dynamic issues, in turn related with both dihydrogen phosphate disorder and water frustration. The simultaneous occurrence of these issues hampered us from obtaining a reliable experimental ED also at 103 K. In any case, a resolution of 1.0 Å^{-1} in $sin\theta/\lambda$ at 103 K (see Table 5.1) is sufficient to provide reliable geometric and Debye-Waller (DB) parameters.

5.3 A Quantum Model for the Chloroquine Dihydrogen Phosphate Dihydrate Salt

The aim of this study is to reveal the role of the non-covalent interactions that dominate the molecular recognition of the chloroquine in the solid state as a model for the possible interatomic interaction in solution between CQ and the heme molecule. Assuming that NCIs are governed by intrinsic ED properties such as electro-/nucleo-philicity, electrostatics, ability of setting up hydrogen bonds, *etc.*, the hypothesis that the crystal structure should qualitatively mimic the essential aspects of the drug-substrate recognition mode is valid.[221–223] In this respect, it should be noted that the action of CQ does not directly involve any metabolic process and relies on purely local chemical factors.[187] Therefore, we first study the crystal structure of the CQH_2^{2+} di-cation which is the prevailing form in the acidic digestive vacuole of *Plasmodium* protozoan. The presence of disorder, diffuse scattering and hydrogen bond frustration (see Section 5.2.1) prevents the possibility

to study the experimental electron density at T = 103 K. To overcome this limitation, we performed quantum mechanical calculations (as explained in Section 5.1.2) to obtain the theoretical charge density. Since the water molecules have just a minor influence on the crystal structure and on the CQ conformation, the full $P2_1/c$ lattice symmetry was exploited in periodic DFT simulations, implying that the frustrated H_2O-chain motif described in Section 5.2 is broken.

Attractive interactions between oppositely charged groups are known to possibly dominate the heme:CQ recognition process.[207] Accordingly, dihydrogen phosphate ions should represent a suitable model for interactions with the propionate groups of heme in solution,[187] while the stacking interactions between adjacent CQH_2^{2+} molecules along the b axis (Figure 5.2) might serve as a probe for the ability of the drug to similarly approach the flat protoporphyrin ring.

5.3.1 On the Validity of the Proposed Periodic DFT Model

To check the reliability of the hypothesis that the water molecules have just a minor influence, a perfectly ordered $P2_1/c$ structure was simulated through all-electron DFT optimizations, starting from the X-ray geometry. Cell edges and angles were kept fixed at the ones estimated at 103 K. To define a sensible starting point, water hydrogen atoms were arbitrarily placed at reasonable positions, according to one of the possible equivalent arrangements in the frustrated structure, while the sites with the highest occupation factors of the disordered P2 $H_2PO_4^-$ ion were selected. Then, all the atomic coordinates were fully relaxed without constraints. The optimized structure and the experimental one are essentially identical: the overall RMSD is 0.55 Å, mostly due to differences in the position of water molecules. Indeed, as the inversion centre is incompatible with the regular H-bonded motif of water, during the simulation the facing O9 atoms are pulled apart to set up a stable NCI network which at the same time satisfies the symmetry requirements of the $P2_1/c$ structure. This results in shifts of O9 and O10 as high as $\approx$ 0.8 Å which eventually destroy the bridging motif between $H_2PO_4^-$ pillars (Figures 5.2 and 5.3). However, the rest of the structure remains basically unaffected. This evidence confirms the validity of our original assumption – namely that water is not significant in determining the intermolecular recognition pattern of CQH_2^{2+}.

5.3.2 Solid-State Molecular Dynamics to Test Water Disorder

To disentangle possible non-stoichiometry issues from dynamic effects related to acoustic phonon modes in the crystal, we performed periodic molecular dynamics P-MD simulations on the $CQH_2^{2+}(H_2PO_4^-)_2\,2\,H_2O$ salt structure at T = 103(2) K.

It should be stressed that the periodic molecular dynamics simulations are completely classic, while quantum effects should be explicitly accounted for to quantitatively describe hydrogen dynamics in the solid state.[224] Indeed, we can not exclude that in the salt a proton is dynamically exchanged between the extremities of the water chain, a process which

might partly occur through quantum tunneling effects. However, classical calculations could be employed at least to qualitatively assess whether large translations and/or rotations of co-crystallized solvent are possible in condensed phase.[225] This P-MD simulation was carried out for 20 ns at T = 103 K, starting from a stoichiometric structure similar to the experimental one in conjunction with an arbitrarily chosen configuration of HBs in the water chain. The last geometry of the thermalized P-MD structure at t = 20 ns is very similar to the experimental one. The backbone conformation of the chloroquine molecule remains the same, with minor differences concerning, as expected, the rotations of the terminal N-ethyl groups. Overall, the average root mean square displacement (RMSD) of CQH_2^{2+} with respect to the experimental structure is quite low and amounts to ≈ 0.3 Å. Slightly larger RMSD estimates are associated to both $H_2PO_4^-$ groups (0.7 Å and 0.6 Å) and water molecules (0.5 Å and 0.7 Å). As expected, this implies that no collective diffusion of $H_2PO_4^-$ or of water takes place. They rather undergo sharp displacements, ≈ 0.21(2) Å large in amplitude in terms of average root mean square fluctuations (RMSF), bouncing back and forth around their crystallographic positions. To the sake of comparison, much larger displacement amplitudes, up to ≈ 1 Å, were found in the hydrated theophylline crystal by classical Monte Carlo simulations.[225] A more complete picture can be given by examining the rotational correlation functions C(t). The latter describe the rotational "memory" of an intermolecular ensemble[225]: in a fully ordered lattice, the molecules will retain their reciprocal orientation, implying C(t) values always very close to 1, while in liquids the values quickly decay to zero from any reference frame.

Figure 5.4 shows that in the present case all the molecules conserve most of their rotational correlation, with average <C(t)> values in the last 10 ns of the simulation as high as 0.999 (CQH_2^{2+}), 0.994 ($H_2PO_4^-$) and 0.974 (H_2O).

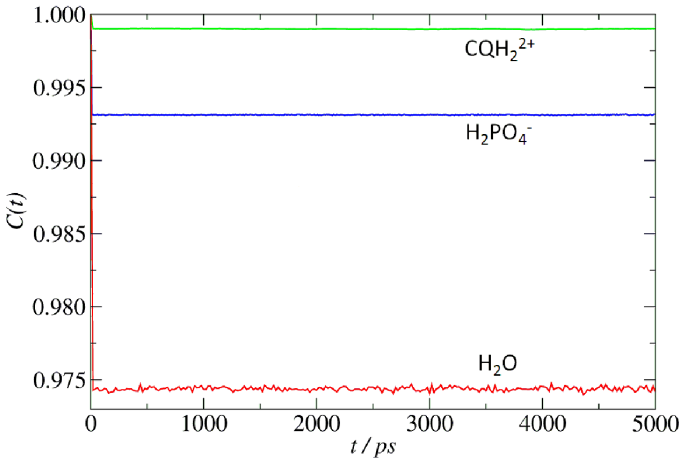


Figure 5.4: rotational correlation functions, C(t), for independent molecules in the title compound (green: chloroquine; blue: phosphate; red: water) over the P-MD simulation. Time zero is the simulation starting point, similar to the experimental structure.

In other words, no significant rotational disorder is present, even for water molecules,

whose maximum loss of rotational correlation amounts to less than 3 %. At the same time, the ≈ 0.2 Å wide average displacements are too low to provide by themselves a satisfactory explanation for the intense Fourier residual between O9 and O10 at 103 K, which lies at ≈ 1.5 Å from O9 and at ≈ 1.3 Å from O10. This also rules out dynamic H_2O disorder, substantiating the picture where non-stoichiometric water might coexist with static disorder of H atoms due to HB frustration (Figure 5.3). It should be noted, however, that this does not mean that hydrogen bond contacts among H_2O units are somewhat weak. On the contrary, surviving of the H_2O relative orientation suggests that a complex and strong HB network is set up to provide a rather stiff water-dihydrogen phosphate framework where $CQH_2{}^{2+}$ can be allocated. Neutron diffraction experiments, even at lower T, are necessary to provide some models to which computational findings can be compared.

5.4 Electron Density Analysis of In-Crystal Intermolecular Interactions

5.4.1 Chloroquine-Dihydrogen Phosphate

$CQH_2{}^{2+}\cdots H_2PO_4{}^-$ interactions are invariably stabilizing owing to strong Coulombic attraction between opposite charges. Figure 5.5 shows the interaction geometries of some most attractive $CQH_2{}^{2+}\cdots H_2PO_4{}^-$ pairs which exploit relevant intermolecular hydrogen bond contacts, as summarized in Table 5.5.

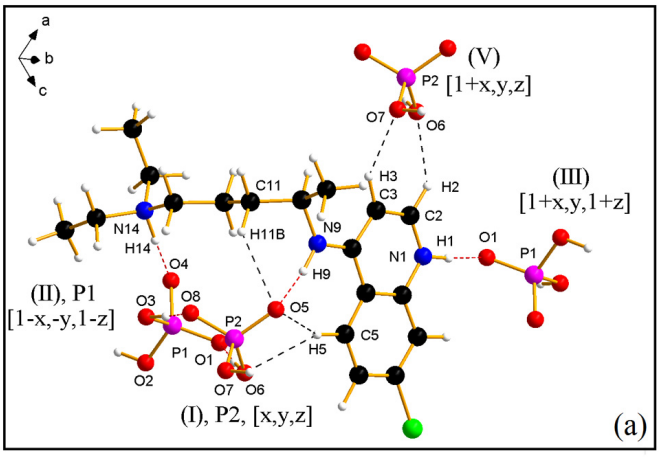


Figure 5.5: most attractive $CQH_2{}^{2+}\cdots H_2PO_4{}^-$ pairs. Roman numerals label pairs in descending order of DFT interaction energies. Atomic interaction lines are shown as dashed lines.

In this context, the contacts are defined relevant if an atomic interaction line (AIL) exists and the charge density at the bond critical point, $\rho_{BCP} \geq 10^{-2}$ *a.u.*. If not otherwise specified, the QTAIM investigation was always performed on the ED of molecular pairs extracted from the crystal. However, the point topological properties in the solid state,

Table 5.5: Bond path lengths, $d_{H\cdots A}$, and topological properties at the bond critical points (BCPs) of relevant non-covalent bonded contacts, as shown in Figure 5.5

Contact[a]	$d_{H\cdots A}$	ρ_{BCP}	$\nabla^2\rho_{BCP}$	G_{BCP}	V_{BCP}	H_{BCP}	BD[b]
Pair I, $CQH_2^{2+}\cdots H_2PO_4^-$ [P2,*E*][c]							
N9–H9$\cdots$O5[d]	**1.8674**	**0.031**	**0.102**	**15.6**	**-15.2**	0.4	**0.02**
C5–H5$\cdots$O5	2.4455	0.011	0.037	4.8	-3.9	0.9	0.13
C5–H5$\cdots$O6	2.4824	0.010	0.034	4.5	-3.6	0.9	0.15
C11–H11B$\cdots$O5	2.5569	0.009	0.031	4.1	-3.2	0.9	0.14
Pair II, $CQH_2^{2+}\cdots H_2PO_4^-$ [P1,*i*][c]							
N14–H14$\cdots$O4[d]	**1.7419**	**0.045**	**0.115**	**22.1**	**-26.1**	-4.0	**-0.14**
Pair III, $CQH_2^{2+}\cdots H_2PO_4^-$ [P1,*E*][c]							
N1–H1$\cdots$O1[d]	**1.6279**	**0.056**	**0.129**	**27.5**	**-34.8**	-7.3	**-0.21**
Pair V, $CQH_2^{2+}\cdots H_2PO_4^-$ [P2,*E*][c]							
C2–H2$\cdots$O6	2.2487	0.015	0.057	7.2	-5.5	1.7	0.18
C3–H3$\cdots$O7	2.2700	0.014	0.053	6.8	-5.2	1.6	0.17

[a] $d_{H\cdots A}$ is expressed in Å; ρ_{BCP}, $\nabla^2\rho_{BCP}$ and the bond degree parameter (BD) are expressed in *a.u.*; G_{BCP} and V_{BCP} are expressed in kcal mol^{-1}Å^{-3}.
[b] Bond degree parameter (BD) according with Espinosa and co-workers, defined as H_{BCP}/ρ_{BCP}, where $H_{BCP} = G_{BCP} + V_{BCP}$.[27,167]
[c] Phosphate number and crystal point symmetry involved in the costruction of the pair reported in square brackets.
[d] Charge-assisted hydrogen bonds (CAHBs) in bold.

obtained using TOPOND software, are in good agreement with those evaluated in vacuo. Both geometry and QTAIM point properties (Table 5.5) confirm that N-H$\cdots$O HBs are significantly stronger than C-H$\cdots$O ones. In fact, the N-H$\cdots$O distances (1.6–1.9 Å) are much shorter (up to 40 %) than the sum of H and O standard van der Waals radii (2.7 Å according to Bondi[226]). The shortest N1-H1$\cdots$O1 and N14-H14$\cdots$O4 interactions are charge-assisted hydrogen bonds since both the protonated amine nitrogens and the phosphate oxygens are formally charged. These contacts show high ρ_{BCP} and negative local energy density H_{BCP} ($H_{BCP} = G_{BCP} + V_{BCP}$) values, differently to N9-H9$\cdots$O5 bond and all the other C-H$\cdots$O reported in Table 5.5. The negative H_{BCP}/ρ_{BCP} ratio, called bond degree (BD) parameter, is usually associated to the incipient formation of covalent bonds[27,167] and thus can be considered as a symptom of partial H$\cdots$O covalency.

The secondary amine substituent of the quinoline ring, N9H9, is involved in the third strong HB with O5 atom of the P2 $H_2PO_4^-$ ion. This bond is expected to be weaker than the other two bonds because the NH group is formally neutral. The electron density associated with its BCP is about 3 times higher than that of C-H$\cdots$O bonds and has energy density H_{BCP} close to zero (see Table 5.5).

A very good agreement emerges between the QTAIM picture and that based on the reduced density gradient. As expected, strong N-H$\cdots$O HBs are invariably associated to contracted, disc-shaped RDG isosurfaces with $\rho(\boldsymbol{r}) \cdot sign(\lambda_2) \ll 0$ (Figure 5.6).[29,31] On the contrary, C-H$\cdots$O interactions show larger and more structured isosurfaces dominated by slightly negative or neutral λ_2 curvatures. It is also worth noting that some very weak

intramolecular H···H contacts are evident in the RDG plots, also in those cases where no AILs are found according to QTAIM.

For instance, one can see contacts in pair (III) between H3···H10 and H5···H9, as shown in Figure 5.6. The latter regions are quite common for weak close-shell interactions,[31,227] but they can hardly be considered significant for molecules self-recognition in solid state. In fact, these interactions are characterized by almost zero or positive values of $\rho(\boldsymbol{r}) \cdot sign(\lambda_2)$.

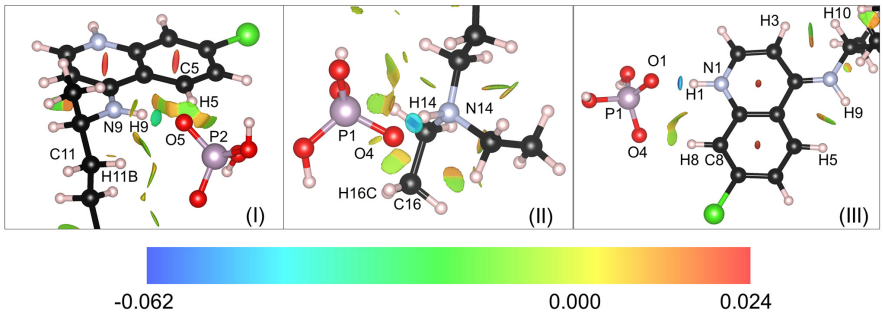


Figure 5.6: $\rho(\boldsymbol{r}) \cdot sign(\lambda_2)$ function plotted on RDG isosurfaces in the $CQH_2{}^{2+}\cdots H_2PO_4{}^-$ most stable pairs (I-III). RDG isosurfaces level is set equal to 0.4. $\rho(\boldsymbol{r}) \cdot sign(\lambda_2)$ values on the chromatic scale are given in *a.u.*.

5.4.2 Chloroquine Self-Recognition

The most repulsive $CQH_2{}^{2+}\cdots CQH_2{}^{2+}$ pair involves inversion-related π-stacked chloroquine molecules and it is shown in Figure 5.7. Various C···X (X = C, N, Cl) and C-H···Cl interactions, each associated to a topological AIL (Table 5.6), are present in this pair.

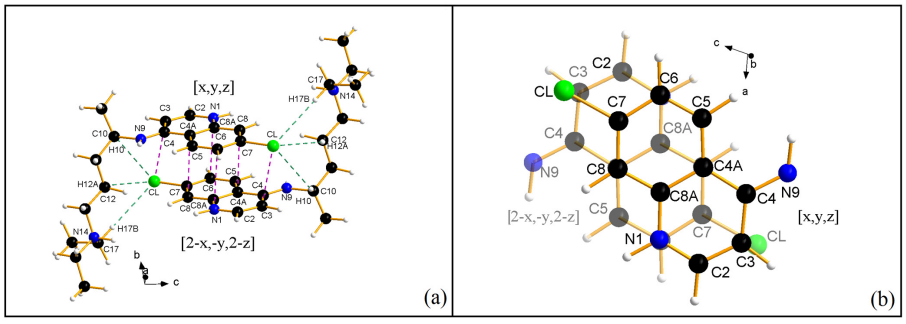


Figure 5.7: (a) most repulsive $CQH_2{}^{2+}\cdots CQH_2{}^{2+}$ pair (d = 5.57 Å, $E_{int,DFT}$ = 153 kcal mol^{-1}) at T = 103 K. AIL are shown as dotted lines. (b) as (a), with focus on the quinoline rings viewed along the C···X AIL.

The ρ_{BCP} of the C···X is approximatively one order of magnitude lower with respect to the N-H···O and C-H···O bonds shown in Table 5.5. However, while C-H···Cl contacts show topological parameters very similar to those of the weakest C-H···O interactions, the C···X ones are systematically a bit stronger in terms of ρ_{BCP} and H_{BCP} estimates. From a geometrical viewpoint, the two flat rings are head-to-tail packed with a least-

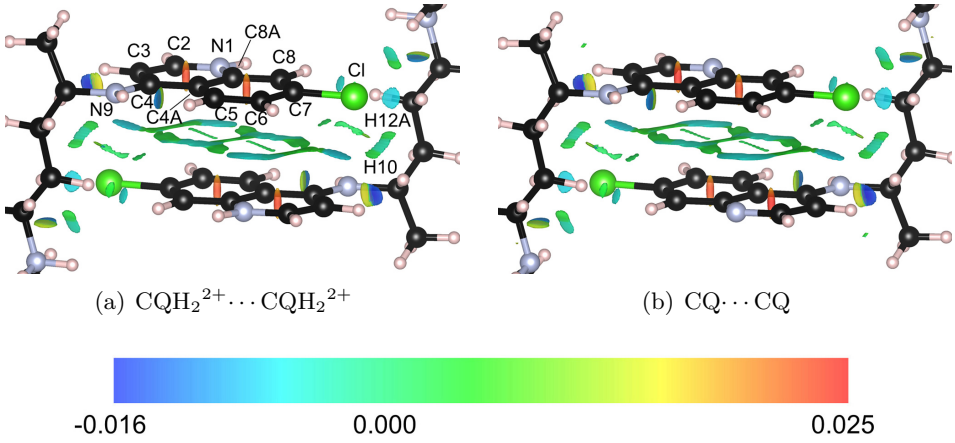
Table 5.6: Bond path lengths, $d_{H\cdots A}$, and topological properties at the bond critical points (BCP) of relevant non-covalent bonded contacts, as shown in Figure 5.7

Contact[a]	$d_{H\cdots A}$	ρ_{BCP}	$\nabla^2\rho_{BCP}$	G_{BCP}	V_{BCP}	H_{BCP}	BD[b]
Pair I, $CQH_2^{2+}\cdots CQH_2^{2+}$ $[i]$[c]							
C3$\cdots$Cl	3.4497	0.007	0.022	3.0	-2.4	0.6	0.13
C6$\cdots$N1	3.4857	0.005	0.021	2.4	-1.6	0.8	0.25
C4A$\cdots$C8	3.5204	0.004	0.019	2.1	-1.2	0.9	0.32
C12-H12A$\cdots$Cl	3.0079	0.007	0.021	2.7	-2.2	0.5	0.12
C17-H17B$\cdots$Cl	3.2650	0.003	0.010	1.3	-1.0	0.3	0.13
C10-H10$\cdots$Cl	3.3902	0.003	0.011	1.3	-0.9	0.4	0.23

[a] $d_{H\cdots A}$ is expressed in Å; ρ_{BCP}, $\nabla^2\rho_{BCP}$ and the bond degree parameter (BD) are expressed in *a.u.*; G_{BCP} and V_{BCP} are expressed in kcal mol^{-1}Å^{-3}.
[b] Bond degree parameter (BD) according with Espinosa and co-workers, defined as H_{BCP}/ρ_{BCP}, where $H_{BCP} = G_{BCP} + V_{BCP}$.[27,167]
[c] Crystal point symmetry involved in the costruction of the pair reported in square brackets.

squares distance between the condensed quinoline moieties as high as 3.37 Å, very similar to the van der Waals contact distance between C atoms ($\approx$ 3.4 Å).[226] Such stacking motifs are very common in systems containing terminal flat aromatic rings bonded to somewhat long hydrocarbon chains, such as in DNA[153] and DNA-protein complexes.[228] These motifs are usually interpreted as a possible signature of $\pi\cdots\pi$ interactions,[37,153,228] while the properties of cage critical points (CCPs) emerging from cyclic AIL patterns in the stacking region are usually correlated to the stabilization energies of the interacting pairs[37].

The presence AIL between facing rings implies that RDG isosurfaces assume a strongly structured motif, as evident in Figure 5.8.



(a) $CQH_2^{2+}\cdots CQH_2^{2+}$ (b) $CQ\cdots CQ$

Figure 5.8: (a) $\rho(\boldsymbol{r}) \cdot sign(\lambda_2)$ function plotted on the RDG isosurfaces (isolevel = 0.4) in the most repulsive $CQH_2^{2+}\cdots CQH_2^{2+}$ pair. (b) Same as (a) for a model $CQ\cdots CQ$ pair, where CQ molecules are placed at the same distance of CQH_2^{2+}. $\rho(\boldsymbol{r}) \cdot sign(\lambda_2)$ values on the chromatic scale are given in *a.u.*.

A bicyclic RDG pattern mirroring the shape of the quinoline backbone appears roughly

halfway between the two chloroquine molecules. The $\rho(\boldsymbol{r}) \cdot sign(\lambda_2)$ quantity is invariably very low and often weakly positive, meaning that $\rho(\boldsymbol{r})$ in the internuclear region is also very low and with a λ_2 curvature close to zero. In this respect, it is instructive to compare Figure 5.8(a) with Figure 5.8(b), where $\rho(\boldsymbol{r}) \cdot sign(\lambda_2)$ is plotted onto the same RDG isosurface computed for a couple of neutral chloroquine molecules, kept exactly at the same positions which their protonated counterparts have in the crystal. The above described features are strictly conserved in the deprotonated molecular pair, together with the corresponding pattern of AILs. This implies that this feature does not correlate with the acid-base status of chloroquine. Indeed, the electrostatic contribution plays here the major role and it is known to be not directly reflected into point properties of the ED topology.[229]

5.5 Interaction Energies

The analysis of interaction energies of molecular pairs at their in-crystal geometries complements the topological picture based on the ED and RDG features discussed in Section 5.4. Table 5.7 summarizes the DFT and semiempirical ECDA interaction energies of relevant pairs in the first coordination shell of CQH_2^{2+}, together with their decomposition into electrostatic (E_{es}), dispersive (E_d) and repulsive (E_r) contributions. The energies of the hydrogen bonds, as estimated through the ECDA model and the natural bond orbital (NBO) analysis of the wavefunction, are also shown in Table 5.7.

The agreement between DFT and ECDA interaction energies is qualitatively good. Absolute differences do not exceed 6.3 kcal mol^{-1}for the closest hydrogen-bonded (I)–(III) chloroquine-phosphate pairs, comparably to what found for p-nitroaniline by Abramov *et al.*.[230]

Larger differences are found for those pairs where no strong interactions occur. The discrepancy between the two approaches in the $CQH_2^{2+}\cdots H_2PO_4^-$ pair (V) is almost the double with respect to the previous case and it is as high as 11.1 kcal mol^{-1}. Similarly, the differences for the charged and neutral CQ pairs are equal to $\approx$ 17 and 21 kcal mol^{-1}, respectively. The reason behind the discrepancy is ascribable to a general underestimation of the electrostatic contributions in the atom-centred multipolar Buckingham approach, on which the ECDA partitioning scheme relies.

In any case, it is not demanded here that the ECDA estimates are quantitatively identical with the higher-grade DFT ones. Rather, the former is used as a model to analyze the possible role of various dispersion, repulsion and electrostatic terms in determining the overall interaction energy on relative grounds.

5.5.1 Coulombic Interactions and the Role of Hydrogen Bonds

The trend of the Coulomb interactions between opposite charges against the centre of mass distances of the $CQH_2^{2+}\cdots H_2PO_4^-$ molecular pairs within d = 20 Å is shown in Figure 5.9.

Table 5.7: Interaction energies *vs.* the centre of mass distance d of the most attractive and repulsive molecular pairs (Figures 5.5 and 5.7) in the first coordination shell of chloroquine in the title compound. Distance values are given in Å and energies in kcal mol^{-1}

d	E_{DFT}	E_{ECDA}	E_{es}	E_r	E_d	Hydrogen Bonds	E(2)[a]	E_{HB}[a]
				$CQH_2^{2+}\cdots H_2PO_4^-$ (I)				
						N9-H9$\cdots$O5	19.3	11.9
4.11	-133.7	-127.4	-140.4	20.0	-7.0	C5-H5$\cdots$O5	1.9	1.5
						C5-H5$\cdots$O6	1.3	1.3
						C11-H11B$\cdots$O5	1.2	1.0
				$CQH_2^{2+}\cdots H_2PO_4^-$ (II)				
4.25	-133.3	-139.1	-158.3	27.1	-7.9	N14-H14$\cdots$O4	33.4	18.4
				$CQH_2^{2+}\cdots H_2PO_4^-$ (III)				
7.66	-108.7	-106.4	-131.4	29.8	-4.8	N1-H1$\cdots$O1	46.9	27.9
				$CQH_2^{2+}\cdots H_2PO_4^-$ (V)				
						C2-H2$\cdots$O6	3.8	3.2
6.30	-83.5	-94.6	-99.1	8.4	-3.8	C3-H3$\cdots$O7	4.2	2.8
				$CQH_2^{2+}\cdots CQH_2^{2+}$				
						C12-H12A$\cdots$Cl	0.5	0.7
5.57	153.5	136.9	136.3	19.9	-19.4	C17-H17B$\cdots$Cl	0.4	0.2
						C10-H10$\cdots$Cl	0.1	0.1
				CQ$\cdots$CQ				
						C12-H12A$\cdots$Cl	0.4	0.7
5.57	9.7	-11.6	-12.4	19.5	-18.7	C17-H17A$\cdots$Cl	0.4	0.2
						C10-H10$\cdots$Cl	0.1	0.1

[a] Hydrogen bond energies, intended as the energies required to break the H$\cdots$acceptor interactions, evaluated within the DFT and ECDA models. E(2) comes from the second-order perturbative estimate of donor-acceptor interactions in the NBO basis. E_{HB} derives from the semiempirical ECDA formalism as the sum of atom-atom H$\cdots$acceptor pairwise attractive terms.

At large centre of mass separations, attractive interactions follow the d^{-1} Coulomb law for the +2/–1 charged chloroquine-phosphate pairs. Specific intermolecular contacts and aspherical features of the electrostatic potential make some DFT interaction energies significantly deviating from the d^{-1} trend at short distances. It is easy to see that this deviation from the d^{-1} trend is especially prominent for pairs (III), (IV) and (VII)–(X). On the contrary, the top-ranking pairs (I) and (II) at d $\approx$ 4 Å are surprisingly close to the E_{LS} curve, despite the extra stabilizing contribution provided by strong CAHBs (Table 5.7). This is likely due to the compensating effect of repulsive interactions which, in tightly bonded pairs, account for the 15-25 % of the total energy and neatly overcome the dispersion term, producing a slight, although significant, stability loss (Table 5.7). On the contrary, pairs (VIII)-(X) are less stable than expected. This evidence might imply some kind of compensating effect operating in the crystal: a large stability gain of some specific pairs could reflect in a bit stability loss of some other pairs. As expected from a thermodynamic viewpoint, what matters as a true structure-determining factor is the overall cohesive energy of the crystal; if some favourable mutual molecular orientations

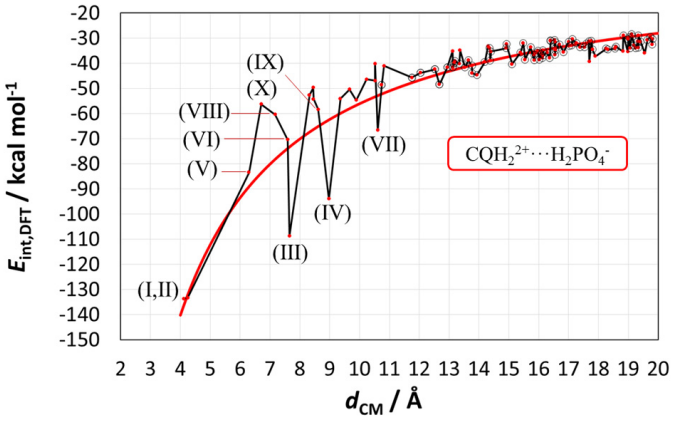


Figure 5.9: B3LYP/pob-TZVP interaction energies ($E_{int,DFT}$) as a function of the centre-of-mass distance (d_{CM}) for $CQH_2^{2+}\cdots H_2PO_4^-$ pairs. The least-square curve has equation $E_{LS}(d)$ = -561(4)/d. The black broken lines serve as a guide for the eyes. Roman numerals indicate some interesting pair in order of increasing energy.

arise, particularly to allow strong CAHBs, this will likely occur at the expense of some other pairs which might be destabilized on a relative basis.

As already shown in Section 5.4, the N-H$\cdots$O CAHBs are somewhat significant as possible structure determinants. Energy outcomes reported in Table 5.7 confirm this result as they account for a consistent part of the total interaction energy. Their relative contributions follow the absolute strength of the interactions. The charged assisted N1–H1$\cdots$O1 and N14–H14$\cdots$O4 interactions provide 43.1 % and 25.1 % of the DFT energy, respectively, while the neutral secondary amine N9–H9$\cdots$O5 HB accounts only 14.4 % of such energy. Interestingly, the strongest hydrogen bond on absolute grounds, N1–H1$\cdots$O1, does not belong to the closest and more stable pairs (I) or (II), but to the (III) one. This occurs frequently in organic crystals without strong HBs.[220] This fact just confirms that there is not, in general, a 1:1 correspondence between very short atom-atom contacts and most stabilizing pairs.

5.5.2 Chloroquine-chloroquine interactions

The last rows in Table 5.7 show the energy contributions associated to the inversion-related π-stacked chloroquine molecules, both for the doubly protonated form found in the crystal and for the neutral CQ placed at the same relative position as the former.

As expected, the total interaction energies are determined by repulsive electrostatic terms. When E_{es} are expressed as multipolar contributions, they are dominated by a large positive monopole (l=0) term contributing $\approx$ 144 kcal mol^{-1}. The promolecular part of the density energy weakly stabilizes the CQH_2^{2+} part by an attractive $\approx$ –10 kcal mol^{-1}contribution. If the l=0 electrostatic term is further decomposed into summations of atomic Mulliken charges of mutually interacting moieties, the relative weights of the various functional groups in determining the gross part of the electrostatic energy can be

appreciated. Numerical values are presented in Table 5.8.

Table 5.8: charge-charge Coulomb contributions to the monopolar l=0 term among different functional groups in charged $CQH_2{}^{2+}\cdots CQH_2{}^{2+}$and neutral $CQ\cdots CQ$ chloroquine pairs. Energies are expressed in kcal mol^{-1}. Percentage values, referred to the absolute value of the total energy, are reported in parentheses.

Pairs	quin-quin[a]	quin-chain[a]	chain-chain[a]	N9H9[a]	total
$CQH_2{}^{2+}$	+55.6 (38.5%)	+99.3 (68.7%)	+47.9 (33.1%)	-58.3 (-40.3%)	+144.5
CQ	-3.1 (-94.0%)	-1.7 (-51.5%)	+2.9 (87.9%)	-1.4 (-42.4%)	-3.3

[a] quin: quinoline, bicyclic condensed aromatic system; chain: atoms belonging to the hydrocarbon chain bonded to N9; N9H9: secondary amine that links quinoline and the lateral chain.

The facing quinoline rings account for +38.5 % (55.6 kcal mol^{-1}) of the total destabilization energy. The contributions of the two long hydrocarbon chains are similar (+33.1 %), while interactions of quinoline and lateral chains with the neutral secondary amine N9H9 lower the energy by –40.3 % (-58.3 kcal mol^{-1}). The largest repulsion (+68.7 %, 99.3 kcal mol^{-1}) arises from cross-interactions between the two charged quinoline rings and the two lateral hydrocarbon chains, as each pair of quinoline-chain functions provides a +34.4 % destabilizing contribution. The same numerical values for the $CQ\cdots CQ$ pair reveal a slightly negative interaction energy (-3.2 kcal mol^{-1}). Stabilization contributions come from the ring-ring and ring-chain interactions, together with the interaction among the neutral secondary amine N9H9 groups. On the contrary, chain-chain interactions slightly destabilize (+2.9 kcal mol^{-1}) the interaction among neutral chloroquine molecules.

Dispersive (E_d) and exchange-repulsion (E_r) terms (Table 5.7) play a secondary role since their numerical contributions are $\approx \pm$ 20 kcal $mol^{-1}$for $CQH_2{}^{2+}\cdots CQH_2{}^{2+}$ pair. Moreover, the two terms exactly cancel each other. This fact leaves the repulsion-dominated electrostatic term as the unique actor on the stage of the interaction energy. Apparently, no positive contribution to the stabilization energy of the pair comes from the dispersion contributions.

$\pi\cdots\pi$ interactions are somewhat elusive as they derive from a mixture of electrostatic[231] and van der Waals[232] terms. The former take into account the effect of the electronic clouds at very large distance through the d^{-1} functional, while the second ones involve time-dependent perturbations of the whole interacting electron density, which can not be properly accounted for by ground-state adiabatic DFT methods.[233] As an approximate estimation of the weight of $\pi\cdots\pi$ interactions in this system, we analyzed the ECDA energy decomposition also for the couple of neutral CQ molecules (last row in Tables 5.7 and 5.8). First, we note that repulsion, mostly due to the closeness of the quinoline cores with the lateral chains, completely counteracts the E_d term also in neutral chloroquine (Table 5.7). Second, a weakly attractive monopole l=0 contribution appears between facing quinoline rings (-3.1 kcal mol^{-1}, Table 5.8), meaning that the mutual arrangement of charges in the facing aromatic systems is slightly favourable. Globally, the $CQ\cdots CQ$ interaction has a negative total energy as high as -3.3 kcal mol^{-1}, due to the sum of

ring-ring, ring-chain and N9H9 contributions (see Table 5.8). The ECDA decomposition also reveals that, similarly to what found for the doubly charged chloroquine pair, the promolecular term contributes ≈ -10 kcal mol^{-1}to the total interaction energy in the pair.

To obtain an estimate of the stabilization energy associated with the stacking motif, we applied the method proposed by Zhikol and co-workers[37] and briefly discussed in Section 1.5.3.

The functional in Equation 1.17 was developed through a least-squares fitting procedure to empirically reproduce the MP2 interaction energies of simple benzene dimers in various stacking geometries. Although chloroquine has a far more complex topology, we believe that the results of this analysis should be trusted at least qualitatively. The results of the application of the Zhikol functional are reported in Table 5.9.

Table 5.9: topological parameters, in atomic units (*a.u.*), of the electron density at the cage critical points CCPs and $E_{\pi\cdots\pi}$ estimates (in kcal mol^{-1}) obtained using the Zhikol functional (Equation 1.17).

System	C	ρ_{CCP}	$\nabla^2\rho_{CCP}$	$E_{\pi\cdots\pi}$
$CQH_2{}^{2+}\cdots CQH_2{}^{2+}$	-0.94555	0.00169	0.01257	4.4(6)
$CQ\cdots CQ$	-0.93716	0.00169	0.01255	4.3(6)

Each ring of the quinoline moiety contributes +2.2(4) kcal mol^{-1}to $E_{\pi\cdots\pi}$. This value is quite close to the Zhikol *et al.*[37] estimate (+2.4 kcal mol^{-1}) for close benzene rings in a similarly staggered conformation. Zhikol functional here suggests that no attractive interactions seem to be present between the two stacked rings. When neutral chloroquine molecules are considered, similar $\pi\cdots\pi$ estimates are found (see last row of Table 5.9).

In conclusion, some attractive contributions, essentially electrostatic in nature, are present in the neutral $CQ\cdots CQ$ π-stacked pair. The latter is mainly due to the promolecular term (≈ –10 kcal mol^{-1}) and to a favorable arrangement of atomic charges. However, the presence of $\pi\cdots\pi$ motif between the rings is difficult to prove without doubts. Some kind of stabilization is detected, although these interactions are so weak to be completely overwhelmed in the charged $CQH_2{}^{2+}$ pair. That said, it is reasonable to believe that in solution $\pi\cdots\pi$ stacking motif should be less significant as structure-driving interaction. Indeed, in a solid-state arrangement, charged molecules are kept in close contact by a mixing of energetic and geometrical factors, while in solution the molecules are free to move and then to minimize their energy moving away.

5.6 A Model for Chloroquine/Heme Interaction

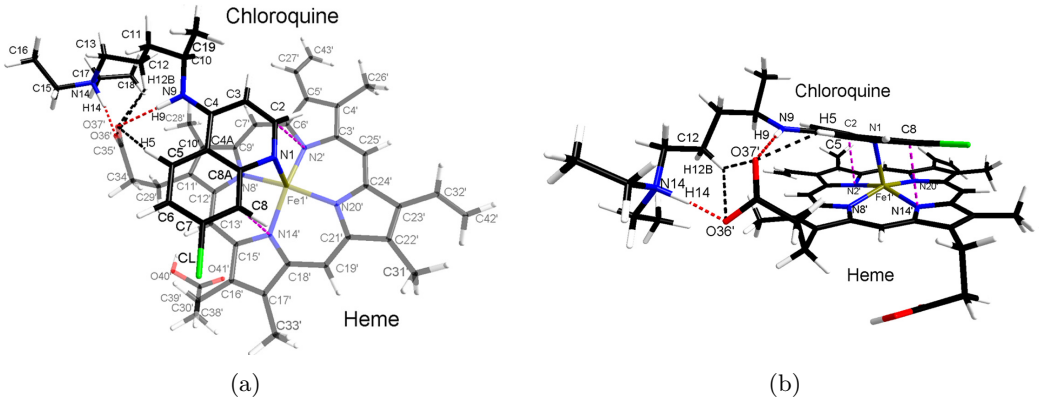
In the acidic digestive vacuole (DV) of the *Plasmodium*, where the pH is ≈ 5, the majority of free heme released from the digestion of hemoglobin is present as monoprotonated neutral hematin, FePPIX(H_2O)H. A water molecule is weakly coordinated to the Fe atom and one propionic acid of the lateral chain of the heme molecule is dissociated (pK_a =

4.3 and 5.5). In a previous work, we have demonstrated that a ligand exchange acid-base reaction is thermodynamically possible in the DV,[187] according to the reaction

$$FePPIX(H_2O)H + CQH_2^{2+} \rightarrow FePPIX(CQH)H^+ + H_3O^+.$$

The FePPIX(CQH)H^+ presents a direct coordinative bond between the N atom of the quinoline ring of CQ and the Fe atom of the heme molecule. The resulting adduct was furthermore stabilized by strong charge-assisted hydrogen bonds between the tertiary protonated amine of the lateral chain of the CQ drug and the propionate group of the protoporphyrin IX scaffold. This structure was compatible with the EXAFS findings in aqueous acidic solutions at room temperature.[187]

The re-optimization step performed in this work led to a structure similar to those found in our previous work with some interesting differences.[187] The graphical representation of the new CQ:heme adduct is shown in Figure 5.10.



(a) (b)

Figure 5.10: Geometry of the CQ:heme adduct optimized at the Grimme corrected B3LYP/pob-TZVP level of theory. Dashes lines represent N-H$\cdots$O CAHBs (red), C-H$\cdots$O (black) and C$\cdots$N (purple).

The main difference between the two adducts resides in the orientation of the quinoline ring with respect to the protophorphyrin scaffold. The dihedral angle in the previous adduct was equal to 78.7°, while in the re-optimized FePPIX(CQH)H^+ the same quantity is now equal to 17.3°. This difference is mainly due to the use of the Grimme dispersion correction which introduces a better description, although empirical, of the van der Waals interactions.

Some similarities emerge between the CQ coordination network in the solid-state crystal and in the CQ:heme adduct. First, from an energetic point of view the interaction energy between the Fe-PPIX moiety and the single-protonated CQH^+ is $\approx$ 106 kcal mol^{-1}. A similar value is found for the $CQH_2{}^{2+}\cdots H_2PO_4{}^-$ pair (III), where the interaction energy is equal to -108.7 kcal mol^{-1}. Second, the strong hydrogen bonds involving the amine groups of the lateral chain, namely N9-H9$\cdots$O and N14-H14$\cdots$O, are conserved (see Tables 5.3, 5.7 and 5.10).

The topological parameters at the BCP of the atomic interaction lines involving the

Table 5.10: Geometrical (in Å), topological (in atomic units *a.u.*) and hydrogen bond energy (in kcal mol^{-1}) properties of intermolecular atom-atom contacts in the chloroquine-heme adduct.

X$\cdots$Y	$d_{X\cdots Y}$	ρ_{BCP}	$\nabla^2\rho_{BCP}$	G_{BCP}	V_{BCP}	H_{BCP}	BD	E(2)
N9-H9$\cdots$O37'	1.869	0.029	0.107	15.8	-14.9	0.9	0.05	4.0
N14-H14$\cdots$O36'	1.429	0.096	0.129	45.0	-69.8	-24.8	-0.41	62.8
C5-H5$\cdots$O37'	2.299	0.015	0.051	6.8	-5.5	1.3	0.14	1.7
C12-H12B$\cdots$O36'	2.522	0.011	0.039	5.2	-4.4	0.8	0.13	0.2
C12-H12B$\cdots$O37'	2.433	0.014	0.049	6.7	-5.8	0.9	0.11	0.2
C2$\cdots$N2'	2.902	0.015	0.055	7.5	-6.5	1.0	0.11	//
C8$\cdots$N14'	2.958	0.013	0.046	6.4	-5.6	0.8	0.10	//
N1$\cdots$Fe	2.100	0.075	0.232	48.2	-60.0	-11.8	-0.25	//

N9 atom are nearly identical in the two systems, suggesting a similarity between the two situations. However, the NBO E(2) estimates reveal that this interaction is weaker in solution than in the crystal by, at least, three times. Results in Table 5.10 show that the energy associated with the N9-H9$\cdots$O5 hydrogen bond is equal to 4.0 kcal mol^{-1}, that is only $\approx$ 4 % of the total interaction energy. For a comparison, the same value in the crystal is predicted to be as high as $\approx$ 19 kcal mol^{-1}(see Table 5.5). This discrepancy reflects the fact that point topological descriptors, such as the bond degree parameter[167], only provide an inherently local picture of the electron density. These properties do neither necessarily nor easily correlate with the interaction energies.[207,234] The analysis of the Mulliken charges of N9-H9$\cdots$O atoms within $CQH_2^{2+}\cdots CQH_2^{2+}$ pair and chloroquine:heme adduct shows that N9 and H9 atoms remain similar, while the negative O charge decreases in magnitude from 0.91 e in the phosphate to 0.67 e in the heme propionate (see Table 5.11). This fact leads, as obvious, to a less favourable Coulomb contribution and then to a lower interaction energy. The energy of the N9-H9$\cdots$O hydrogen bond significantly changes without affecting the local properties of the electron density distribution at the bond critical point.

Table 5.11: Mulliken charges of the atoms involved in the N-H$\cdots$O CAHBs for the pairs extracted from the crystal (pair I and II) and for the chloroquine:heme adduct (FePPIX(CQH))

Atom	$CQH_2^{2+}\cdots H_2PO_4^-$	FePPIX(CQH)
N9	-0.516	-0.530
H9	0.237	0.214
O5/O37'	-0.907	-0.667
N14	-0.706	-0.799
H14	0.405	0.510
O4/O36'	-0.936	-0.751

On the other hand, the N14-H14$\cdots$O CAHB is strongly reinforced in the drug-heme adduct. This is evident looking at the much shorter H$\cdots$O distance ($\approx$ 1.4 Å) and at

the large and negative bond degree parameter in Table 5.10. As already discussed before, negative values of BD parameter indicate an increased degree of covalency for the studied bond. In this CQ:heme adduct, the N14-H14$\cdots$O contact provides a contribution of $\approx$ 63 kcal mol^{-1}, which is almost the 45 % of the global stabilization energy of the adduct.

Also the electrostatic weight becomes more favourable in this interaction. The lower absolute charge of the oxygen propionate in the heme:CQ complex is contrasted by the increased H14 positive charge and the decreased H$\cdots$O distance (Tables 5.10 and 5.11).

The N1-Fe AIL is associated with the coordinative bond with the basic N1 centre of the quinoline ring. It shows a very large negative BD parameter (-0.25 *a.u.*), indicating a relevant degree of covalency in that bond. As a comparison, the strong CAHBs N14-H14$\cdots$O36' has a BD parameter as high as -0.41 *a.u.*.

The close proximity of the quinoline ring to the protoporphyrin system implies the appearance of some atomic interaction lines connecting C atoms of the quinoline ring with the N atoms of the heme (Table 5.10). C2$\cdots$N2' and C8$\cdots$N14' AILs imply ρ_{BCP} estimates at least one order of magnitude higher than in the crystal (compare Tables 5.6 and 5.10). C8A and N8' do not present any atomic interaction line because they are too far from each other. Accordingly, the energy density G_{BCP} and V_{BCP} are from two to three times larger in magnitude, indicating that these interactions are stronger in the complex.

Similarly to what observed in the crystal, two cage critical points are present, located near the protoporphyrine IX aromatic system. Zhikol functional[37] predicts, in contrast with the results found in the crystal, a total stabilizing $E_{\pi\ldots\pi}$ as large as -31(2) kcal mol^{-1}(see Table 5.12).

Table 5.12: topological parameters, in atomic units (*a.u.*), of the electron density at the cage critical points CCPs and $E_{\pi\ldots\pi}$ estimates (in kcal mol^{-1}) obtained using the Zhikol functional (Equation 1.17) for the chloroquine:heme adduct.

System	C	ρ_{CCP}	$\nabla^2\rho_{CCP}$	$E_{\pi\ldots\pi}$
FePPIX(CQH) (1)	0.38058	0.00470	0.02640	-10.9(8)
FePPIX(CQH) (2)	0.55766	0.00526	0.02896	-20(1)
Total	//	//	//	-31(2)

This attractive interaction corresponds to $\approx$ 22.7 % of the total energy predicted by the DFT functional. This result can be only in part explained on the basis of the shorter average distance of the quinoline ring from the least-squares plane of the protoporphyrin (3.1(1) Å, *vs* $\approx$ 3.4 Å for the $CQH_2^{2+}\cdots CQH_2^{2+}$ system). It is worth noting that the two interacting rings are not parallel in the drug-heme adduct. Although this arrangement is not uncommon *per se*,[235] the Zhikol functional was developed from a training set of planar and parallel benzene rings. For this reason, results obtained through functional expressed in Equation 1.17 might be not accurate in such a complex system and the value obtained may be an overestimation of the real interaction energy. In any case, both geometrical and topological parameters agree in pointing out that stacking interactions

between the two cycles are stronger in the adduct than in the $CQH_2^{2+}\cdots CQH_2^{2+}$ solid state pair. However, these $\pi\cdots\pi$ interactions are still weaker than the CAHBs formed by the tertiary amine group of chloroquine, which remain the most important intermolecular contacts also in the adduct.

The reduce density gradient analysis, depicted in Figure 5.11, shows the results already obtained in a more clear way.

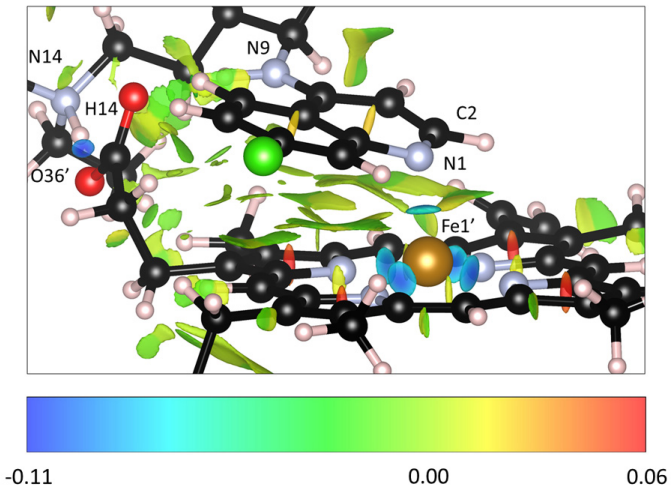


Figure 5.11: $\rho(\boldsymbol{r}) \cdot sign(\lambda_2)$ function plotted on the RDG isosurfaces (isolevel = 0.4) in the CQ:heme adduct. $\rho(\boldsymbol{r}) \cdot sign(\lambda_2)$ values on the chromatic scale are given in *a.u.*.

Five sharp circular-shape isosurfaces with large negative $\rho(\boldsymbol{r}) \cdot sign(\lambda_2)$ values are evident around the Fe atom, four corresponding to the Fe-N coordinative bonds set up by the metal ion with the protoporphyrin scaffold and one associated with the coordinative bond with the quinoline system. A similar feature appears between H14 and O36', confirming that the N14-H14$\cdots$O36' CAHB is extremely strong and bears features comparable to the coordinative N-Fe interactions. On the other hand, the region between the aromatic systems is dominated by unstructured large and flat RDG isosurfaces, bearing slightly negative $\rho(\boldsymbol{r}) \cdot sign(\lambda_2)$ values. Some kind of attractive $\pi\cdots\pi$ interactions seem to occur and are significantly stronger than those established between chloroquine molecules, where the $\rho(\boldsymbol{r}) \cdot sign(\lambda_2)$ function is positive or even close to zero on the RDG isosurfaces (see Figure 5.8).

5.7 Conclusions

In this chapter and in the related work (Reference [175]), a combined experimental and theoretical study was carried out on the chloroquine dihydrogen phosphate dihydrate salt at T = 103(2) K. This investigation allowed us to explore the NCI network, which determines how the drug interacts with its neighbourhood in the solid state. The analysis of the solid state self-recognition is taken here as a model for the drug-substrate interaction

because, likely, the same structure-determining factors are active also in solution.

In literature, $\pi\cdots\pi$ interactions are claimed to be the major players in the drug:target recognition process. In the acidic digestive vacuole of the parasite, where the chloroquine is supposed to be active, the hematin Fe atom is coordinated by a water molecule in the axial position, while the drug is doubly protonated. Although hematin is overall neutral, the metallic centre of the protoporphyrin IX ring remains positively charged, as Fe(III) net charge prevails over the two negative charges carried by the equatorial pyrrole functions. Thus, interactions between the quinoline system of $CQH_2{}^{2+}$ and the FePPIX scaffold should be repulsive in agreement to what observed for the closed-packed $CQH_2{}^{2+}\cdots CQH_2{}^{2+}$ arrangement.

Even though some weak electrostatic attractive interactions might be present, due to local anisotropies in the electron density distribution, they are totally overwhelmed by Coulomb repulsion between the quinoline ring and the Fe(III) ion. Such a repulsive arrangement is even more unlike in solution, where molecules can alleviate electrostatic repulsion by simply keeping moving.

The picture is quite different when a coordinative Fe-N bond is allowed to form. The chloroquine loses its formal charge on the quinoline ring and no net electrostatic repulsion exists. The direct bond between chloroquine and heme keeps the quinoline system very close to the protoporphyrin ring. From topological, structural and energetic point of view, $\pi\cdots\pi$ interactions become stronger in the heme:drug adduct than in the $CQH_2{}^{2+}\cdots CQH_2{}^{2+}$ solid-state pair. Using Zhikol functional, the total interaction energy due to $\pi\cdots\pi$ stacking is estimated to be ≈ 23 %.

Nevertheless, the $CQH_2{}^{2+}\cdots H_2PO_4{}^{-}$ hydrogen bonds involving the amine group of chloroquine are responsible for the stable dihydrogen phosphate salt structure. Both the protonated basic functions in the crystal are involved in strong CAHBs with the negatively charged phosphates, even though there is not a direct correspondence between the strength of the hydrogen bonds and the interaction energies of the corresponding molecular pairs. In the chloroquine:heme complex, hydrogen bonds found in the solid-state are maintained. CAHB between the protonated amine of the drug and the charged propionate function of the protoporphyrin ring accounts for more than 45 % of the total interaction energy, confirming its leading importance in the stabilization of the adduct.

In conclusions, CAHBs are the most important interactions in determining the drug-substrate recognition mode, both in the solid-state and in the acidic chemical environment of the digestive vacuole. The $\pi\cdots\pi$ interactions are intrinsically weaker than the former but they positively co-operate in solution to the stabilization of the chloroquine:heme adduct. The B3LYP/pob-TZVP optimized structure of the adduct is compatible with previous EXAFS findings[187]. Moreover, it might reconcile the two mechanisms of interaction proposed in literature, namely the direct coordinative Fe-N bond and the $\pi\cdots\pi$ motif.

From an engineering point of view, the pharmacophore should have at least three important features. Firstly, a long hydrocarbon chain, able to reach the propionate groups

of the heme molecule, is needed. This chain has to have a strong protonable base, able to form hydrogen bond interaction. Secondly, the drug should have a weak Lewis base on the other side of the molecule, such as the quinoline N atom. This allows the coordination of this base with the metallic centre of the heme. Thirdly, an aromatic system, able to set up $\pi \cdots \pi$ interactions, seems to be helpful for the global stabilization of the drug:substrate adduct.

As a final conclusion, it was here shown how a combined electron density and energetic analysis is a useful approach to shed light on non-covalent interactions features. Both quantities give information on the mode in which the molecules can recognize each other, but from different points of view. The combination of these analyses is the key point to tackle such a complicated topic as the non-covalent interactions.

References

[1] G. N. Lewis, *J. Am. Chem. Soc.* **1916**, *38*, 762.

[2] *Applications of Topological Methods in Molecular Chemistry*, (Eds.: R. Chauvin, C. Lepetit, B. Silvi, E. Alikhani), Springer International Publishing, Switzerland, **2016**, pp. 1–586.

[3] P. L. A. Popelier, *On Quantum Chemical Topology* in *Applications of Topological Methods in Molecular Chemistry, Challenges and Advances in Computational Chemistry and Physics*, *Vol. 22*, (Eds.: R. Chauvin, C. Lepetit, B. Silvi, E. Alikhani), Springer International Publishing, Switzerland, **2016**, pp. 23–52.

[4] *Modern Charge Density Analysis*, (Eds.: C. Gatti, P. Macchi), Springer Dordrecht Heidelberg London New York, **2012**.

[5] G. Saleh, D. Ceresoli, G. Macetti, C. Gatti, *Chemical Bonding Investigation for Materials* in *Computational Materials Discovery*, (Eds.: A. R. Oganov, G. Saleh, A. G. Kvashin), The Royal Society of Chemistry, **2019**, pp. 117–175.

[6] R. F. W. Bader, *Atoms in Molecules: a Quantum Theory*, *Vol. 22*, Oxford University Press, Oxford, **1990**.

[7] P. L. A. Popelier, *Atoms in Molecules: An Introduction*, Prentice Hall, **2000**.

[8] *The Quantum Theory of Atoms in Molecules: from Solid State to DNA and Drug Design*, (Eds.: C. F. Matta, R. J. Boyd), Wiley-VCH, **2007**.

[9] P. L. A. Popelier, F. M. Aicken, *Chem. Phys. Chem* **2003**, *4*, 824–829.

[10] A. D. McNaught, A. Wilkinson., IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Blackwell Scientific Publications, Oxford, **1997**.

[11] P. Hohenberg, W. Kohn, *Phys. Rev. B* **1964**, *136*, 864–871.

[12] T. S. Koritsanszky, P. Coppens, *Chem. Rev.* **2001**, *101*, 1583–1627.

[13] R. F. Stewart, *J. Chem. Phys.* **1973**, *55*, 1668–1676.

[14] N. K. Hansen, P. Coppens, *Acta Cryst. A* **1978**, *34*, 909–921.

[15] A. Volkov, P. Macchi, L. J. Farrugia, C. Gatti, P. Mallinson, T. R. T. koritsanszky, XD2016 Manual, **2016**.

[16] C. Gatti, *Z. Kristallogr.* **2005**, *220*, 399–457.

[17] G. Arfken, *Mathematical Methods for Physicists*, Academic Press, **1985**.

[18] R. F. W. Bader, *J. Phys. Chem. A* **1998**, *102*, 7314–7323.

[19] A. Martín Pendás, E. Francisco, M. A. Blanco, C. Gatti, *Chem. Eur. J.* **2007**, *13*, 9362–9371.

[20] C. R. Wick, T. Clark, *J. Mol. Model* **2018**, *24*, 142 1–9.

[21] C. Gatti, *Phys. Script.* **2013**, *87*, 048102.

[22] R. J. Gillespie, *Molecular Geometry*, Van Nostrand Reinhold, London, **1972**.

[23] C. Gatti, V. R. Saunders, C. Roetti, *J. Chem. Phys.* **1994**, *101*, 10686–10696.

[24] P. Macchi, A. Sironi, *Coord. Chem. Rev.* **2003**, *238-239*, 383–412.

[25] J. C. Slater, *J. Chem. Phys.* **1933**, *1*, 687–691.

[26] R. F. W. Bader, H. Hessén, *J. Chem. Phys.* **1984**, *80*, 1943–1960.

[27] E. Espinosa, I. Alkorta, J. Elguero, E. Molins, *J. Chem. Phys.* **2002**, *117*, 5529–5542.

[28] D. Cremer, E. Kraka, *Croat. Chem. Acta* **1984**, *57*, 1259–1281.

[29] E. R. Johnson, S. Keinan, P. Mori-Sanchez, J. Contreras-García, A. J. Cohen, W. Yang, *J. Am. Chem. Soc.* **2010**, *132*, 6498–6506.

[30] J. Contreras-García, W. Yang, E. R. Johnson, *J. Phys. Chem. A* **2011**, *115*, 12983–12990.

[31] G. Saleh, C. Gatti, L. Lo Presti, J. Contreras-García, *Chem. Eur. J.* **2012**, *18*, 15523–15536.

[32] A. D. Becke, *Exchange-Correlation Approximation in Density-Functional Theory, Modern Electronic Structure Theory* in *Adv. Series in Physcal Chemistry*, *Vol. 2*, (Ed.: D. R. Yarcony), World Scientific, River Edge, NY, **1995**, pp. 1022–1046.

[33] A. Zupan, K. Burke, M. Ernzerhof, J. P. Perdew, *J. Chem. Phys.* **1997**, *106*, 10184–10193.

[34] A. Zupan, J. P. Perdew, K. Burke, M. Causá, *Int. J. Quantum Chem.* **1997**, *61*, 835–845.

[35] G. Saleh, C. Gatti, L. Lo Presti, *Comput. Theor. Chem.* **2012**, *998*, 148–163.

[36] S. Grimme, *J. Comput. Chem.* **2006**, *27*, 1787–1799.

[37] P. A. Zhikol, O. V. Shishkin, K. A. Lyssenko, J. Leszczynski, *J. Chem. Phys.* **2005**, *122*, 144104 1–8.

[38] M. A. Spackman, *J. Chem. Phys.* **1986**, *85*, 6579–6586.

[39] F. London, *Z. Phys. Chem.* **1930**, *11*, 222–251.

[40] J. C. Slater, J. G. Kirkwood, *Phys. Rev.* **1931**, *37*, 682–697.

[41] Q. Wu, W. Yang, *J. Chem. Phys.* **2002**, *116*, 515–524.

[42] A. Volkov, P. Coppens, *J. Comput. Chem.* **2004**, *25*, 921–934.

[43] A. Volkov, X. Li, T. Koritsansky, P. Coppens, *J. Phys. Chem. A* **2004**, *108*, 4283–4300.

[44] M. A. Spackman, *J. Chem. Phys.* **1986**, *85*, 6587–6601.

[45] M. A. Spackman, *J. Phys. Chem.* **1987**, *91*, 3179–3186.

[46] M. A. Spackman, H. P. Weber, B. M. Craven, *J. Am. Chem. Soc.* **1988**, *110*, 775–782.

[47] M. A. Spackman, Keynote Lecture KN23.28, IUCR2005, Book of Abstract, Firenze, Italy.

[48] M. A. Spackman, *Chem. Phys. Lett.* **2006**, *418*, 158–162.

[49] A. D. Buckingham, *Electronic Moments of Molecules* in *Physical Chemistry: An Advanced Treatise*, *Vol. IV: Molecular Properties*, (Ed.: D. Henderson), Academic Press, NewYork, **1970**, Chapter 8, pp. 349–386.

[50] G. Bruno, G. Macetti, L. Lo Presti, C. Gatti, MSc thesis, Università degli Studi di Milano, **2018**.

[51] T. Helgaker, P. Jørgensen, J. Olsen, *Molecular Electronic Structure Theory*, John Wiley & Sons, Chichester, **2000**.

[52] W. Koch, M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, 2nd, Wiley-VCH, Weinheim, **2001**.

[53] C. R. Jacob, M. Reiher, *Int. J. Quantum Chem.* **2012**, *112*, 3661–3684.

[54] M. Reiher, *Chimia* **2009**, *63*, 140–145.

[55] K. Boguslawski, C. R. Jacob, M. Reiher, *J. Chem. Theory Comput.* **2011**, *7*, 2740–2752.

[56] G. K. L. Chan, J. J. Dorando, D. Ghosh, J. Hachmann, E. Neuscamman, H. Wang, T. Yanai, *An Introduction to the Density Matrix Renormalization Group Ansatz in Quantum Chemistry* in *Frontiers in Quantum Systems in Chemistry and Physics*, *Vol. 18*, (Eds.: S. Wilson, P. J. Grout, J. Maruani, G. Delgado-Barrio, P. Piecuch), Springer, Dordrecht, **2008**, pp. 49–65.

[57] K Boguslawski, K. H. Marti, Ö Legeza, M Reiher, *J. Chem. Theory Comput.* **2012**, *8*, 1970–1982.

[58] B. Gillon, P. Becker, *Magnetization Densities in Material Science* in *Modern Charge Density Analysis*, (Eds.: C. Gatti, P. Macchi), Springer Dordrecht Heidelberg London New York, **2012**.

[59] J. M. Gillet, P. J. Becker, P. Cortona, *Phys. Rev. B* **2001**, *63*, 1–7.

[60] J. M. Gillet, *Acta Cryst. A* **2007**, *A63*, 234–238.

[61] J.-M. Gillet, T. Koritsanszky, *Past, Present and Future of Charge Density and Density Matrix Refinements* in *Modern Charge Density Analysis*, (Eds.: C. Gatti, P. Macchi), Springer Dordrecht Heidelberg London New York, **2012**.

[62] R. F. Bader, C. Gatti, *Chem. Phys. Lett.* **1998**, *287*, 233–238.

[63] C. Gatti, *The Source Function Descriptor as a Tool to Extract Chemical Information from Theoretical and Experimental Electron Densities* in *Electron Density and Chemical Bonding, Vol. 2*, (Ed.: D. Stalke), Springer-Verlag Berlin Heidelberg, **2012**, pp. 193–285.

[64] R. F. Stewart, J. Bentley, B. Goodman, *J. Chem. Phys.* **1975**, *63*, 3786–3793.

[65] C. Gatti, G. Saleh, L. Lo Presti, *Acta Cryst. B* **2016**, *72*, 180–193.

[66] C. Gatti, A. Orlando, L. Lo Presti, *Chem. Sci.* **2015**, *6*, 3845–3852.

[67] C. Gatti, A. M. Orlando, E. Monza, L. Lo Presti, *Exploring Chemistry Through the Source Function for the Electron and the Electron Spin Densities* in *Applications of Topological Methods in Molecular Chemistry, Challenges and Advances in Computational Chemistry and Physics, Vol. 22*, (Eds.: R. Chauvin, C. Lepetit, B. Silvi, E. Alikhani), Springer International Publishing, Switzerland, **2016**, pp. 101–129.

[68] C. Gatti, G. Macetti, L. Lo Presti, *Acta Cryst. B* **2017**, *73*, 565–583.

[69] G. Macetti, L. Lo Presti, C. Gatti, *J. Comput. Chem.* **2018**, *39*, 587–603.

[70] C. Gatti, L. Bertini, F. Cargnoni, *J. Comput. Chem.* **2003**, *24*, 422–436.

[71] D. Lasi, E. Ortoleva, C. Gatti, MSc thesis, Università degli Studi di Milano, **2006**.

[72] R. S. Mulliken, *J. Chem. Phys.* **1955**, *23*, 1833–1846.

[73] R. Ponec, F. Uhlik, *J. Mol. Struct.* **1997**, *391*, 159–168.

[74] X. Fradera, M. A. Austen, R. F. W. Bader, *J. Phys. Chem. A* **1999**, *103*, 304–314.

[75] R. Bochiccio, R. Ponec, L. Lain, A. Torre, *J. Phys. Chem. A* **2000**, *104*, 9130–9135.

[76] C. F. Matta, *J. Comput. Chem.* **2014**, *35*, 1165–1198.

[77] C. F. Matta, *Theor. Comput. Chem.* **2018**, *1124*, 1–14.

[78] P. Coppens, *X-ray Charge Densities and Chemical Bonding*, Oxford University Press, **1997**.

[79] P. Coppens, *Phys. Script.* **2015**, *90*, 058001.

[80] P. Coppens, *Angewandte Chemie - International Edition* **2009**, *48*, 4280–4281.

[81] J. A. Duffy, M. J. Cooper, *Spin and the Complementary Worlds of Electron Position and Momentum Densities* in *Modern Charge Density Analysis*, (Eds.: C. Gatti, P. Macchi), Springer Dordrecht Heidelberg London New York, **2012**.

[82] M. Deutsch, B. Gillon, N. Claiser, J. M. Gillet, C. Lecomte, M. Souhassou, *IUCrJ* **2014**, *1*, 194–199.

[83] B. Gillon, C. Sangregorio, A. Caneschi, D. Gatteschi, R. Sessoli, E. Ressouche, Y. Pontillon, *Inorg. Chim. Acta* **2007**, *360*, 3802–3806.

[84] M. Deutsch, N. Claiser, S. Pillet, Y. Chumakov, P. Becker, J. M. Gillet, B. Gillon, C. Lecomte, M. Souhassou, *Acta Cryst. A* **2012**, *68*, 675–686.

[85] P. Macchi, J. M. Gillet, F. Taulelle, J. Campo, N. Claiser, C. Lecomte, *IUCrJ* **2015**, *2*, 441–451.

[86] J. Conradie, A. Ghosh, *J. Phys. Chem. B* **2007**, *111*, 12621–12624.

[87] M. Aebersold, O. Plantevin, L. Pardi, O. Kahn, P. Bergerat, I. V. Seggern, F. Tuczek, L. Öhrström, A. Grand, E Lelievre-Berna, *J. Am. Chem. Soc.* **1998**, *120*, 5238–5245.

[88] C. Aronica, E. Jeanneau, H. El Moll, D. Luneau, B. Gillon, A. Goujon, A. Cousson, M. A. Carvajal, V. Robert, *Chem. Eur. J.* **2007**, *13*, 3666–3674.

[89] M. J. Frisch et al., Gaussian 09, Revision A.01, Gaussian Inc. Wallingford CT, **2009**.

[90] M. J. Frisch et al., Gaussian 16, Revision A.03, Gaussian Inc. Wallingford CT, **2016**.

[91] M. F. Peintinger, D. V. Oliveira, T. Bredow, *J. Comput. Chem.* **2013**, *34*, 451–459.

[92] R. F. W. Bader, AIMPAC, **2017**, `http://www.chemistry.mcmaster.ca/aimpac/`.

[93] F. W. Biegler-König, T. T. Nguyen-Dang, Y. Tal, R. F. W. Bader, A. J. Duke, *J. Phys. B: At. Mol. Phys.* **1981**, *14*, 2739–2751.

[94] F. W. Biegler-könig, R. F. W. Bader, T.-H. Tang, *J. Comput. Chem.* **1982**, *3*, 317–328.

[95] C. Gatti, G. Macetti, "AIMPAC Modified to Evaluate Source Function Contributions", Available on request: carlo.gatti@istm.cnr.it, **2018**.

[96] H. Putz, K. Brandenburg, Diamond - Crystal and Molecular Structure Visualizaton, Crystal Impact: Bonn, Germany, **1997-2012**.

[97] K. Momma, F. Izumi, *J. Appl. Crystallogr.* **2011**, *44*, 1272–1276.

[98] R. Bianchi, C. Gatti, V. Adovasio, M. Nardelli, *Acta Cryst. B* **1996**, *52*, 471–478.

[99] L. Lo Presti, R. Destro, *J. Chem. Phys.* **2008**, *128*, 044710.

[100] S. Elliot, *The Physics and Chemistry of Solids*, John Wiley and Sons, **1998**.

[101] J. Miralles, O. Castell, R. Caballol, J.-P. Malrieu, *Chem. Phys.* **1993**, *172*, 33–43.

[102] J. P. Malrieu, R. Caballol, C. J. Calzado, C. De Graaf, N. Guihéry, *Chem. Rev.* **2014**, *114*, 429–492.

[103] P. Mori-Sánchez, A. J. Cohen, W. Yang, *Phys. Rev. Lett.* **2008**, *100*, 1–4.

[104] F. Neese, *Coord. Chem. Rev.* **2009**, *253*, 526–563.

[105] L. H. Dos Santos, A. Lanza, A. M. Barton, J. Brambleby, W. J. Blackmore, P. A. Goddard, F. Xiao, R. C. Williams, T. Lancaster, F. L. Pratt, S. J. Blundell, J. Singleton, J. L. Manson, P. Macchi, *J. Am. Chem. Soc.* **2016**, *138*, 2280–2291.

[106] C. Gatti, G. Macetti, R. Boyd, C. Matta, *J. Comput. Chem.* **2018**, *39*, 1112–1128.

[107] E. Schrödinger, *What is Life?*, Cambridge University Press: Cambridge, **1944**.

[108] F. Miescher, *Medicinish-Chemische Untersuchungen* **1871**, *4*, 441–460.

[109] O. T. Avery, C. M. MacLeod, M. McCarthy, *J. Exp. Med.* **1944**, *79*, 137–158.

[110] E. Vischer, E. Chargaff, *J. Biol. Chem.* **1948**, *176*, 715–734.

[111] E. Chargaff, E. Vischer, R. Doniger, C. Green, F. Misani, *J. Biol. Chem.* **1949**, *177*, 405–416.

[112] E. Chargaff, *Experientia* **1950**, *6*, 201–209.

[113] R. E. Franklin, R. G. Gosling, *Nature* **1953**, *171*, 740–741.

[114] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737–8.

[115] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 964–967.

[116] F. H. C. Crick, J. D. Watson, *Proc. R. Soc. Lond.* **1954**, *223*, 80–96.

[117] J. D. Watson, *Molecular Biology of the Gene, 2nd ed.* W. A. Benjamin, Inc., New York, **1970**.

[118] J. D. Watson, *The Double Helix: A Personal Account of the Discovery of the Structure of DNA* in, (Ed.: G. S. Stent), W. W. Norton & Co., New York, **1980**.

[119] L. Pauling, R. B. Corey, *Nature* **1953**, *171*, 346.

[120] L. P. ad R. B. Corey, *Proc. Natl. Acad. Sci. USA* **1953**, *39*, 84–97.

[121] R. Parthasarathi, R. Amutha, V. Subramanian, B. U. Nair, T. Ramasami, *J. Phys. Chem.A* **2004**, *108*, 3817–3828.

[122] *Oxford Handbook of Nucleic Acid Structure*, (Ed.: S. Neidle), Oxford University Press, Oxford, **1999**.

[123] J. Šponer, J. Leszczynski, P. Hobza, *J. Biomol. Struct. Dyn.* **1996**, *14*, 117–135.

[124] J. Šponer, I. Beger, N. Spackova, J. Leszczynki, P. Hobza, *J. Biomol. Struct. Dyn.* **2000**, *17*, 1–24.

[125] P. Hobza, J. Šponer, *J. Am. Chem. Soc.* **2002**, *124*, 11802–11808.

[126] P.-O. Löwdin, *Rev. Mod. Phys.* **1963**, *35*, 724–732.

[127] P.-O. Löwdin, *Adv. Quantum Chem.* **1965**, *2*, 213–360.

[128] R. R. Dogonadze, Y. I. Kharkats, J. Ulstrup, *Chem. Phys. Lett.* **1976**, *37*, 360–364.

[129] E. J. Brändas, E. S. Kryachko, *Fundamental World of Quantum Chemistry: A Tribute to the Memory of Per-Olov Löwdin*, Kluwer Academic Publishers, Dordrecht, **2003**.

[130] E. S. Kryachko, J. R. Sabin, *Int. J. Quantum Chem.* **2003**, *91*, 695–710.

[131] L. Gorb, Y. Podolyan, J. Leszczynski, W. Siebrand, A. Fernández-Ramos, *Biopolymers* **2002**, *61*, 77–83.

[132] L. Gorb, Y. Podolyan, P. Dziekonski, W. A. Sokalski, J. Leszczynski, *J. Am. Chem. Soc.* **2004**, *126*, 10119–10129.

[133] Y. Podolyan, L. Gorb, J. Leszczynski, *Int. J. Mol. Sci.* **2003**, *4*, 410–421.

[134] D. N. Stamos in *Quantum Biochemistry: Electronic Structure and Biological Activity*, *Vol. 2*, (Ed.: C. F. Matta), Wiley-VCH, Weinheim, **2010**, Chapter 31, pp. 837–871.

[135] *Quantum Biochemistry: Electronic Structure and Biological Activity*, *Vol. 1 and 2*, (Ed.: C. F. Matta), Wiley-VCH, Weinheim, **2010**.

[136] L.-Y. Fu, G.-Z. Wang, B.-G. Ma, H.-Y. Zhang, *Biochem. Biophys. Res. Commun.* **2011**, *409*, 367–371.

[137] O. O. Brovarets', R. O. Zhurakivsky, D. M. Hovorun, *Chem. Phys. Lett.* **2013**, *578*, 126–132.

[138] O. O. Brovarets', D. M. Hovorun, *Mol. Phys.* **2014**, *112*, 3033–3046.

[139] O. O. Brovarets', D. M. Hovorun, *J. Biomol. Struct. Dyn.* **2015**, *33*, 925–945.

[140] O. O. Brovarets', D. M. Hovorun, *J. Biomol. Struct. Dyn.* **2015**, *33*, 2297–2315.

[141] J. J. Ladik, *Int. J. Quantum Chem.* **1975**, *9*, 133–.

[142] L. Mejía-Mazariegos, J. Hernández-Trujillo, *Chem. Phys. Lett.* **2009**, *482*, 24–29.

[143] A. A. Arabi, C. F. Matta, *Phys. Chem. Chem. Phys.* **2011**, *13*, 13738–13748.

[144] J. P. Cerón-Carrasco, D. Jacquemin, *Phys. Chem. Chem. Phys.* **2013**, *15*, 4548–4553.

[145] J. P. Cerón-Carrasco, D. Jacquemin, *Chem. Commun.* **2013**, *49*, 7578–7580.

[146] J. P. Cerón-Carrasco, J. Cerezo, D. Jacquemin, *Phys. Chem. Chem. Phys.* **2014**, *16*, 8243–8246.

[147] L. P. Candeias, S. Steenken, *J. Am. Chem. Soc.* **1992**, *114*, 699–704.

[148] J. Gu, Y. Xie, H. F. Schaefer, *J. Phys. Chem. B* **2005**, *109*, 13067–13075.

[149] B. Pullman, A. Pullman, *Nature (London)* **1961**, *189*, 725–727.

[150] R. Parthasarathi, V. Subramanian, *Struct. Chem.* **2005**, *16*, 243–255.

[151] R. Parthasarathi, V. Subramanian, *Chem. Phys. Lett.* **2006**, *418*, 530–534.

[152] M. P. Waller, A. Robertazzi, J. A. Platts, D. E. Hibbs, P. A. Williams, *J. Comput. Chem.* **2006**, *27*, 491–.

[153] C. F. Matta, N. Castillo, R. J. Boyd, *J. Phys. Chem. B* **2006**, *110*, 563–578.

[154] C. Gatti, D. Lasi, *Faraday Discuss.* **2007**, *135*, 55–78.

[155] C. Gatti, L. Bertini, *Acta Cryst. A* **2004**, *60*, 438–449.

[156] E. Monza, C. Gatti, L. Lo Presti, E. Ortoleva, *J. Phys. Chem. A* **2011**, *115*, 12864–12878.

[157] F. Meng, C. Liu, W. Xu, *Chem. Phys. Lett.* **2003**, *373*, 72–78.

[158] J Overgaard, B Schiott, F. K. Larsen, B. B. Iversen, B Schiøtt, *Chem. Eur. J.* **2001**, *7*, 3756–3767.

[159] G. Gilli, P. Gilli, *J. Mol. Struct.* **2000**, *552*, 1–15.

[160] P. L. A. Popelier, *Quantum Chemical Topology* in *The Chemical Bond II, 100 Years Old and Getting Stronger*, (Eds.: D. Mingos, P. Michael), Springer International Publishing, Switzerland, **2016**, pp. 71–117.

[161] M. A. Blanco, A. M. Pendás, E. Francisco, *J. Chem. Theor. Comput.* **2005**, *1*, 1096–1109.

[162] A. Martín Pendás, M. A. Blanco, E. Francisco, *J. Chem. Phys.* **2006**, *125*, 1–20.

[163] A. Martin Pendás, E. Francisco, M. A. Blanco, *J. Phys. Chem. A* **2006**, *110*, 12864–12869.

[164] E. Francisco, A. M. Pendás, M. A. Blanco, *J. Chem. Theor. Comput.* **2006**, *2*, 90–102.

[165] A. Asensio, N. Kobko, J. J. Dannenberg, *J. Phys. Chem. A* **2003**, *107*, 6441–6443.

[166] S. F. Boys, F. Bernardi, *Mol. Phys.* **1970**, *19*, 553–566.

[167] E. Espinosa, E. Molins, C. Lecomte, *Chem. Phys. Lett.* **1998**, *285*, 170–173.

[168] P. Abbott, *Mathematica* **2005**, *9*, 689–691.

[169] V. I. Lebedev, D. N. Laikov, *Dokl. Math.* **1999**, *59*, 477–481.

[170] R. H. Hardin, N. J. A. Sloane, W. D. Smith, Tables of spherical codes with icosahedral symmetry, **2018**, `http://NeilSloane.com/icosahedral.codes/`.

[171] G. Voronoi, *Reine Angew. Math.* **1908**, *134*, 198–287.

[172] N. O. Malcolm, P. L. Popelier, *J. Comput. Chem.* **2003**, *24*, 1276–1282.

[173] W. Tang, E. Sanville, G. Henkelman, *J. Phys.: Condens. Matter* **2009**, *21*, 084204 1–7.

[174] M. Yu, D. R. Trinkle, *J. Chem. Phys.* **2011**, *134*, 064111 1–8.

[175] G. Macetti, L. Loconte, S. Rizzato, C. Gatti, L. Lo Presti, *Cryst. Growth Des.* **2016**, *16*, 6043–6054.

[176] A. F. Slater, W. J. Swiggard, B. R. Orton, W. D. Flitter, D. E. Goldberg, A. Cerami, G. B. Henderson, *Proc. Natl. Acad. Sci. U.S.A* **1991**, *88*, 325–329.

[177] D. C. Warhurst, J. C. Steele, I. S. Adagu, J. C. Craig, C. Cullander, *J. Antimic. Chemother.* **2003**, *52*, 188–193.

[178] A. Leed, K. DuBay, L. M. B. Ursos, D. Sears, A. C. de Dios, P. D. Roepe, *Biochemistry* **2002**, *41*, 10245–10255.

[179] A. C. de Dios, R. Tycko, L. M. B. Ursos, P. D. Roepe, *J. Phys. Chem. A* **2003**, *107*, 5821–5825.

[180] J Gildenhuys, T le Roex, T. J. Egan, K. a. de Villiers, *J. Am. Chem. Soc.* **2013**, *135*, 1037–1047.

[181] L. Kořený, M. Oborník, J. Lukěs, *PLoS Pathog.* **2013**, *9*, e1003088 1–6.

[182] S. Kumar, U. Bandyopadhyay, *Toxical. Lett.* **2005**, *157*, 175–188.

[183] M. J. Dascombe, M. G. B. Drew, H. Morris, P. Wilairat, S. Auparakkitanon, W. A. Moule, S. Alizadeh-Shekalgourabi, P. G. Evans, M. Lloyd, A. M. Dyas, P. Carr, F. M. D. Ismail, *J. Med. Chem.* **2005**, *48*, 5423–5436.

[184] M. S. Walczak, K. Lawniczak-Jablonska, A. Wolska, M. Sikora, A. Sienkiewicz, L. Suárez, A. J. Kosar, M.-J. Bellemare, D. S. Bohle, *J. Phys. Chem. B* **2011**, *115*, 4419–4426.

[185] C. Portela, C. M. M. Afonso, M. M. M. Pinto, M. J. Ramos, *FEBS Lett.* **2003**, *547*, 217–222.

[186] M. Asghari-Khiavi, J. Vongsvivut, I. Perepichka, A. Mechler, B. R. Wood, D. McNaughton, D. S. Bohle, *J. Inorg. Biochem.* **2011**, *105*, 1662–1669.

[187] G. Macetti, S. Rizzato, F. Beghi, L. Silvestrini, L. Lo Presti, *Phys. Script.* **2015**, *91*, 023001 1–13.

[188] T. J. Egan, *J. Inorg. Biochem.* **2006**, *100*, 916–926.

[189] E. Guantai, K. Chibale, *Curr. Drug. Delivery* **2010**, *7*, 312–323.

[190] M. Chinappi, A. Via, P. Marcatili, A. Tramontano, *PLoS One* **2010**, *5*, e14064 1–12.

[191] M.-C. Robert, O. Vidal, J.-M. Garda-Ruiz, *Crystallization in Gels and Related Methods* in *Crystallization of Nucleic Acids and Proteins. A Practical Approach*, (Eds.: A. Ducruix, R. Giegé), IRL PRess, Oxford, **1992**, pp. 149–176.

[192] SAINT+, Bruker AXS Inc., Madison, Wisconsin, USA, **2012**.

[193] SADABS, Bruker AXS Inc., Madison, Wisconsin, USA, **2014**.

[194] XPREP, Bruker AXS Inc., Madison, Wisconsin, USA, **2014**.

[195] G. M. Sheldrick, *Acta Cryst. A.* **2015**, *71*, 3–8.

[196] R. Dovesi, R. Orlando, A. Erba, C. M. Zicovich-Wilson, B. Civalleri, S. Casassa, L. Maschio, M. Ferrabone, M. De La Pierre, P. D'Arco, Y. Noël, M. Causà, M. Rérat, B. Kirtman, *Int. J. Quantum Chem.* **2014**, *114*, 1287–1317.

[197] R. Dovesi, V. R. Saunders, C. Roetti, R. Corlando, C. M. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N. M. Harrison, I. J. Bush, P. D'Arco, M. Llunell, M. Causà, Y. Noël, CRYSTAL14 User's Manual, University of Torino, Torino, Italy, **2014**.

[198] C. M. Zicovich-Wilson, A. Bert, C. Roetti, R. Dovesi, V. R. Saunders, *J. Chem. Phys.* **2002**, *116*, 1120–1127.

[199] E. Aprà, M. Causà, M. Prencipe, R. Dovesi, V. R. Saunders, *J. Phys.: Condens. Matter* **1993**, *5*, 2969–2976.

[200] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100.

[201] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B: Conden. Matter Mater. Phys.* **1988**, *37*, 785–789.

[202] H. B. Schlegel, *J. Comput. Chem.* **1982**, *3*, 214–218.

[203] C. Gatti, S. Casassa, TOPOND14 User's Manual, CNR-ISTM of Milano, Milano, Italy, **2014**.

[204] L. Lo Presti, A. Ellern, R. Destro, B. Lunelli, *J. Phys. Chem. A* **2009**, *113*, 3186–3196.

[205] L. Lo Presti, R. S. M. Longhi, E. Ortoleva, *Acta Cryst. B* **2010**, *66*, 527–543.

[206] L. Lo Presti, M. Sist, L. Loconte, A. Pinto, L. Tamborini, C. Gatti, *Cryst. Growth Des.* **2014**, *14*, 5822–5833.

[207] A. Gavezzotti, L. Lo Presti, *Cryst. Growth Des.* **2015**, *15*, 3792–3803.

[208] M. Barzaghi, PAMoC, Online User's Manual, version 2010-07-27, CNR-ISTM, Institute of Molecular Science and Tecnologies, **2010**, `http://www.istm.cnr.it/pamoc`.

[209] E. D. Glendening, A. E. Reed, J. E. Carpenter, F. Weinhold, NBO Version 3.1.

[210] G. Saleh, L. Lo Presti, C. Gatti, D. Ceresoli, *J. Appl. Crystallogr.* **2013**, *46*, 1513–1517.

[211] M. J. Abraham, D. van der Spoel, E. Lindahl, B. Hess, Gromacs development team, GROMACS User Manual, version 5.0.7, **2015**.

[212] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, J. Hermans, *Interaction models for water in relation to protein hydration* in *Intermolecular Forces*, (Ed.: B. Pullman), Reidel Publishing Company, Dordrecht, **1981**, pp. 331–342.

[213] W. L. Jorgensen, J. Tirado-Rives, *J. Am. Chem. Soc.* **1988**, *110*, 1463–1472.

[214] B. Hess, H. Bekker, H. J. C. Berendsen, J. G.E. M. Fraaje, *J. Comput. Chem.* **1997**, *18*, 1463–1472.

[215] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089–10092.

[216] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684–3690.

[217] M. Parrinello, A. Rahman, *J. Appl. Phys.* **1981**, *52*, 7182–7190.

[218] H. S. Preston, J. M. Stewart, *J. Chem. Soc. D* **1970**, 1142–1143.

[219] J. M. Karle, I. L. Karle, *Acta Crystallogr. C* **1988**, *44*, 4–7.

[220] A. Gavezzotti, L. Lo Presti, *Cryst. Growth Des.* **2016**, *16*, 2952–2962.

[221] R. Destro, R. Soave, M. Barzaghi, L. Lo Presti, *Chem. Eur. J.* **2005**, *11*, 4621–4634.

[222] G. Saleh, R. Soave, L. Lo Presti, R. Destro, *Chem. Eur. J* **2013**, *19*, 3490–3503.

[223] P. Politzer, J. S. Murray, Z. Peralta-Inga, *Int. J. Quantum Chem.* **2001**, *85*, 676–684.

[224] F. Fillaux, *Int. Rev. Phys. Chem.* **2000**, *19*, 553–564.

[225] A. Gavezzotti, *CrystEngComm* **2013**, *15*, 4027–4035.

[226] A. Bondi, *J. Phys. Chem.* **1964**, *68*, 441–451.

[227] R. A. Boto, J. Contraras-García, J. Tierny, J.-P. Piquemal, *Mol. Phys.* **2016**, *114*, 1406–1414.

[228] C. D. M. Churchill, L. R. Rutledge, S. D. Wetmore, *Phys. Chem. Chem. Phys.* **2010**, *12*, 14515–14526.

[229] M. Palusiak, B. Bankiewicz, *Comput. Theor. Chem.* **2011**, *966*, 113–119.

[230] Y. A. Abramov, A. Volkov, G. Wu, P. Coppens, *Acta Cryst. A* **2000**, *56*, 585–591.

[231] C. A. Hunter, J. K. M. Sanders, *J. Am. Chem. Soc.* **1990**, *112*, 5525–5534.

[232] A. L. Ringer, M. O. Sinnokrot, R. P. Lively, C. D. Sherrill, *Chem. Eur. J.* **2006**, *12*, 3821–3828.

[233] D. S. Sholl, J. A. Steckel, *Density Functional Theory, A Practical Introduction*, Wiley, New York, **2009**, p. 29.

[234] M. A. Spackman, *Cryst. Growth. Des.* **2015**, *15*, 5624–5628.

[235] F. H. Herbstein, *Crystalline Molecular Complexes and Compounds, Structures and Principles*, *Vol. 1*, Oxford University Press, Oxford, UK, **2005**, p. 113.