An Electron Density Source-Function Study of DNA Bases Pairs in their Neutral and Ionized Ground States†

Carlo Gatti,[(a,b)]* Giovanni Macetti,[(c)] Russell J. Boyd,[(d)] Chérif F. Matta,[(d-g)]**

[(a)] CNR-ISTM, Istituto di Scienze e Tecnologie Molecolari, via Golgi 19, 20133 Milano, Italy. [(b)] Istituto Lombardo Accademia di Scienze e Lettere, via Brera, 28, 20121 Milano, Italy. [(c)] Dipartimento di Chimica, Università degli Studi di Milano, via Golgi 19, 20133 Milano, Italy. [(d)] Dept. of Chemistry, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4J3. [(e)] Dept. of Chemistry and Physics, Mount Saint Vincent University, Halifax, Nova Scotia, Canada B3M 2J6. [(f)] Dept. of Chemistry, Saint Mary's University, Halifax, Nova Scotia, Canada B3H 3C3. [(g)] Dép. de chimie, Université Laval, Québec, Québec, Canada G1V 0A6.

* E-mail: c.gatti@istm.cnr.it; Telephone: (+39) 02 503 14277; Fax: (+39) 02 503 14300.

** E-mail: cherif.matta@msvu.ca; Telephone: (+1) 902 457 6142; Fax: (+1) 902 457 6134.

ABSTRACT

The source function (SF) decomposes the electron density at any point into contributions from all other points in the molecule, complex, or crystal. The SF "illuminates" those regions in a molecule that most contribute to determine the electron density at a given point of reference. When this point of reference is the bond critical point (BCP), a commonly used surrogate of chemical bonding, then the SF analysis broken down at an atomic resolution within the framework of Bader's Quantum Theory of Atoms in Molecules (QTAIM) returns the contributions (positive or negative) of the various atoms in the system to the electron density at that BCP. The most important region that controls a hydrogen bond emerges naturally from this analysis which is the point of this paper as the SF analysis is applied to DNA Watson-Crick dimers (adenine:thymine (AT), and guanine:cytosine (GC)). Both AT and GC dimers are studied in their neutral and their ionized doublet states, radical cationic and radical anionic states, that can result from the interaction with ionizing radiation or with a solvated electron, respectively. The hydrogen bonds in the two dimers are shown to be delocalized to various extents and, surprisingly, the effects of gaining or loosing an electron have similar net effects on some of the hydrogen bonds concealing subtle compensations that are only revealed by examining individual atomic sources

† This paper is dedicated to the memory of Professor Robert G. Parr (1921-2017).

contributions. Coarser levels of resolutions (at the level of the rings and even at the level of monomers-in-dimers) reveal that distant atoms, groups, and rings often have non-negligible effects especially on the weaker hydrogen bonds such as the relatively recently discovered third (weak) hydrogen bond in the AT dimer of the type C-H···O. Interestingly, neither the purine nor the pyrimidine in the neutral or either of the ionized forms dominate any given interactions despite that the former has more atoms that can act as source or sink for the density at a given hydrogen bond.

Keywords: Hydrogen bonding, Bader's quantum theory of atoms in molecules (QTAIM), electron density, adenine-thymine base pair, guanine-cytosine base pair, ionization of DNA.

Introduction

Arguably, few books have had the impact of Erwin Schrödinger's 1944 "*What is Life*".[1] In an extraordinary leap of intuition, Schrödinger realizes that the hereditary mechanism must be mediated by a linear "code-script", which in modern parlance is known as the "genetic code",[2-5] necessary for the transmission and copying of genetic information. The "genetic code" is a fitting description indeed since it is amenable, for example, to a description with Shannon information theory as an abstract language of four letters (adenine (A), guanine (G), thymine (T) or uracil (U), and cytosine (C)) that transcends its physical realization.[6]

While a *complex* of DNA (deoxyribonucleic acid) with proteins was isolated from leukocytes' nuclei in 1869 by Friedrich Miescher (who called this substance "nuclein"),[7,8] Schrödinger did not have the benefit of knowing the chemical nature of the genetic material. That nature was only elucidated by Avery, MacLeod, and McCarty[9] when they established that Griffith's pneumococcal transforming principle[10] was DNA in nature, findings first reported simultaneously with the appearance of Schrödinger's book. (Avery's discovery was not immediately appreciated though and may have been somewhat premature in the mid-forties of last century (see Ref. [11])).

Four years later, Chargaff first reported the rules now carrying his name regarding the ratio and complementarity of nucleic acid.[12-14] By 1952, the chemical composition

of the genetic material was now known and so were Chargaff's rules, however, a crucial ingredient of jigsaw puzzle was still missing: the *structure* of the DNA molecule. Without a structure there could be no more than speculations regarding DNA's stability and the mechanisms of translation, copying, and transmission of the genetic instructions.

Meanwhile, Schrödinger's 1944 book inspired a generation of young scientists to study structural biochemistry in the search of a model of the structure of DNA, the hereditary molecule that carries genetic information. Among those scientists were the then unknown biologist James D. Watson and physicist Francis H. C. Crick.

Using crude (by today's standards) model-building techniques but much ingenuity, Watson and Crick (WC) built the now iconic model for the structure of DNA as a double anti-parallel intertwined α-helices.[15-20] The double helix was proposed by WC on the basis of X-ray diffraction scattering images obtained by Rosalind Franklin.[21] Chargaff's rules guided them in the assignment of complementarity of the DNA bases which, according to their now accepted model, are hydrogen bonded and are nearest to the helical axis in contrast to Pauling and Corey's proposal.[22,23] WC place the phosphate groups at the periphery with respect to the helical axis (the furthest possible from the axis) on the basis of an electrostatic repulsion argument between the negatively charged $-PO_4^{2-}$ groups, again contrary to Pauling and Corey's model.

Watson and Crick's pairing in DNA requires adenine (A) to be bonded by two hydrogen bonds to thymine (T) which can be denoted by $\uparrow A = T \downarrow$ (where the arrows – that can both be reversed in the symbol - indicate the 5'-3' direction of the phosphate-sugar backbone), and guanine (G) to cytosine (C) through three hydrogen bonds which can be denoted by $\uparrow G \equiv C \downarrow$ (where the two arrows can of course also be revered). In 2004, Parthasarathi *et al.* report a third (weak) hydrogen bonding path of the type C-H···O=C linking A and T.[24]

According WC's model, hydrogen bonding in DNA strictly determines specificity and complementarity with an automatic satisfaction of Chargaff's rules as a result. Inter-strand hydrogen bonding interactions and stacking interactions between adjacently stacked base pairs contribute to the stability of the helix structure and hence of the genetic code itself.[25-28]

The asymmetric double well potential of a hydrogen bond in DNA begets a potential

energy barrier that prevents the classical (over the barrier) proton hopping from one base to its complementary base at appreciable rates at physiological temperature. This barrier accounts for the extraordinary stability of the genetic code and the high fidelity of its transmission. Protons, however, are sufficiently small to exhibit some quantum effects such as tunneling. Löwdin was the first to realize that protons can tunnel at a small but finite rate through the barriers of DNA's hydrogen bonds.[29-40]

If quantum mechanical proton tunneling occurs in DNA immediately prior to cell division, it can lead to permanent changes in the genetic code, *i.e.* to a mutation that can be passed on to the daughter cells, which can be harmful to the organism. This mechanism of spontaneous (or induced) genetic error generation is known in the literature as the Löwdin mechanism.[29-40]

Unfavorable charge separation reduces the probability of a single proton tunneling from one base to the opposite side, but a concerted double protons tunneling in opposite directions would preserve charge neutrality of the individual bases and has a higher probability. Such double proton transfer alters the pattern of hydrogen bonding by accompanying the conversion of the bases to their less stable tautomers (denoted by the asterisks A*═T* and G*≡C*).[29,30,32,37,38,40] If double proton hopping occurs during DNA replication, then the less stable tautomer in the template helix will hydrogen bond with the "wrong" base, that is, A*═C, A═C*, G*=T, or G═T*.[41-44] If non immediately lethal, this proton tunneling is an accepted mechanism of ageing through the accumulation of mutational error in the genetic code.[29-32,41-46] The induction of mutation by double proton hopping between the monomers constituting a WC dimer can be accelerated by the agency of externally applied electromagnetic fields.[40-44,47-50]

Nucleic acids (DNA and RNA) are constructed from nitrogenous bases (purines: A, G; and pyrimidines: T, U, C) each attached to a sugar and the sugars in their turn are linked sequentially via phosphate bridges whereby the alternating phosphate and sugar constitute the backbone of the nucleic acid strand. Cartesian analytical philosophical tradition of breaking a problem into smaller ones drove scores of computational chemists to investigate the building blocks of nucleic acids separately. To reduce the size of quantum chemical calculations, the bases, base pairs, triads, etc. - whether Watson-Crick, Hoogsteen, or else - are often studied in isolation from their attached sugar and phosphate moieties after capping the atoms that connect the two bases to

their respective sugar.

That being said, one notes that Cartesian tradition and expediency are *not* the only justifications for the study of isolated bases (or base pairs) rather than complete nucleosides or nucleotides. This approach can be experimentally justified.[51]

Thus, Candeias and Steenken (CS) use 193 nm photons to provoke single-photon ionization-induced photolysis of purine nucleosides and nucleotides in aqueous solution as a model simulating the effect of ionizing radiation on nucleic acids. This wavelength is sufficiently longer than the wavelength that marks the onset of photolysis of water (180 nm) to avoid complications of secondary ionization(s) in the aqueous solutions of these nucleosides and nucleotides. CS found that the (deoxy)ribose sugar, the phosphate (especially the di-anion at pH 9), and the (deoxy)ribose phosphate are all efficiently ionized by single 193 nm photons judging from quantum yields, but the extinction coefficients of the base moieties are two orders of magnitude higher than those of the (deoxy)ribose and phosphate.[51] Similar results were obtained for a model pyrimidine (1,3,5,6-tertramethyluracil) as free and when incorporated into its nucleoside or nucleotide. CS conclude that ionization of nucleosides or nucleotides occur primarily (90%) at the base moiety and neither at the sugar nor at the phosphate group.[51] Furthermore, density functional theory (DFT)[52] calculations by Schaefer III *et al.*[53] demonstrate that the addition of the sugar moiety to the bases have minimal effects on the AT base pairing energies whether in the neutral or anionic form. Thus the pairing energies for deoxyribose-A:deoxyribose-T (dAdT) vs. AT pairs or for their anionic counter parts ($dAdT^-$ vs. AT^-) are equal within 1 kcal/mol at the B3LYP/DZP++ level of theory.[53] This suggests that the hydrogen bonding between the A and T is marginally perturbed by the distant group in the sugar and that it is essentially dominated by proximal and local factors. Given these considerations, the relevance of the present study of *isolated WC base pairs* in their neutral and ionized states to biology (where they form part of nucleotides) has experimental and theoretical justifications.

The five nitrogenous bases (A, G, U, T, C) constitute the letters of the DNA/RNA language in which the genetic code is inscribed. These bases have been the subject of countless theoretical investigations since the dawn of quantum chemistry.[54] Molecular orbital studies of these bases have traditionally been used to estimate the energies,

thermodynamic stabilities of different tautomers, geometries, bond orders and atomic electric charges, dipole moments (and transition dipoles), electronic and vibrational transition energies, ionization potentials and electron affinities.[31,41-44,46] More recently, the topology and topography of the electron density, aromaticity, and conceptual DFT properties have also been investigated.[24,37,55-59]

The present work explores the effect of ionization on the hydrogen bonding interactions in the canonical (WC) base pairs using the tools of Bader's quantum theory of atoms in molecules (QTAIM)[60-62] in conjunction with the analysis of the Bader-Gatti source function.[63-70]

The source function (SF) analysis unravels, for example, the contribution of each individual atom to the density at the DNA hydrogen bonding bond critical point. The hypothesis being tested is that distant atoms can have a considerable influence on the strengths of the individual hydrogen bonds in systems such as the Watson-Crick base pairs. Revealing the atoms with the strongest influence on the hydrogen bonds holding the base pairs can be used to potentially predict substituent effects[71] on this bonding, that is, what would be the effect of replacing an atom or group that could be distant on the hydrogen bonding of the base pair. Further, it is known that the gain or loss of an electron can initiate DNA damage. Hence, another goal of this work is to evaluate the effect of ionization on the hydrogen bonds and their main contributing atoms and groups in the dimers.

Source Function (SF) Analysis

The Source Function (SF) introduces a novel perspective and one rich of chemical insight on how the electron density (ED) of a system originates.[64,68-70] In particular, it shows, within a *cause-effect* view, how the various atoms or groups of atoms in the system contribute to determine the electron density at a given point of the system. If such point is a bond critical point (BCP), assumed as the most representative electron density location for a given chemical interaction, the pattern of the atomic SF contributions provides a visible representation of the more or less delocalized nature of the interaction.[64,68-70]

The SF analysis may be applied to any chemical interaction. Yet, it emerges as of

particular interest for those cases where conventional covalent bonding - characterized by dominant sources from the pair of bonded atoms at BCP - does not apply or where chemical bonding exhibits a large variety of nature, depending on some well identified chemical variables.[64,68,69] This is clearly the case of hydrogen bonds (HBs).

According to the proton-donor and proton-acceptor nature and their linked molecular backbones, hydrogen bonds may range from being short, strong and largely covalent in nature to be long, weak and with electrostatic and/or dispersion forces playing the major energetic role.[72-74] The SF tool has proved to be particularly useful to classify and rank hydrogen bonds,[64,68,69] according to smaller or larger SF locality degree and to the sign and relative weight of sources from the proton-donor, the proton-acceptor and the H atom directly involved in the hydrogen bond (these three atoms will be hereinafter denoted as the hydrogen bond *triad* of atoms).

A later section briefly summaries previous SF results on paradigmatic hydrogen bonded systems[64,68] and shows SF pattern trends for prototypical N-H···N, OH···N and CH···O systems with fixed donor to acceptor distances, in the range of those observed for the hydrogen bonds of the investigated DNA base pairs. These trends will then serve as a model reference for discussing the SF features of the hydrogen bonds binding the DNA bases. Basic aspects of the SF tool are succinctly recalled below.

As shown by Bader & Gatti already 20 years ago,[63] the electron density $\rho(\mathbf{r})$ at a point $\mathbf{r}$ can be described as determined by contributions from a local source LS($\mathbf{r}$,$\mathbf{r}$'), operating at any other point $\mathbf{r}$' in R^3:

$$\rho(\mathbf{r}) = \int \mathrm{LS}(\mathbf{r},\mathbf{r}')\ d\mathbf{r}' \qquad (1)$$

and given by:

$$\mathrm{LS}(\mathbf{r},\mathbf{r}') = -\,(4\pi\cdot|\,\mathbf{r}-\mathbf{r}'\,|)^{-1}\cdot\nabla^2\rho(\mathbf{r}'), \qquad (2)$$

where $(4\pi\cdot|\mathbf{r}-\mathbf{r}'|)^{-1}$ is an *influence* function[75] measuring the effectiveness of $\nabla^2\rho(\mathbf{r}')d\mathbf{r}'$ in determining the effect $\rho(\mathbf{r})$. The concentration or dilution of the electron density, as expressed by its Laplacian at all other points $\mathbf{r}$', is the *cause* for the *effect*, the value of the electron density at $\mathbf{r}$.

By replacing the integration of the local source (LS) over R^3 (Eq. 1) with a sum of disjoint integrations over the QTAIM atomic basins Ω

$$\rho(\mathbf{r}) = \int \mathrm{LS}(\mathbf{r},\mathbf{r}')\ d\mathbf{r}' = \sum_{\Omega} \int_{\Omega} \mathrm{LS}(\mathbf{r},\mathbf{r}')\ d\mathbf{r}' = \sum_{\Omega} \mathrm{SF}(\mathbf{r},\Omega)\,, \qquad (3)$$

the electron density at **r** may be written as a sum of atomic contributions, SF(**r**;Ω), each of them called the *Source Function* from atom Ω to the electron density at **r**.[63]

While the electron density appears as a local property (left hand side of Eq. (3)), in reality it is determined by all parts of the system (the three expressions in Eq. (3)). Whether the influence of these other parts (e.g. atomic basins) is large or small is expressed by their SF values.

From the tenets of DFT one knows that the electron density is uniquely mapped to the external potential which is itself determined from the positions and charges of all nuclei in the system (in addition to any externally-applied fields). The SF tool enables the expression of such non local dependence *in terms of chemistry*, that is, in terms of the influence the various atoms or group of atoms in the system have in determining $\rho(\mathbf{r})$. Furthermore, the SF tool also enables us to quantify how the atoms or group of atoms contribute to determine the local changes of the electron density induced by a given system's perturbation (chemical substitution, geometry change, change of environment, change of the total number of electrons as in the present study, *etc*.)

A final aspect, which is often misunderstood, needs to be clarified. SF atomic contributions (Eq. 3) do not represent the *direct* contribution to the electron density at **r** from their own atom-centered basis set functions, but the capability of the electron distribution within the various atomic basins to determine such electron density value. Said in other words, for a given electron density and **r** value the SF values will be always the same, regardless the way this electron density is expressed (through an atomic basis set or through a multipolar model expansion or given numerically on a grid, *etc*.)

Though not explored here, it is worth mentioning that the SF analysis represents a natural choice for comparing, on the same grounds, results from studies of intermolecular interactions using experimentally and theoretically derived electron densities.[68,70] In fact, SF values are also amenable to experimental determination[68,70,76] provided an accurate $\nabla^2\rho$ distribution derived from high-quality single-crystal X-ray diffraction intensity data is available (Eqs. 1-3).

Source Function representations

Atomic SF contributions are customarily analyzed in terms of Source Function percentage contributions, SF%,

$$\mathrm{SF\%}(\mathbf{r},\Omega)=\frac{\mathrm{SF}(\mathbf{r},\Omega)}{\rho(\mathbf{r})}\times 100 \tag{4}$$

conveying the capability of an atom or group of atoms Ω to determine the electron density at **r,** relative to all other atoms or groups in the system.[64,68,69] The bond critical points (BCPs) are the usually adopted *reference points* (RPs) when electron density reconstructions through SF or SF% are investigated to discuss chemical bonding.[64,68,69] Yet many other choices are possible and in some cases even more convenient.[68,70,76]

To overcome the arbitrariness inherent to any reference points selection, very recently, the interesting concept of SF reconstructed partial electron densities has been introduced.[76] The reconstructed partial electron density is defined:[76]

$$\rho\{\Omega,\ \mathrm{subset}\}(\mathbf{r})=\sum_{\Omega,\mathrm{subset}}\int \mathrm{LS}(\mathbf{r},\mathbf{r'})\,d\mathbf{r'}=\sum_{\Omega,\mathrm{subset}}\mathrm{SF}(\mathbf{r},\Omega)\,. \tag{5}$$

These reconstructed partial electron densities are electron distributions representing that part of the total electron density distribution that are determined by only a given subset of atoms in the system. At the limit, when such subset includes all system's atoms, these distributions become equal to the total electron density if the reconstruction procedure is free of numerical errors. This limit has been verified to be almost the case,[76] especially where $\rho(\mathbf{r})$ is > 0.001 atomic unit (a.u.)

The SF reconstructed partial densities may be conveniently visualized as 2D contour plots. When compared with the corresponding electron density representations, they provide a clear picture of the major or minor role a given subset of atoms plays in determining the electron density in the various system's regions in the plan of the map.[76] *The SF partially reconstructed densities should not be confused with the partial density representations obtained by using only the contributions from the basis functions of a subset of atoms in an* ab initio *computation.* The partial densities represent a rigorous *cause-effect* picture of atomic contributions in partially determining a given electron density. In contrast with *standard* partial densities (that are dependent on the underlying expansion of the electron density in terms of a specific set of basis functions), the SF partial densities represent a rigorous

decomposition of the total electron density that is fully independent of the specific expression adopted to obtain it.

In the analysis of the bonding in neutral and ionized DNA base pairs given below, use will be made both of the conventional ball-and stick SF% representations[64,68-70] and of the SF reconstructed partial densities.[76] Though these latter may be portrayed as such or also as percentage reconstructed densities, only the former option will be considered.

Calculations

Electronic structure calculations

The geometries of the two Watson-Crick base-pair dimers (AT = Adenine-Thymine; GC = Guanine-Cytosine) were fully optimized with no constraints at the density functional theory (DFT)[52] B3LYP/6-31+G(*d*,*p*) level of theory[77,78] followed by the calculation of harmonic vibrational frequencies that revealed no imaginary frequencies. Energies and electron densities were further refined by performing single point calculations at the B3LYP/6-311++G(d,p) level of theory at the geometry optimized with the smaller basis set.

Electron densities and energies of the singly ionized radical species (doublet cations and doublet anions, namely, AT^+, AT^-, GC^+, and GC^-) were obtained using the unrestricted DFT formalism UB3LYP/6-311++G(d,p) after geometry optimization (adiabatic ionization) at the level of theory denoted by UB3LYP/6-311++G(*d*,*p*)//UB3LYP/6-31+G(*d*,*p*). Ball-and-stick representations of the optimized geometries of the neutral and charged base pairs along with the atomic and ring labeling are shown in Fig. 1.

Model hydrogen bonded systems (H_2N-H···NH_3, H_2NH··· OH_2 and H_3CH···OH_2) were evaluated at the same theoretical level of the DNA base pairs and at fixed donor to acceptor distances (H_2N-H···NH_3, 2.8, 3.0 and 3.2 Å; HOH···NH_3, 2.6, 2.8, 3.0 and 3.2 Å; H_3CH···OH_2, 3.6, 3.8, 4.0 and 4.2 Å), covering the range of distances found in the DNA base pairs investigated, at their neutral and charged geometries. All other geometrical parameters were optimized within the C_s symmetry constraint.

All electronic structure calculations and geometry optimizations were performed

using Gaussian 09.[79] Source Function analyses were performed as described in the next sections.

Source Function calculations

Topological analysis of the electron density (ρ) scalar field and the evaluations of percentage atomic SF contributions were performed through a modified version[80] of the AIMPAC program package.[81-83] An in-house code (PLOTDEN2016) has been used to evaluate the SF reconstructed partial electron densities and to plot them. PLOTDEN2016 is an updated and unpublished version of the PLOTDEN2013 code (also unpublished, but with brief description in the Supplementary Information of one previous paper by one of us).[84] The Diamond code[85] was employed to draw all the atomic SF percentages ball-and-stick pictures.

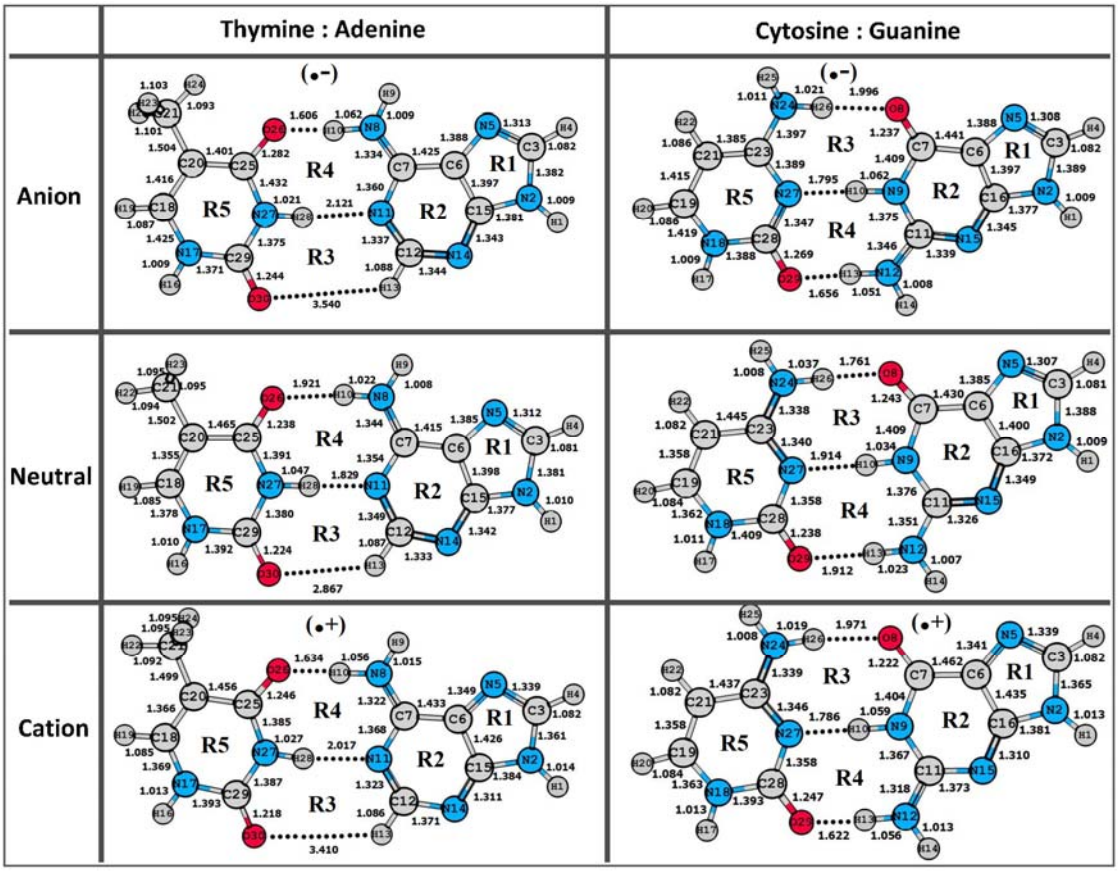

Fig. 1 A display of the optimized geometries (with bond length in Å) along with atomic and rings labeling (rings (R1-R5)) for both neutral and charged AT and GC base pairs. In a nucleoside or nucleotide, the sugar C2' carbon would substitute H16 of thymine (bonded to N17), H1 of adenine (to be bonded to N2), H17 of cytosine (bonded to N18), and H1 of guanine (bonded to N2).

Results and Discussion

Hydrogen bonded systems and their Source Function description

Reference [[68]] provides a detailed summary of previous hydrogen bond studies using the SF. Increasing SF contributions to the hydrogen bond BCP electron density from the hydrogen bond triad of atoms (D-H···A) and a parallel decrease of those from other atoms with increasing energy, covalency and local character of the hydrogen bond were anticipated from such studies. However, other features, which provide additional insights on the nature of the hydrogen bond, were less predictable. In particular, the exceptionally wide variation of the role played by the hydrogen atom directly involved in the hydrogen bond, depending on the hydrogen bond nature,[64,68] or the specific features characterizing the low-barrier hydrogen bonds (LBHB)[86] or, also, the mechanisms controlling the peculiarities of the resonance-assisted hydrogen bonds (RAHBs)[64,68,87]

A systematic study on a number of paradigmatic hydrogen bond systems was carried out by Gatti *et al*.[64] The reaction path of two water molecules approaching each other within the linear dimer C_s constraint serve, in that study, as a model to analyse the evolution of the SF atomic contributions to the electron density at the hydrogen bond BCP when the H···O and O···O distances change from the values typical of the weak isolated hydrogen bonds (IHBs) to those occurring in the charge assisted H-bonds, CAHBs.[87]

Though the total source function contributions from the donor and the acceptor water molecules remain almost constant and comparable along the whole reaction path, the individual SF% contributions were found to change dramatically. In particular, the SF% from the hydrogen atom involved in the hydrogen bond represents a clear marker of nature of the hydrogen bond. Its value is largely negative at equilibrium distance (O···O = 3.020 Å), then keeps diminishing in absolute value with decreasing donor to acceptor distance and eventually becomes slightly positive for very short O···O and H···O distances.

For a standard covalent bond, the sum of the SF% contributions from the two bonded atoms to the electron density at their intervening BCP is usually around 80-90%. Conversely, the cumulative SF% contribution from the hydrogen atom and its

acceptor oxygen atom to the electron density of the hydrogen bond BCP was found to be (significantly) negative at equilibrium distance, almost zero at distances (O···O = 2.750 Å) typical of the long chains of O-H···O bonds in water and alcohols dominated by σ-bond cooperativity,[88] and finally becomes positive, yet not larger than 50%, at the O···O = 2.250 Å – a distance typical of CAHBs.[87]

To reach values comparable to those of a typical covalent bond, one has to include in the SF% sum also the contribution from the donor oxygen atom and to consider donor to acceptor distances as low as 2.5 Å. Thus, even at such low distances, the hydrogen bond retains at least a three-center nature,[64,68,73,87] as also corroborated recently by Popelier *et al*.[89] using Interacting Quantum Atoms (IQA) theory[90-94] and QTAIM delocalization indexes.[95]

At equilibrium, the sources from atoms other than the hydrogen bond triad of atoms yield almost half of the electron density at the hydrogen bond BCP. This fairly delocalized picture of sources has been related[64,68] to the dominant electrostatic nature of the hydrogen bond at equilibrium distance, to be contrasted with the partial covalent character of the bond when the SF%(H+A+D) contribution becomes dominant and the atomic sources become much more localized, as is for CAHBs.

Evolution of atomic sources with hydrogen bond nature change can be depicted as a ball and stick representation of sources with the balls' volumes proportional to the SF% contributions.[64] Sources are visibly seen as spread almost over the entire hydrogen bonded dimer at equilibrium distance and progressively more localized on the D-H···A triad of atoms by decreasing O···O distances.[64] The dramatic change along the reaction path of the SF% contribution from the hydrogen atom involved in the hydrogen bond, and in particular the occurrence of negative and even very large values for such contribution might be surprising. Yet, all this was shown to have a profound chemical meaning.[64,68]

For an isolated atom Ω, SF(**r**,Ω) and SF%(**r**,Ω) is always >0, because only Ω determines the electron density for any **r,** but for an atom in a polyatomic system the situation may be different. Usually local sources are found, on average, to dominate the local sinks, leading to positive SF and SF% values for BCP densities reconstruction. However, when the Laplacian distribution of atom in combination is deeply distorted from the spherical symmetry, the reverse may also occur, depending

on the location of the reference point. Indeed, the hydrogen atom involved in weak to moderate OH···O hydrogen bonds is notably asymmetric in shape and has a quite inhomogeneous $\nabla^2\rho$ distribution, with a large area of positive $\nabla^2\rho$ close to and surrounding the hydrogen bond BCP and, conversely, a localized and highly negative $\nabla^2\rho$ region along the O-H bond. It is this highly asymmetric $\nabla^2\rho$ distribution which leads to negative SF(**r**,H) and SF%(**r**,H) contributions for the electron density at the hydrogen bond BCP and to the usual largely positive contributions to the electron density at the O-H bond BCP . However, when the O···O distance decreases, the O-H and H···O distances approach each other in value and both the shape of the hydrogen basin and its $\nabla^2\rho$ distribution become more symmetric along the two bonding directions, eventually leading to positive sources from the hydrogen atom to both the O-H and the hydrogen bond BCP electron densities.

The evolution of the SF% contribution from the hydrogen atom to its hydrogen bond BCP recap the complex change of shape, size and $\nabla^2\rho$ distribution within the basin of the hydrogen atom involved in the hydrogen bond related to the progressive equalization of the O-H and H···O bonds and to the increasing partial covalent character of the H···O bond with decreasing O···O distance.

The source function patterns recovered for the model of the two approaching water molecules, were also compared to those obtained on a series of prototypical OH···O hydrogen bond complexes,[64] including positively (+CAHB) and negatively (-CAHB) charge-assisted hydrogen bonds, RAHBs, polarization-assisted hydrogen bonds (PAHB) and isolated hydrogen bonds, according to Gilli & Gilli's classification.[87] The comparison revealed similar qualitative patterns, as a function of the O···O distance, but with appreciably enhanced variations of the SF%(H) value relative to the water dimer model.

Additional subtle features were observed for the RAHB systems. They are characterized by a significantly decreased SF%(H+D+A) contribution from the hydrogen bond triad of atoms at the hydrogen bond BCP electron density relative to that expected at a similar O···O separation distance in the water dimer model. The deviation from the expected trend highlights the ability of the SF tool to reflect the electron conjugation mechanisms operative in the RAHBs. Two recent studies have confirmed the ability of the SF to reflect electron delocalization effects, either using

theoretically[67,70] or experimentally[70] derived electron densities.

Gatti *et al*.[64] also studied a few hydrogen bond systems where either the donor (D) or the acceptor (A) atom, or both of them are not oxygen atoms. Though exhibiting trends of SF% contributions qualitatively similar to those observed for OH···O hydrogen bond complexes, as a function of the donor to acceptor distance, the relative weights of such contributions also show a marked dependence on the specific nature of the proton-donor and proton-acceptor atoms. Table 1 lists the SF% values for the electron density at the hydrogen bond BCP for three model hydrogen bond systems, (H_2N-H···NH_3, H_2NH··· OH_2 and H_3CH···OH_2) having the same D and A atom pairs as in the studied DNA base pairs and spanning the same range of donor to acceptor atom distances. Fig. 2 illustrates ball and stick representations of these SF% contributions.

Table 1 Source Function percentage (SF%) contributions to the electron density at the hydrogen bond BCP as a function of the donor to acceptor atom distance in the H_2N-H···NH_3, H_2NH··· OH_2 and H_3CH···OH_2 model complexes [(a,b)]

$R_{D\cdots A}$	$R_{D\text{-}H}$	$R_{H\cdots A}$	ρ_{BCP}	SF%(H)	SF%(H+A)	SF%(H+D+A)	$SF\%_{mol(D)}$
				H_2N-H···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	39.6	46.8
2.8	1.010	1.790	0.0429	-5.9	17.5	57.0	46.6
				H_2N-H···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-H···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) Distances in Å; ρ_b in atomic units (a.u.)

(b) H, D and A are, respectively, the hydrogen atom directly involved in the hydrogen bond, the proton-donor and the proton-acceptor atoms while $SF\%_{mol(D)}$ is the percentage source contribution from the proton-donor molecule in the complex.

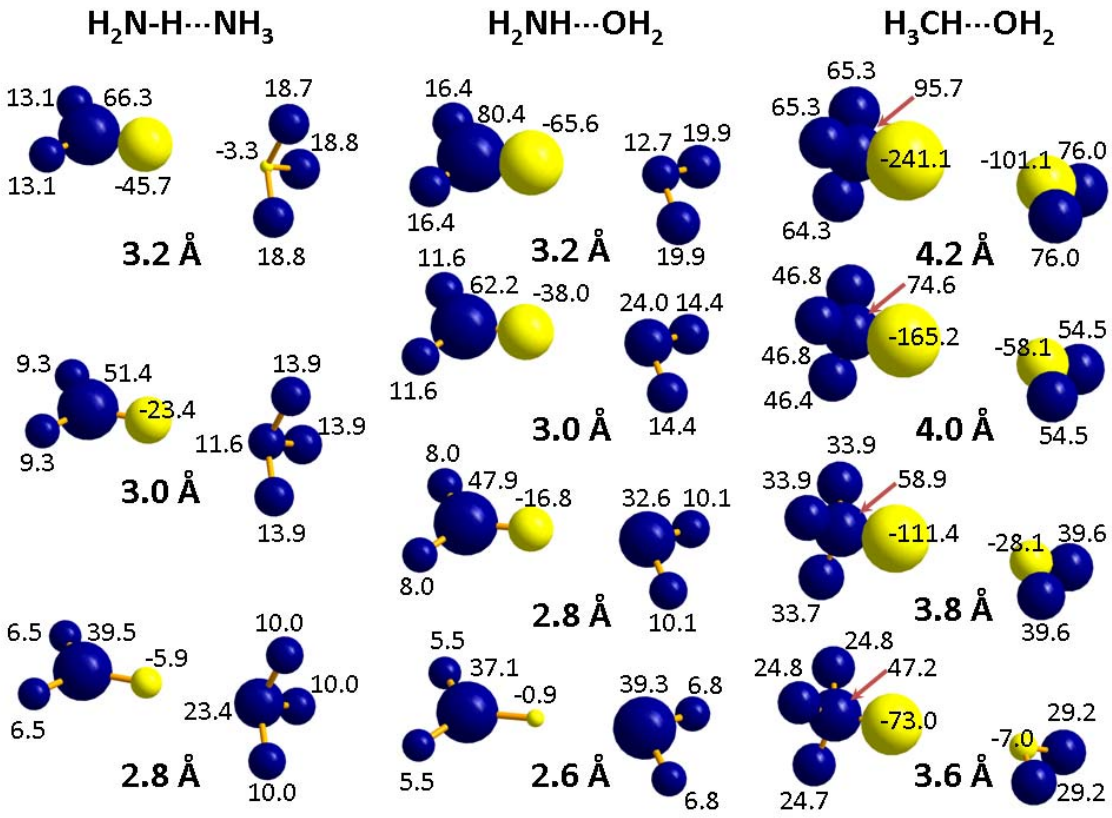

Fig. 2 Source function percentage atomic contributions to the electron density at the hydrogen bond BCP as a function of the donor to acceptor atom distance in the H_2N-H···NH_3, H_2NH··· OH_2 and H_3CH···OH_2 model complexes. Atoms are displayed as spheres whose volume is proportional to the SF percentage contribution from Ω to the electron density at BCP. Positive (sources) in blue/dark and negative (sinks) in yellow/light.

Data in Table 1 confirm that the global SF% contributions from the donor and the acceptor molecules remain almost constant and comparable for all model complexes and at all studied D···A distances, whereas the SF% contributions from selected group of atoms (H; H+A; H+D+A) change dramatically with the D···A distance and significantly as a function of the D and A kind of atoms.

The SF% from the hydrogen atom involved in the hydrogen bond is confirmed as a clear marker of the nature of the hydrogen bond. In the range of investigated D···A distances, it is everywhere negative and its magnitude decreases with decreasing D···A distance, approaching almost zero for the strongest hydrogen bond (NH···O at the shortest N···O distance of 2.6 Å). For the very weak CH···O hydrogen bond, the SF%(H) is largely negative at all examined C···O distances and only at the shortest one (3.6 Å) the global SF% contribution from the hydrogen bond triad of atoms becomes

comparable (42%) to that from all other atoms in the complex. Only for the NH···O hydrogen bond and at the shortest distance considered, is the SF% contribution from the hydrogen bond triad almost approaching a source percentage value similar to that from the two bonded atoms of a conventional covalent bond. Ball and stick representations in Fig. 2 summarize the SF% patterns differences in the three examined model complexes. They serve as a visual reference to interpret SF% data for hydrogen bonds in the DNA base pairs, the object of this study.

Note that for the weaker CH···O hydrogen bonds, besides the source function percentages from the hydrogen atom involved in the hydrogen bond, also those of the oxygen acceptor atom are always largely negative, approaching zero only at the shortest C···O distance. Fig. 2 also shows that even at the shortest analysed distance, all complexes exhibit a fairly delocalized pattern of sources. The delocalized pattern increases with decreasing hydrogen bond strength (NH···O < NH···N < CH···O).

DNA base pairs

It is usual to consider the Watson and Crick DNA base pairs as being bound by two hydrogen bonds in AT and by three in GC. As a result, the AT dimer has O30 (Fig. 1 numbering) of thymine base as a totally available hydrogen bond acceptor on the side of the minor-groove, a fact noted by Crick as a possible source of DNA structural instability.[96] Geometric arguments led Leonard *et al*. to predict a third hydrogen bond in the AT pair with an approximate stabilization energy of 1 kcal/mol.[97] In 2004, and on the basis of a study of the topography and topology of the electron density, Parthasarathi *et al.* discovered the third hydrogen bonding bond path in the AT dimer, with the triad O30-H13-C12 (according to the numbering of Fig. 1).[24] Hence, there exist three hydrogen bond paths in each of the two WC dimers.

The energetic stability imparted by individual hydrogen bonds in the Watson-Crick dimers has been estimated with MP2/D95(d,p) level after counterpoise correction of the basis set superposition error (BSSE)[98] by Asensio *et al*.[99] The study of these workers included energetic estimates of the third hydrogen bond in the AT dimer. The energy of a given hydrogen bond was equated to the total binding energy of the dimer after constraining the two monomers to adopt perpendicular planes along the axis of the hydrogen bond of interest.[99] The values obtained by Asensio *et al*. generally agree

with the values obtained by Matta *et al*.[58] who use Espinosa, Molins, and Lecomte's approximation[100] to estimate the energetic contribution of a given hydrogen bond. The two literature estimates, just mentioned, agree in that the weak hydrogen bond in AT contributes indeed *ca.* 1 kcal/mol in stability as formerly predicted.

The total dimerization energy of the AT dimer is approximately 13/15 kcal/mol (where the first estimate is of Matta *et al*.[58] and the second that of Asensio *et al.*,[99] the same convention that will be used in the following discussion when quoting energies of individual hydrogen bonds in the dimers).

The three hydrogen bonds that link the two base pairs, AT and GC, give rise to two rings (R3 and R4) in each WC dimers. In both dimers, the central N-H···N hydrogen bond is shared by rings R3 and R4 in both systems (Fig. 1). In AT this hydrogen bond is estimated to contribute ca. 6 to 5 kcal/mol to the system's dimerization energy. The other strong hydrogen bond in the AT dimer, of the N-H···O type, contributes *ca*. 6/6 kcal/mol to the dimerization energy bond. Finally, the weak relatively recently discovered hydrogen bond of the type C-H···O contributes 1/1 kcal/mol. Whereas in the GC pair, the N-H···N hydrogen bond contributes ca. 9/9 kcal/mol, while the other two (both of the N-H···O type, but antiparallel) contribute 11/5 kcal/mol (N12-H13···O29) and 6/13 kcal/mol (N24-H25···O8), respectively.

Tables 3 and 4 summarize the principal geometric changes and changes in the SF contributions to the individual hydrogen bonds upon ionization. Upon uptake or removal of an electron, the central N-H···N hydrogen bond weakens in the AT pair ($R_{D\cdots A}$ is 2.877 Å in AT and elongates up to 3.036 in AT^+ and even to 3.136 Å in AT^-), while it is similarly reinforced in the GC pair ($R_{N\cdots N}$ is 2.947 Å in GC and shortens to 2.844 and 2.857 Å in GC^+ and GC^-, respectively). Regardless of the sign of the charged complex, elongation of the central N-H···N hydrogen bond in AT is accompanied by a significant strengthening of the N-H···O bond, with a $R_{N\cdots O}$ decrease of almost 0.3 Å and by the rupture of the weak C-H···O bond (a BCP is lacking in AT^+ and AT^-) due to the very large $R_{C\cdots O}$ increase (GC: 3.689 Å; GC^+ and GC^-, 4.129 and 4.249 Å, respectively). In the GC base pair, uptake or removal of an electron leads to a significant strengthening of one N-H···O hydrogen bond, similarly to the AT complex, and a clear weakening, though not a rupture, of the other N-H···O HB.

The SF% contributions at the hydrogen bond densities at the BCPs portrayed in Fig. 3 (AT and charged forms) and Fig. 4 (GC and charged forms) illustrate such changes. The SF% contributions from the hydrogen atom involved in the hydrogen bond are negative for all hydrogen bonds in the AT and GC neutral complexes and range from -137.8% for the very weak C-H···O bond in AT down to only -3.4% and -7.1% for the strongest N27-H28···N11 and N24-H26···O8 hydrogen bonds in AT and GC, respectively.

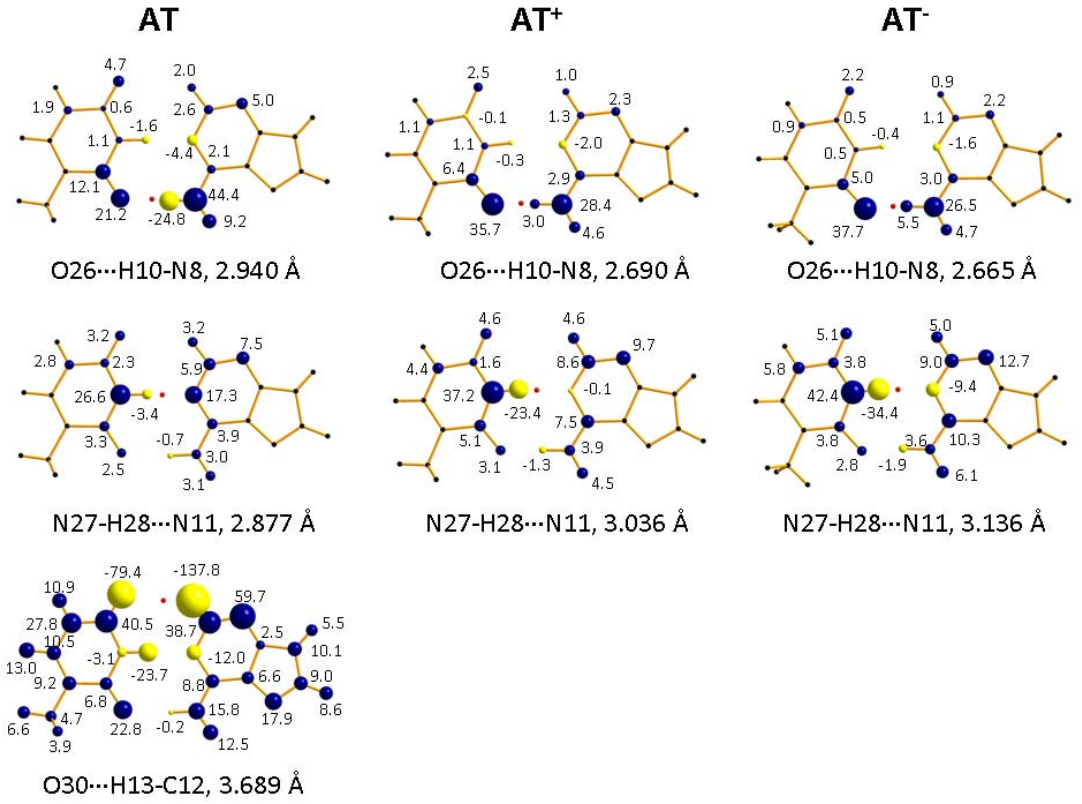

Fig. 3 Source function percentage atomic contributions to the electron density at the hydrogen bonds BCPs (denoted as small (red) dots when they act as reference point for the SF electron density reconstruction) in the AT base pair (neutral and charged forms). Atoms are displayed as spheres whose volume is proportional to the SF percentage contribution from Ω to the electron density at BCP. Positive (sources) in blue/dark and negative (sinks) in yellow/light. (The orientation of the dimer in this figure is rotated 180° along the horizontal central axis with respect to the orientation in Fig. 1).

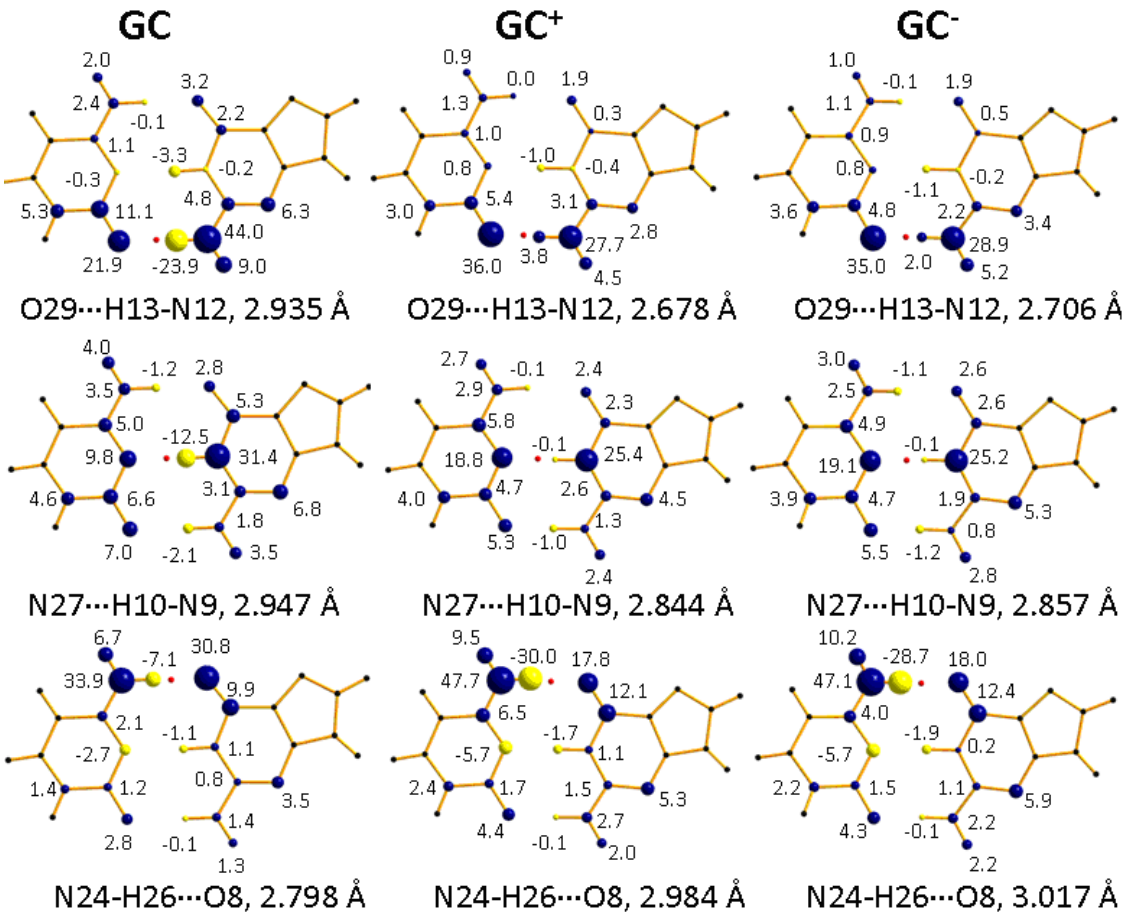

Fig. 4 Source function percentage atomic contributions to the electron density at the hydrogen bonds BCP s (denoted as small (red) dots when they act as reference point for the SF electron density reconstruction) in the GC base pair (neutral and charged forms). Atoms are displayed as spheres whose volume is proportional to the SF percentage contribution from Ω to the electron density at BCP. Positive (sources) in blue/dark and negative (sinks) in yellow/light. (The orientation of the dimer in this figure is the same as in Fig. 1).

Upon uptake or removal of an electron, hydrogen bond weakening implies, as expected, more negative SF%(H) contributions, while hydrogen bond strengthening leads to almost zero or even positive SF%(H) contributions. For instance, lengthening of the central N27-H28···N11 hydrogen bond in AT decreases SF%(H) from -3.4% in the neutral form to -23.4% and -34.4% in AT^+ and AT^- , while shortening of the central N-9-H10···N27 hydrogen bond in GC increases SF%(H) from -12.5% in the uncharged complex to -0.1% in both GC+ and GC^-. Weakening or strengthening of the hydrogen bonds, upon electron attachment or detachment, is also accompanied by, respectively, an increase or a decrease of the spread of sources (Figs. 3 and 4). This is quantitatively summarized by the SF%(H+D+A) values in Tables 2 and 3.

Table 2 Adenine:Thymine (AT) Source Function percentage (SF%) contributions to the electron density at the hydrogen bond BCPs in the neutral and charged base pairs at their fully optimized geometries[a,b]

Bond/system	$R_{D\cdots A}$	$R_{D\text{-}H}$	$R_{H\cdots A}$	ρ_{BCP}	SF%(H)	SF%(H+A)	SF%(H+D+A)	$SF\%_{Ade}$
N8-H10···O26								
AT	2.940	1.022	1.921	0.0262	-24.8	-3.6	40.8	47.3
AT^+	2.690	1.056	1.634	0.0519	3.0	38.7	67.1	46.3
AT^-	2.665	1.062	1.606	0.0564	5.5	43.2	69.7	46.4
N27-H28···N11								
AT	2.877	1.047	1.830	0.0406	-3.4	13.9	40.5	53.6
AT^+	3.036	1.027	2.017	0.0257	-23.4	-23.5	13.7	51.8
AT^-	3.136	1.021	2.121	0.0209	-34.4	-43.8	-1.4	52.6
C12-H13···O30								
AT	3.689	1.087	2.868	0.0043	-137.8	-217.2	-178.5	47.1
AT^+	4.129	1.086	3.410	No BCP				
AT^-	4.249	1.088	3.540	No BCP				

(a) Distances in Å; ρ_b in atomic units (a.u.)

(b) H, D and A are, respectively, the hydrogen atom directly involved in the hydrogen bond, the proton-donor and the proton-acceptor atoms while $SF\%_{Ade}$ is the percentage source contribution from the adenine monomer in the dimer.

Table 3 Guanine:Cytosine (GC) Source Function percentage (SF%) contributions to the electron density at the hydrogen bond BCP s in the neutral and charged base pairs at their fully optimized geometries[a,b]

Bond/system	$R_{D\cdots A}$	$R_{D\text{-}H}$	$R_{H\cdots A}$	ρ_{BCP}	SF%(H)	SF%(H+A)	SF%(H+D+A)	$SF\%_{Gua}$
N12-H13···O29								
GC	2.935	1.022	1.913	0.0268	-23.9	-2.0	42.0	46.2
GC^+	2.678	1.056	1.622	0.0538	3.8	39.8	67.5	46.7
GC^-	2.706	1.051	1.656	0.0511	2.0	37.0	65.9	46.8
N9-H10···N27								
GC	2.947	1.034	1.914	0.0330	-12.5	-2.7	28.7	45.7
GC^+	2.844	1.059	1.786	0.0443	-0.1	18.7	44.1	46.6
GC^-	2.857	1.062	1.795	0.0438	-0.1	19.0	44.2	46.6
N24-H26···O8								
GC	2.798	1.038	1.761	0.0387	-7.1	23.7	57.6	52.5
GC^+	2.984	1.019	1.971	0.0235	-30.0	-12.2	35.5	51.8
GC^-	3.017	1.021	1.996	0.0236	-28.7	-10.7	36.4	52.3

(a) Distances in Å; ρ_b in atomic units (a.u.)

(b) H, D and A are, respectively, the hydrogen atom directly involved in the hydrogen bond, the proton-donor and the proton-acceptor atoms while $SF\%_{Gua}$ is the percentage source contribution from the guanine monomer in the dimer.

For the central N27-H28···N11 bond in AT, SF%(H+D+A) decreases from the already low value of 40.5% in AT down to 13.7% and even -1.4% in AT^+ and AT^-, respectively, while for the central N-9-H10···N27 hydrogen bond in GC, SF%(H+D+A) increases from 28.7% in the neutral form up to more than 44% in the charged forms.

Hydrogen bonds in AT and GC complexes exhibit, in general, fairly delocalized sources. Even for the strongest hydrogen bonds, SF%(H+D+A) does not reach 70%. For the weakest N-H ···N hydrogen bond SF%(H+D+A) is close to zero and equal to -1.4% (AT^- , N27-H28···N11) and for the weakest hydrogen bond, C12-H30···O30, it is largely negative (-178.5%).

A question that arises in this context is that of which of the bases in each complex is giving the larger SF% contribution at a given HB. Said differently, is the purine or the pyrimidine base the principle responsible for the glue that binds the dimer?

Data in Tables 2 and 3, where the global SF% contribution from the atoms of the adenine and the guanine moieties are reported ($SF\%_{Ade}$ and $SF\%_{Gua}$) show that the contributions form the pair of bases in each complex are similar, the dominating one never exceeding 54%. Hence, there is no *clear* dominant monomer type that can be termed the principle responsible of the binding of the dimer. However, and similar to the case of the H_2N-H···NH_3, H_2NH···OH_2 and H_3CH···OH_2 in the model complexes, it is *the donor moiety of the complex that appears to always yield a slightly lower contribution* (slightly below 50%).

The above general conclusions apply to both the neutral and the charged forms of the base pairs. For instance, in the neutral form of AT, $SF\%_{Ade}$ is 53.6% for the central N27-H28···N1 hydrogen bond, where the adenine acts as an acceptor moiety, while $SF\%_{Ade}$ is about 47% for the other two hydrogen bonds where the adenine acts as an hydrogen donor. AT^+ and AT^- exhibit a similar behavior (See Table 2). Turning now to individual rings-in-molecules (RIMs) contributions, Tables 4 and 5 list the SF% contributions from the atoms of rings R1-R5 (defined in Fig. 1) to the electron densities at hydrogen bonds' BCPs in both complexes.

Table 4 Adenine:Thymine (AT) Source Function percentage (SF%) ring contributions to the electron density at the hydrogen bond BCPs in the neutral and charged base pairs at their fully optimized geometries[a,b]

Bond/system	$R_{D\cdots A}$	SF%(H+D+A)	SF%(R1)	SF%(R2)	SF%(R3)	SF%(R4)	SF%(R5)
N8-H10···O26							
AT	2.940	40.8	6.4	10.0	4.8	**53.1**	21.4
AT^+	2.690	67.1	3.3	5.7	3.6[c]	**75.2**	11.4
AT^-	2.665	69.7	2.9	5.7	3.2[c]	**76.1**	10.1
N27-H28···N11							
AT	2.877	40.5	7.4	37.8	**53.9**	**54.6**	38.0
AT^+	3.036	13.7	11.4	28.6	**33.0**[c]	**32.1**	53.2
AT^-	3.136	-1.4	13.3	26.3	**21.5**[c]	**17.3**	62.2
C12-H13···O30							
AT	3.689	-178.5	48.2	106.1	**-188.1**	26.6	88.5
AT^+	4.129	No BCP					
AT^-	4.249	No BCP					

(a) SF%(RX, X = 1 - 5) is the SF% contribution from all atoms belonging to ring RX. For Ring labeling, see Fig. 1. Other quantities retain the same meaning as in Table 2 and, for the donor to acceptor distance, also the same units.

(b) The SF% of the ring(s) formed by the triad (H+A+D) of atoms directly involved in the hydrogen bond is (are) denoted in bold.

(c) R3 is not a "topological" ring for the charged AT^+ and AT^- pairs at their optimized geometries.

Table 5 Guanine:Cythosine (GC): Source Function percentage (SF%) Ring contributions to the electron density at the hydrogen bond BCP s in the neutral and charged base pairs at their fully optimized geometries[a,b]

Bond/system	$R_{D\cdots A}$	SF%(H+D+A)	SF%(R1)	SF%(R2)	SF%(R3)	SF%(R4)	SF%(R5)
N12-H13···O29							
GC	2.935	42.0	5.6	10.5	4.6	**50.2**	21.3
GC^+	2.678	67.5	3.0	6.2	3.9	**75.5**	12.4
GC^-	2.706	65.9	3.0	6.4	3.9	**72.5**	12.4
N9-H10···N27							
GC	2.947	28.7	6.9	44.5	**43.8**	**42.3**	33.0
GC^+	2.844	44.1	5.1	35.5	**56.9**	**57.1**	37.9
GC^-	2.857	44.2	5.2	36.0	**55.5**	**55.7**	37.6
N24-H26···O8							
GC	2.798	57.6	5.4	14.9	**66.4**	2.7	6.8
GC^+	2.984	35.5	8.0	22.2	**47.8**	3.8	10.4
GC^-	3.017	36.4	8.8	22.7	**45.4**	1.6	8.3

(a) SF%(RX, X = 1 - 5) is the SF% contribution from all atoms belonging to ring RX. For Ring labeling, see Fig. 1. Other quantities retain the same meaning as in Table 4 and, for the donor to acceptor distance, also the same units.

(b) The SF% of the ring(s) formed by the triad (H+A+D) of atoms directly involved in the hydrogen bond is (are) denoted in bold.

Note that the SF% contributions from the various rings do not sum up to 100%, because the contributions from atoms shared by two rings are counted twice and those

from the H atoms not involved in the hydrogen bonds are not included in the sum. SF%(RX, X=1 - 5) data shown in Tables 4 and 5 serve to distinguish the various hydrogen bonds in terms of the different involvement the rings have in determining the electron density at the hydrogen bonds' BCPs.

While the electron density at the N-H···O hydrogen bonds is dominated by R4 (N8-H10···O26 in AT and N12-H13···O29 in GC) or R3 (N24-H26···O8 in GC) sources, for the central N-H···N bond in GC and AT both R3 and R4 sources become relevant. This is in some way expected. Yet comparable sources also come from R2 and R5, denoting the peculiar delocalized nature of the central N-H···N bond in the two complexes, both in the neutral and charged forms. The very weak C-H···O bond in neutral AT has then extremely delocalized sources. Its associated ring R3 yields a very negative SF% contribution, which is overbalanced by large positive contributions from the other rings, in particular R2 and R5.

A question one can pose is whether the source function patterns for the hydrogen bonds in the AT and GC pairs closely resemble, at similar donor to acceptor distances, those of the model systems discussed above or whether they exhibit some specific trait. Comparison of the SF% data in Tables 2-3 with those listed in Table 1 and the SF% portraits shown in Figs. 3-4 *vs* those illustrated in Fig. 2 reveals important differences that hold true both for neutral and charged base pairs.

Firstly, at similar $R_{D\cdots A}$ distances, the hydrogen atoms involved in the N-H···N and NH···O hydrogen bonds, are significantly displaced towards their acceptor atoms in the DNA base pairs. For instance, $R_{D\cdots H}$ for these kinds of hydrogen bonds does not exceed 1.010 Å in the model complexes, regardless of the $R_{D\cdots A}$ value, while it is higher than 1.020 and even as large as 1.062 Å in the DNA base pairs.

The CH···O bond behaves differently since $R_{D\cdots H}$ even slightly decreases in the GC and AT systems relative to the values in the $H_3CH\cdots OH_2$ model complex. Lengthening of $R_{D\cdots H}$ and corresponding shortening of $R_{H\cdots A}$ for the N-H···N and NH···O hydrogen bonds in the GC and AT complexes leads to an increase of the SF% from the H involved in the hydrogen bond, which becomes less negative or even positive in the DNA base pairs relative to the model complexes at similar $R_{D\cdots A}$ distances. N-H···N and NH···O hydrogen bonds are clearly strengthened in GC and AT with respect to the

model complexes.

The effect of strengthening is however apparently less evident for SF%(H+A) and in particular for SF%(H+D+A). As for SF%(H), the former quantity is also generally higher in value than in the model complexes of similar $R_{D\cdots A}$ distances, indicating that the H···A bond has slightly increased its covalency in the DNA base pairs. Instead, the SF%(H+D+A) is always much lower in these pairs than in the model complexes of similar $R_{D\cdots A}$ distances. This turns out to be especially true for the central N-H···N bond in both AT and GC pairs. SF%(H+D+A) becomes even negative in value (-1.4%) in AT^- ($R_{D\cdots A}$ = 3.136 Å), while it is still clearly positive (17.3%) in the model complex at a longer $R_{D\cdots A}$ = 3.2 Å. At $R_{D\cdots A}$ distances around 2.84-2.86 Å (GC^+, GC^-) and 2.88 Å (AT), SF%(H+D+A) does not exceed 45%, while it is almost as large in the model complex at a much longer $R_{D\cdots A}$ distance of 3.0 Å.

The low SF%(H+D+A) values in the DNA base pairs, especially for the central N-H···N hydrogen bond is another sign of the enhanced delocalization of such bonds, relative to the case of the model complexes, besides that of the unexpectedly large SF% contributions to the hydrogen bond BCP density from rings other than R3 and R4 (Tables 4 and 5). Such an enhanced delocalization may be easily appreciated by visual comparison of Figs. 3-4 and 2. It also explains why SF%(H+A) does increase less evidently than SF%(H) for N-H···N and NH···O hydrogen bonds, relative to the corresponding values in the model complexes at similar $R_{D\cdots A}$ distances.

The case of the CH···O hydrogen bond in the AT complex is different. As said before, $R_{D\cdots H}$ shortens, rather than lengthening relative to the model complex at similar $R_{D\cdots A}$ distance and, as a consequence, SF%(H) becomes largely more negative than in $H_3CH\cdots OH_2$. Also SF%(H+A) and SF%(H+D+A) remain negative and even more so than SF%(H), at variance with the case of the model complex. The CH···O hydrogen bond in AT is thus largely weakened relative the bond with corresponding $R_{D\cdots A}$ distance in the model complex. This is also revealed by a comparison of the corresponding ρ_{BCP} values ($H_3CH\cdots OH_2$, ρ_{BCP} = 0.0059 a.u. at $R_{D\cdots A}$ = 3.8 Å; AT, ρ_{BCP} = 0.0043 a.u. despite the shorter $R_{D\cdots A}$ value of 3.689 Å). Eventually, this leads to the lack of a BCP in the charged AT species and at $R_{D\cdots A}$ distances where a clear BCP is still present in the model complex.

Interesting insights on the localised/delocalised hydrogen bonds nature in the neutral and charged AT and GC base pairs may be also gained from the SF reconstructed partial electron density maps shown in Fig. 5 (AT), Fig. 6 (GC), Fig. 7 (AT^+) and Fig. 8 (AT^-). As discussed before, the advantage of such maps is that of enlarging our view besides that recovered at some selected reference points. Partial reconstructions from only ring R3+R4 atoms sources, $\rho\{R3+R4\}$, or from only R3, $\rho\{R3\}$, or only R4, $\rho\{R4\}$, atoms sources are compared in these figures with those including all atoms sources, $\rho\{all\}$. This latter density equals ρ within absolutely negligible numerical errors.

Figs. 5 and 6 both highlight three important points. Firstly, there is a large influence of rings 3 sources, $\rho\{R3\}$, to the electron density in the region of the external hydrogen bond in ring 4 and, analogously, a large influence of ring 4 sources, $\rho\{R4\}$, to the electron density in the region of the external hydrogen bond in ring 3. Secondly, significantly large regions of negative electron density occur in the centers of the rings not only when the ring 3 or the ring 4 sources alone are taken into account, but also when both ring 3 and ring 4, $\rho\{R3+R4\}$, sources are included. Such regions become positive only when the sources from the excluded atoms of rings 1-2 and those of ring 5 are also considered. Third, the electron density along the NH···O, and in particular along the NH···N central bond paths is appreciably lowered relative to $\rho\{all\}$, even when both R3 and R4 atoms sources are included. This becomes dramatically evident for the weak CH···O hydrogen bond in AT, where a significant region of negative sources, rather than of positive ED, occurs along the bond path. Similar general considerations apply to the partial electron density maps of charged AT complexes (Figs. 7 and 8) and of charged GC base pairs (not shown), though geometry changes due to electron uptake or detachment, lead to visible variations.

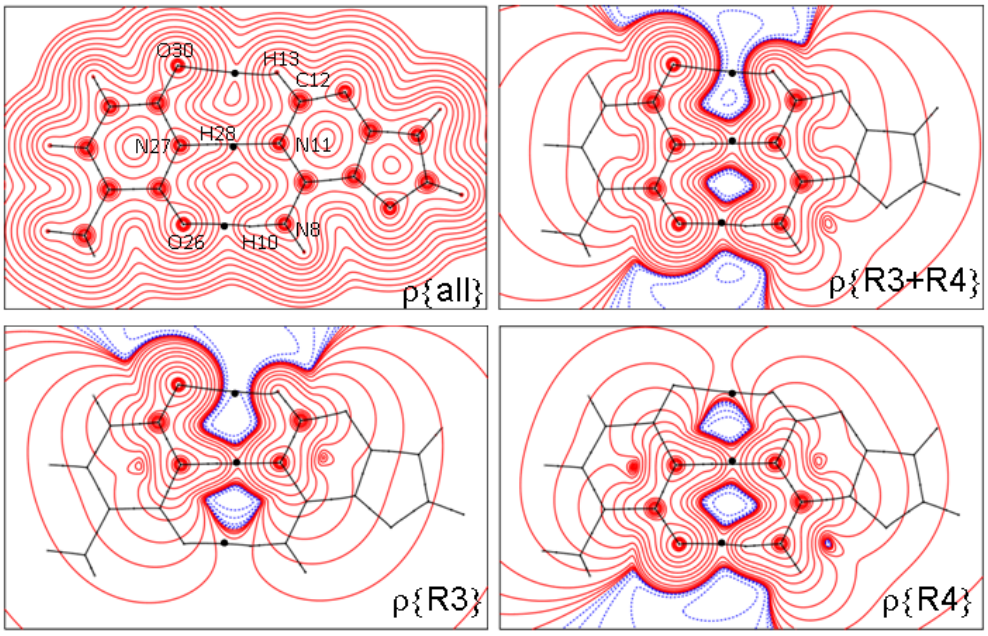

Fig. 5 Adenine:thymine neutral complex: Source function reconstructed partial electron densities for various atomic subsets $\{\Omega\}$ in the least squares molecular plane. Positions of the hydrogen bonds' bond critical points are denoted by full black dots. Bond paths are also shown. Contours are drawn at interval of $\pm(2,4,8)\cdot 10^{n}$, $-3 \le n \le 0$ atomic units (a.u.) Red positive and, dotted blue, negative contour values. (The orientation of the dimer in this figure is rotated 180° along the horizontal central axis with respect to the orientation in Fig. 1).

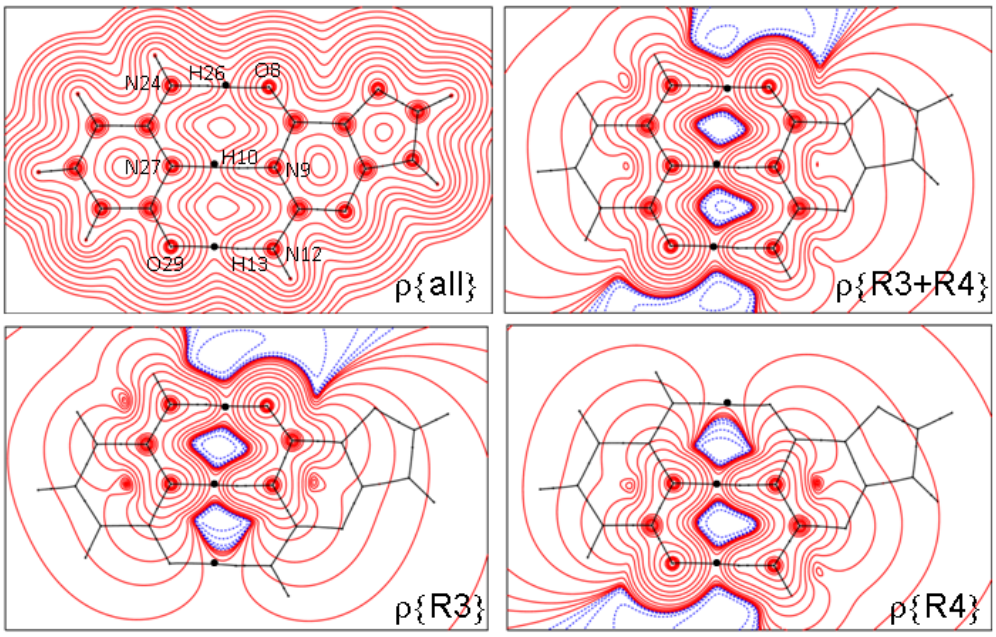

Fig. 6 Guanine:cytosine neutral complex: Source function reconstructed partial electron densities for various atomic subsets $\{\Omega\}$ in the least squares molecular plane. Positions of the hydrogen bonds' bond critical points are denoted by full black dots. Bond paths are also shown. Contours are drawn at interval of $\pm(2,4,8)\cdot 10^{n}$, $-3 \le n \le 0$ atomic units (a.u.) Red positive and, dotted blue, negative contour values. (The orientation of the dimer in this figure is the same as in Fig. 1).

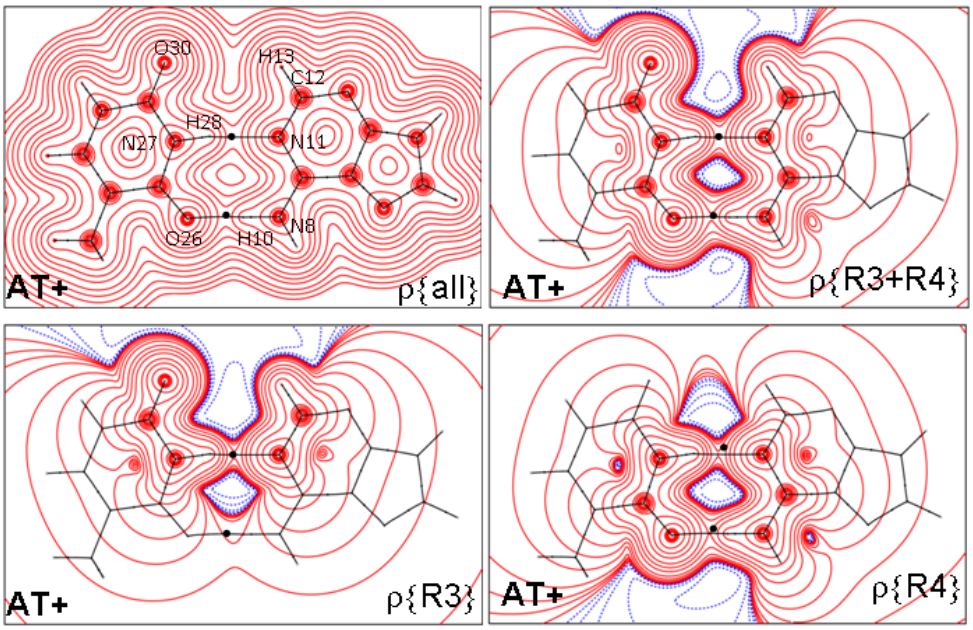

Fig. 7 Adenine:thymine positively charged complex, AT^{+}: Source function reconstructed partial electron densities for various atomic subsets $\{\Omega\}$ in the least squares molecular plane. Positions of the Hydrogen bonds bond critical points are denoted by full black dots. Bond paths are also shown. Contours are drawn at interval of $\pm(2,4,8)\cdot 10^{n}$, $-3 \le n \le 0$ atomic units (a.u.) Red positive and, dotted blue, negative contour values. (The orientation of the dimer in this figure is rotated 180° along the horizontal central axis with respect to the orientation in Fig. 1).

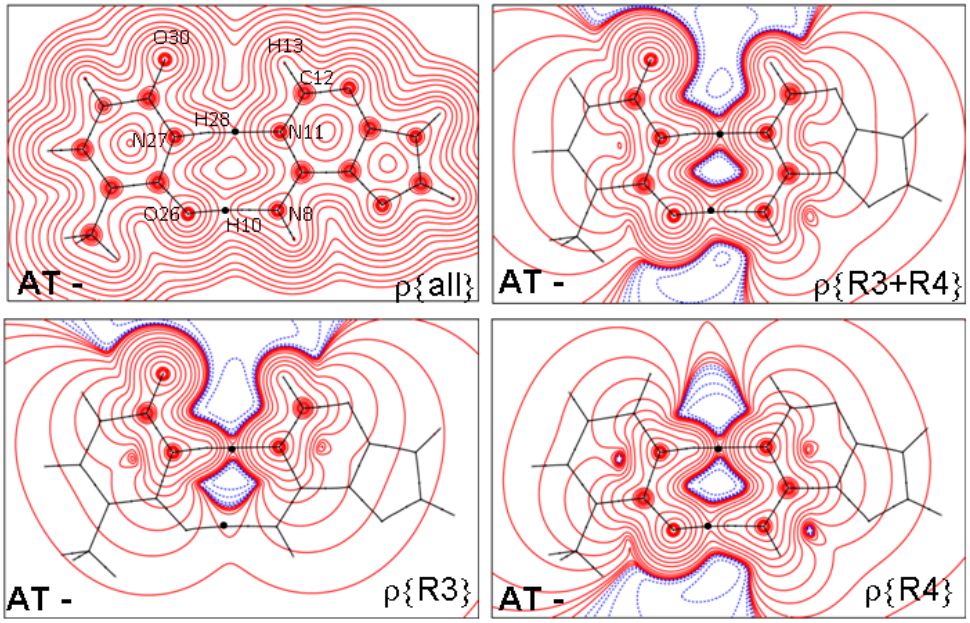

Fig. 8 Adenine:thymine negatively charged complex, AT^{-}: Source function reconstructed partial electron densities for various atomic subsets $\{\Omega\}$ in the least squares molecular plane. Positions of the Hydrogen bonds bond critical points are denoted by full black dots. Bond paths are also shown. Contours are drawn at interval of $\pm(2,4,8)\cdot 10^{n}$, $-3 \le n \le 0$ atomic units (a.u.) Red positive and, dotted blue, negative contour values. (The orientation of the dimer in this figure is rotated 180° along the horizontal central axis with respect to the orientation in Fig. 1).

To conclude, Fig. 9 shows, as an example, the partial electron density due to only the sources from the donor, acceptor and hydrogen atoms of the central NH···N hydrogen bond in the AT neutral and charged complexes. The three maps in Fig. 9 are qualitatively similar, yet immediately reveal the significant NH···N hydrogen bond weakening on going from the neutral to the positively charged and to the negatively charged complex.

The two regions of negative sources present in the ring 3 and 4 in AT approach each other in AT^+ and finally almost merge in AT^-. Comparison with ρ{R3+R4} (and ρ{R3} or ρ{R4}) maps in Figs. 5, 7, and 8 shows the relevant role that atoms in rings 3 and/or 4, other than the hydrogen bond triad, play in determining the electron density in the central NH···N hydrogen bond region, providing further conclusive evidence for the quite delocalised nature of such hydrogen bonded interaction.

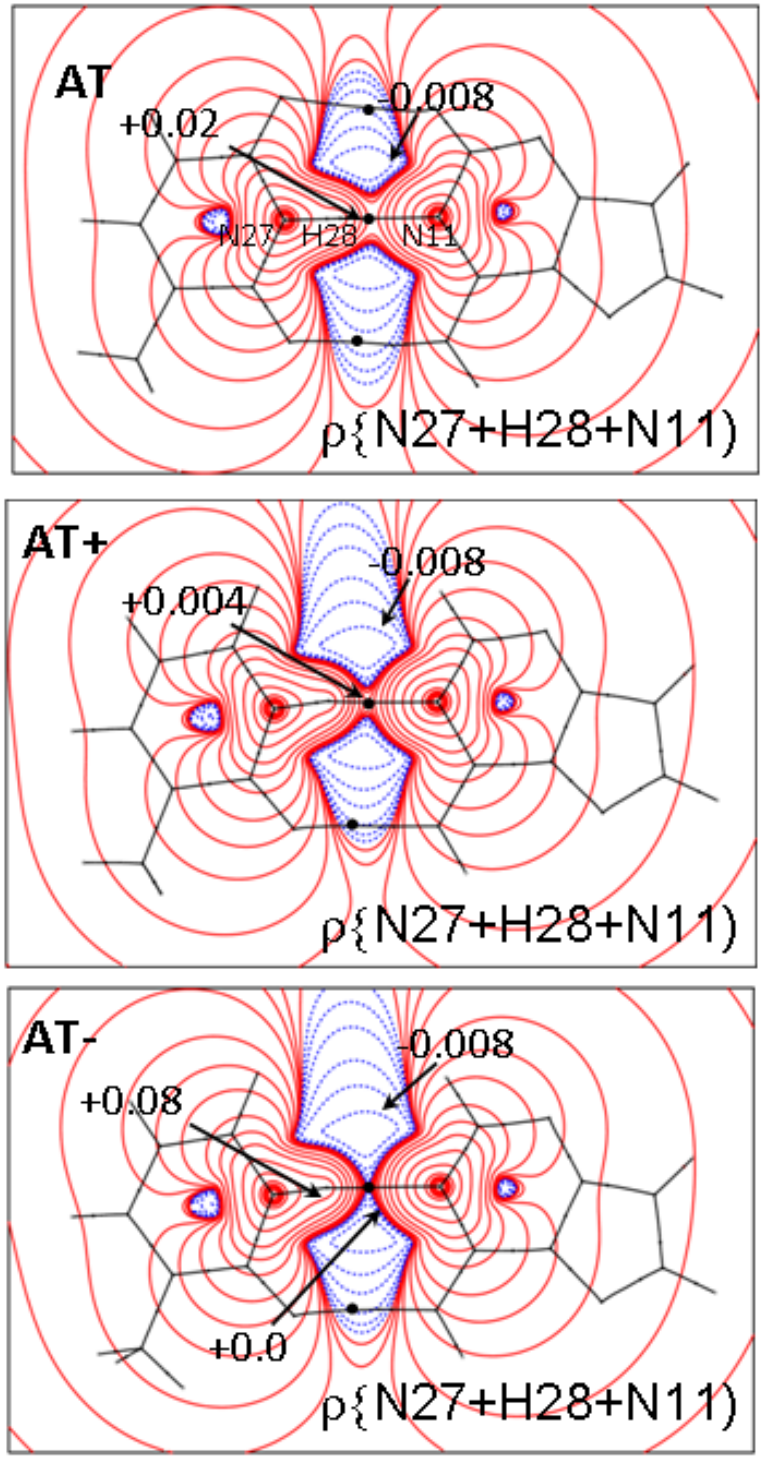

Fig. 9 Adenine:thymine neutral and charged complexes: Source function reconstructed partial electron densities in the least squares molecular plane for the triad of atoms (N27, H28,N11) involved in the central N-H···N bond. Positions of the hydrogen bonds' bond critical points are denoted by full black dots. Bond paths are shown. Contours are drawn at interval of $\pm(2,4,8)\cdot 10^{n}$, $-3 \leq n \leq 0$ atomic units (a.u.) Red positive and, dotted blue, negative contour values (with selected contour values are given in a.u.) (The orientation of the dimer in this figure is rotated 180^{o} along the horizontal central axis with respect to the orientation in Fig. 1).

Closing Remarks

It is possible for a scientific question not to ever occur if there exists no conceivable approach to answer it in the first place. DNA is without a doubt one of the most important and most heavily studied biopolymers. A simple internet search on the keyword "DNA" returns 474,000,000 results in 0.49 s.[101] Hence research papers on DNA must count by the tens of millions. With this volume of accumulated knowledge, it is fair to ask what is the gap in knowledge that this paper is attempting to fill regarding DNA?

The gap in knowledge that this paper addresses can be translated into the following questions that refer to the two Watson-Crick DNA dimers (AT and GC):

(1) What are the atoms that are responsible for most of the binding through the hydrogen bonds in the dimer?

(2) Are there dominant atoms? Dominant rings? Or even a dominant monomer that drives the binding of the dimer together?

(3) What is the effect of ionization by either losing an electron to ionizing radiation or gaining one from, say, a solvated electron on the order of dominance in atomic or group contributions?

These questions would not have had a rigorous and model-independent answer prior to the pioneering discovery of the source function and its analysis. The best that could have been done in the past was to calculate basis function contributions, or stock-holder contributions to a point of reference with all the well-known problems these approaches have (non-uniqueness, basis set dependency, *etc*.) The SF analysis, by being unique, is also unique in providing the bridge to experiment.

Applied to the WC DNA dimers, the SF analysis uncovers truly remarkable facts - previously unknown – about the extent to which a given hydrogen bond in a dimer (or its ion) depends on distant atoms far from the hydrogen bond triad *per se*. A notable example is that ionization can have little net effects on some of the hydrogen bonds that hide a subtle by significantly different interplay of sources and sinks behind the scenes. That resiliency of WC hydrogen bonds in DNA could be a mechanism to maintain DNA's stability in the face of environmental assaults at least to give time for the polaron to travel though the α-helix, be dissipated to the surroundings, or for the repair mechanisms to start. Evolution may have, therefore, fine tuned the atomic

composition of the bases to minimize the effect of ionization on the ensuing hydrogen bonds. This hypothesis will be the subject of a future study.

Another finding of interest is that coarser grained sources reveal that there is no particular advantage for the purine or the pyrimidine base in determining the strength of the hydrogen bonding whether in the neutral or the ionized dimers. The binding of either of the WC dimers, whether in the neutral, singly cationic, or singly anionic forms, is equally determined by each of the monomers composing it irrespective of that monomer's size.

The double helix was conceived by Watson and Crick in 1953 on the basis of the X-ray diffraction pattern obtained by Rosalind Franklin. In the 65 years to the time of writing, X-ray diffraction has moved light years ahead on all fronts. Similar spectacular advances have occurred during this time in computer technology and in the theory and practice of quantum chemistry. While the techniques themselves diverge quickly by being increasingly more specialized, paradoxically, the object of their study - the electron density - unite them more than ever. We now have highly accurate experimental and theoretical electron density maps to both of which the SF analysis can be applied. Comparison of the SF results from theory and experiment can serve in the bench-marking of new theoretical methods or basis sets being developed. The SF is also a promising molecular (or unit cell descriptor) that is well suited for quantitative structure activity relationship (QSAR) studies.[102]

One principal advantage of the SF should be clear: It goes beyond the picture of the electron density as just a "local" scalar function and expresses it, instead, as an effect of a cause distributed in all space. In effect, it provides a language to discuss electron delocalization[67,70] that can be applied to both theoretically-generated and experimentally-fitted electron densities[70] as has been recently emphasized in an article regarding molecules as networks.[102]

The SF is blossoming at a time when "Quantum Crystallography", the science of constraining the densities refined from X-ray diffraction to mathematical requirements from quantum mechanics, is seeing a strong renewed interest.[103-111] The entire field of study of the electron density is, in more than one way, undergoing a fast evolution on the front of obtaining quantum mechanically correct experimental electron densities (and density matrices) that can then be compared with those obtained from theory in

benchmarking using tools such as the SF analysis as a base for comparison. The recent extension[84,112] of the SF tool to the analysis of the electron spin densities provide further precious insights in such comparison benchmarking.[76,84,112]

Acknowledgements

C.G. and G.M. acknowledge funding from Danmarks Grundforskningsfond (award No. DNRF93). C.F.M. and R.J.B. acknowledge the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canada Foundation for Innovation (CFI), Mount Saint Vincent University for funding. C.F.M. is also grateful to the Killam Trusts for an Izaak Walter Killam Postdoctoral Award in the period 2004-2006 when this work began.

References

[1] E. Schrödinger, What is Life?; Cambridge University Press: Cambridge, **1944**.

[2] F. H. C. Crick. The genetic code: III. *Sci. Am.* **1966**, *215 (No. 4, Oct.)*, 55-63.

[3] M. V. Volkenstein. The genetic coding of protein structure. *Biochim. Biophys. Acta* **1966**, *119*, 421-424.

[4] M. V. Volkenshtein, Molecules and Life: An Introduction to Molecular Biology; Plenum Press: New York, **1970**.

[5] M. V. Volkenstein, Molecular Biophysics; Academic Press, Inc.: New York, **1977**.

[6] L. L. Gatlin, Information Theory and the Living System; Columbia University Press: New York, **1972**.

[7] F. Miescher. Ueber die chemische zusammensetzung der eiterzellen (On the chemical composition of pus cells). *Medicinisch-chemische Untersuchungen* **1871**, *4*, 441-460.

[8] R. Dahm. Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. *Hum. Genet.* **2008**, *122*, 565-581.

[9] O. T. Avery, C. M. MacLeod, M. McCarty. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. *J. Exp. Med.* **1944**, *79*, 137-158.

[10] F. Griffith. The significance of Pneumococcal types. *J. Hyg.* **1928**, *27*, 113-159.

[11] G. S. Stent. Prematurity and uniqueness in scientific discovery. *Sci. Am.* **1972**, *227 (No. 6, Dec.)*, 84-93.

[12] E. Vischer, E. Chargaff. The composition of the pentose nucleic acids of yeast and pancreas. *J. Biol. Chem.* **1948**, *176*, 715-734.

[13] E. Chargaff, E. Vischer, R. Doniger, C. Green, F. Misani. The composition of the desoxypentose nucleic acids of thymus and spleen. *J. Biol. Chem.* **1949**, *177*, 405-416.

[14] E. Chargaff. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* **1950**, *6*, 201-209.

[15] J. D. Watson, F. H. C. Crick. The structure of DNA. *Cold Spring Harbor Symposia on Quantitative Biology* **1953**, *XVIII*, 123-131.

[16] J. D. Watson, F. H. C. Crick. A structure for deoxyribose nucleic acid. *Nature* **1953**, *171*, 737-738.

[17] J. D. Watson, F. H. C. Crick. Genetical implications of the structure of deoxyribose nucleic acid. *Nature* **1953**, *171*, 964-967.

[18] F. H. C. Crick, J. D. Watson. The complementary structure of deoxyribonucleic acid. *Proceedings of the Royal Society (London)* **1954**, *A 223*, 80-96.

[19] J. D. Watson, Molecular Biology of the Gene (Second Edition); W. A. Benjamin, Inc.: New York, **1970**.

[20] J. D. Watson, The Double Helix: A Personal Account of the Discovery of the Structure of DNA (Edited by G. S. Stent); W. W. Norton & Co.: New York, **1980**.

[21] R. E. Franklin, R. G. Gosling. Molecular configuration in sodium thymonucleate. *Nature* **1953**, *171*, 740-741.

[22] L. Pauling, R. B. Corey. Structure of the nucleic acids. *Nature* **1953**, *171*, 346.

[23] L. Pauling, R. B. Corey. A proposed structure for the nucleic acids. *Proc. Natl. Acad. Sci. USA* **1953**, *39*, 84-97.

[24] R. Parthasarathi, R. Amutha, V. Subramanian, B. U. Nair, T. Ramasami. Bader's and reactivity descriptors' analysis of DNA base pairs. *J. Phys. Chem. A* **2004**, *108*, 3817-3828.

[25] S. Ed. Neidle, Oxford Handbook of Nucleic Acid Structure; Oxford University Press: Oxford, **1999**.

[26] J. Sponer, J. Leszczynski, P. Hobza. Hydrogen bonding and stacking in DNA bases: A review of quantum-chemical *ab initio* studies. *J. Biomol. Struct. Dynam.* **1996**, *14*, 117-135.

[27] J. Sponer, I. Berger, N. Spackova, J. Leszczynski, P. Hobza. Aromatic base stacking in DNA: From *ab initio* calculations to molecular dynamics simulations. *J. Biomol. Struct. Dynam.* **2000**, *S2*, 383-407.

[28] P. Hobza, J. Sponer. Toward true DNA base-stacking energies: MP2, CCSD(T), and complete basis set calculations. *J. Am. Chem. Soc.* **2002**, *124*, 11802-11808.

[29] P.-O. Löwdin. Proton tunneling in DNA and its biological implications. *Rev. Mod. Phys.* **1963**, *35*, 721-733.

[30] P.-O. Löwdin. Quantum genetics and the aperiodic solid: some aspects on the biological problems of heredity, mutation, aging, and tumors in view of the quantum theory of the DNA molecule. *Adv. Quantum Chem.* **1965**, *2*, 213-360.

[31] R. R. Dogonadze, Y. I. Kharkats, J. Ulstrup. A dynamic quantum mechanical model for the elementary act of spontaneous and induced mutations. *Chem. Phys. Lett.* **1976**, *37*, 360-364.

[32] 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**.

[33] E. Kryachko, J. R. Sabin. Quantum chemical study of the hydrogen-bonded patterns in A.T base pair of DNA: Origins of tautomeric mispairs, base flipping, and Watson-Crick → Hoogsteen conversion. *Int. J. Quantum Chem.* **2003**, *91*, 695-710.

[34] L. Gorb, Y. Podolyan, J. Leszczynski, W. Siebrand, A. Fernández-Ramos, Z. Smedarchina. A quantum-dynamics study of the prototropic tautomerism of guanine

and its contribution to spontaneous point mutations in Escherichia coli. *Biopolymers* **2002**, *61*, 77-83.

[35] L. Gorb, Y. Podolyan, P. Dziekonski, W. A. Sokalski, J. Leszczynski. Double-proton transfer in adenine-thymine and guanine-cytosine base pairs. A post-Hartree-Fock ab initio study. *J. Am. Chem. Soc.* **2004**, *126*, 10119-10129.

[36] Y. Podolyan, L. Gorb, J. Leszczynski. *Ab initio* study of the prototropic tautomerism of cytosine and guanine and their contribution to spontaneous point mutations. *Int. J. Mol. Sci.* **2003**, *4*, 410-421.

[37] L. Mejía-Mazariegos, J. Hernández-Trujillo. Electron density analysis of tautomeric mechanisms of adenine, thymine and guanine and the pairs of thymine with adenine or guanine. *Chem. Phys. Lett.* **2009**, *482*, 24-29.

[38] D. N. Stamos. "Quantum indeterminism, mutation, natural selection, and the meaning of life", Chapter 31 in: *Quantum Biochemistry: Electronic Structure and Biological Activity (Vol. 2)*; C. F. Matta, Ed., Wiley-VCH: Weinheim, 2010; pp. 837-871.

[39] C. F. Matta (Ed.) Quantum Biochemistry: Electronic Structure and Biological Activity (Vols. 1 & 2); Wiley-VCH: Weinheim, **2010**.

[40] L.-Y. Fu, G.-Z. Wang, B.-G. Ma, H.-Y. Zhang. Exploring the Common Molecular Basis for the Universal DNA Mutation Bias: Revival of Lowdin Mutation Model. *Biochem. Biophys. Res. Commun.* **2011**, *409*, 367-371.

[41] O. O. Brovarets, R. O. Zhurakivsky, D. M. Hovorun. The physico-chemical mechanism of the tautomerisation via the DPT of the long Hyp...Hyp WatsonCrick base pair containing rare tautomer: A QM and QTAIM detailed look. *Chem. Phys. Lett.* **2013**, *578*, 126-132.

[42] O. O. Brovarets', D. M. Hovorun. Does the G·G*syn DNA mismatch containing canonical and rare tautomers of the guanine tautomerise through the DPT? A QM/QTAIM microstructural study. *Mol. Phys.* **2014**, *112*, 3033-3046.

[43] O. O. Brovarets', D. M. Hovorun. How many tautomerization pathways connect WatsonCrick-like G*· T DNA base mispair and wobble mismatches? *J. Biomol. Struct. Dynam.* **2015**, *33*, 2297-2315.

[44] O. O. Brovarets', D. M. Hovorun. The nature of the transition mismatches with WatsonCrick architecture: the G*·T or G·T* DNA base mispair or both? A QM/QTAIM perspective for the biological problem. *J. Biomol. Struct. Dynam.* **2015**, *33*, 925-945.

[45] B. Pullman, A. Pullman, Quantum Biochemistry; Interscience Publishers: New York, **1963**.

[46] J. J. Ladik. Toward the electronic structure of real DNA. *Int. J. Quantum Chem.* **1975**, *9 (Issue S2)*, 133-143.

[47] A. A. Arabi, C. F. Matta. Effects of external electric fields on double proton transfer kinetics in the formic acid dimer. *Phys. Chem. Chem. Phys. (PCCP)* **2011**, *13*, 13738-13748.

[48] J. P. Cerón-Carrasco, D. Jacquemin. Electric field induced DNA damage: An open door for selective mutations. *Chem. Commun.* **2013**, *49*, 7578-7580.

[49] J. P. Cerón-Carrasco, D. Jacquemin. Electric-field induced mutation of DNA: A theoretical investigation of the GC base pair. *Phys. Chem. Chem. Phys. (PCCP)* **2013**, *15*, 4548-4553.

[50] J. P. Cerón-Carrasco, J. Cerezoa, D. Jacquemin. How DNA is damaged by external electric fields: Selective mutation vs. random degradation. *Phys. Chem. Chem. Phys.*

(PCCP) **2014**, *16*, 8243-8246.

[51] L. P. Candeias, S. Steenken. Ionization of purine nucleosides and nucleotides and their components by 193-nm laser photolysis in aqueous solution: model studies for oxidative damage of DNA. *J. Am. Chem. Soc.* **1992**, *114*, 699-704.

[52] R. G. Parr, W. Yang, Density-Functional Theory of Atoms and Molecules; Oxford University Press: Oxford, **1989**.

[53] J. Gu, Y. Xie, H. F. I. Schaefer. Structural and energetic characterization of a DNA nucleoside pair and its anion: deoxyriboadenosine (dA)-deoxyribothymidine (dT). *J. Phys. Chem. B* **2005**, *109*, 13067-13075.

[54] B. Pullman, A. Pullman. Submolecular structure of the nucleic acids. *Nature (London)* **1961**, *189*, 725-727.

[55] R. R. Parthasarathi, V. Subramanian. Stacking interactions in benzene and cytosine dimers: From molecular electron density perspective. *Struct. Chem.* **2005**, *16*, 243-255.

[56] R. R. Parthasarathi, V. Subramanian. Hydrogen bonding of DNA base pairs and information entropy: From molecular electron density perspective. *Chem. Phys. Lett.* **2006**, *418*, 530-534.

[57] M. P. Waller, A. Robertazzi, J. A. Platts, D. E. Hibbs, P. A. Williams. Hybrid density functional theory for -stacking interactions: Application to benzenes, pyridines, and DNA bases. *J. Comput. Chem.* **2006**, *27*, 491-504.

[58] C. F. Matta, N. Castillo, R. J. Boyd. Extended weak bonding interactions in DNA: π-Stacking (base-base), base-backbone, and backbone-backbone interactions. *J. Phys. Chem. B* **2006**, *110*, 563-578.

[59] C. F. Matta. Modeling biophysical and biological properties from the characteristics of the molecular electron density, electron localization and delocalization matrices, and the electrostatic potential. *J. Comput. Chem.* **2014**, *35*, 1165-1198.

[60] R. F. W. Bader, Atoms in Molecules: A Quantum Theory; Oxford University Press: Oxford, U.K., **1990**.

[61] P. L. A. Popelier, Atoms in Molecules: An Introduction; Prentice Hall: London, **2000**.

[62] C. F. Matta, R. J. Boyd (Eds.) The Quantum Theory of Atoms in Molecules: From Solid State to DNA and Drug Design; Wiley-VCH: Weinheim, **2007**.

[63] R. F. W. Bader, C. Gatti. Green's function for the density. *Chem. Phys. Lett.* **1998**, *287*, 233-238.

[64] C. Gatti, F. Cargnoni, L. Bertini. Chemical information from the source function. *J. Comput. Chem.* **2003**, *24*, 422-436.

[65] C. Gatti, L. Bertini. The local form of the source function as a fingerprint of strong and weak intra- and intermolecular interactions. *Acta Cryst. A* **2004**, *60*, 438-449.

[66] C. Gatti, D. Lasi. Source function description of metal-metal bonding in d-block organometallic compounds. *Farad. Discuss.* **2007**, *135*, 55-78.

[67] E. Monza, C. Gatti, L. Lo Presti, E. Ortoleva. Revealing electron delocalization through the source function. *J. Phys. Chem. A* **2011**, *115*, 12864-12878.

[68] C. Gatti. The source function descriptor as a tool to extract chemical information from theoretical and experimental electron densities. *Struct. Bond. (Electron Density and Chemical Bonding II - Theoretical Charge Density Studies, D. Stalke (Ed.))* **2012**, *147*, 193-286.

[69] C. Gatti. Challenging chemical concepts through charge density of molecules and crystals. *Phys. Scr.* **2013**, *87*, 048102.

[70] C. Gatti, G. Saleh, L. Lo Presti. Source function applied to experimental densities reveals subtle electron delocalization effects and appraises their transferability properties in crystals. *Acta Cryst. A* **2016**, *72*, 180-193.

[71] F. Meng, C. Liu, W. Xu. Substituent effects of R (R = CH_3, CH_3O, F and NO_2) on the A:T and G:C base pairs: a theoretical study. *Chem. Phys. Lett.* **2003**, *373*, 72-78.

[72] S. N. Vinogradov, R. H. Linnell. Hydrogen Bonding; Van Nostrand Reinhold Co.: New York, **1971**.

[73] G. A. Jeffrey. An Introduction to Hydrogen Bonding; Oxford University Press: Oxford, **1997**.

[74] G. Gilli, P. Gilli. The Nature of the Hydrogen Bond: Outline of a Comprehensive Hydrogen Bond Theory; Oxford University Press: Oxford, **2009**.

[75] G. Arfken. Mathematical Methods for Physicists; Academic Press: Orlando, Florida, **1985**.

[76] C. Gatti, G. Macetti, L. Lo Presti. Insights on spin delocalization and spin polarization mechanisms in crystals of azido copper(II) dinuclear complexes through the electron spin density Source Function. *Acta Cryst. B* **2017**, *73*, 565-583.

[77] C. Lee, W. Yang, R. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron-density. *Phys. Rev. B* **1988**, *37*, 785-789.

[78] A. Becke. Density-functional thermochemistry.3. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648-5652.

[79] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian 09, Revision B.01 (Gaussian Inc., Wallingford CT, 2010).

[80] C. Gatti, AIMPAC modified to evaluate SF contributions (unpublished, available on request: carlo.gatti@istm.cnr.it); NRC: Milano, **2018**.

[81] Bader, R. F. W., AIMPAC http://www.chemistry.mcmaster.ca/aimpac/ (accessed 10 July **2017**).

[82] 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.

[83] F. W. Biegler-König, R. F. W. Bader, T.-H. Tang. Calculation of the average properties of atoms in molecules. II. *J. Comput. Chem.* **1982**, *13*, 317-328.

[84] C. Gatti, A. M. Orlando, L. Lo Presti. Insights on spin polarization through the spin density Source Function. *Chem. Sci.* **2015**, *6*, 3845-3852.

[85] H. Putz, K. Brandenburg, Diamond - Crystal and Molecular Structure Visualization (http://www.crystalimpact.com/diamond); Crystal Impact, GbR: Kreuzherrenstrasse

102, 53227 Bonn, Germany, **2012**.

[86] J. Overgaard, B. Schiųtt, F. K. Larsen, B. B. Iversen. The charge density distribution in a model compound of the catalytic triad in serine proteases. *Chem. Eur. J.* **2001**, *7*, 3756-3767.

[87] G. Gilli, P. Gilli. Towards an unified hydrogen-bond theory. J. Mol. Struct. *J. Mol. Struct.* **2000**, *552*, 1-15.

[88] G. A. Jeffrey, W. Saenger. Hydrogen Bonding in Biological Structures; Springer Verlag: Berlin, **1991**.

[89] P. L. A. Popelier. "Quantum chemical topology" in: The Chemical Bond II - 100 Years Old and Getting Stronger (Mingos, D. and Michael, P, Eds.) - Struct. Bond. 170. (pp. 71-117); Springer International Publishing AG: Switzerland, **2016**.

[90] Blanco M. A., A. Martín Pendás, E. Francisco. Interacting quantum atoms: A correlated energy decomposition scheme based on the quantum theory of atoms in molecules. *J. Chem. Theor. Comput.* **2005**, *1*, 1096-1109.

[91] A. Martín Pendás, M. A. Blanco, E. Francisco. The nature of the hydrogen bond: A synthesis from the interacting quantum atoms picture. *J. Chem. Phys.* **2006**, *125*, 184112.

[92] A. Martín Pendás, E. Francisco, M. A. Blanco. Binding energies of first row diatomics in the light of the interacting quantum atoms approach. *J. Phys. Chem. A* **2006**, *110*, 12864-12869.

[93] E. Francisco, A. Martín Pendás, M. A. Blanco. A molecular energy decomposition scheme for atoms in molecules. *J. Chem. Theory Comput.* **2006**, *2*, 90-102.

[94] A. Martín Pendás, E. Francisco, M. A. Blanco, C. Gatti. Bond paths as privileged exchange channels. *Chem. Eur. J.* **2007**, *13*, 9362-9371.

[95] X. Fradera, M. A. Austen, R. F. W. Bader. The Lewis model and beyond. *J. Phys. Chem. A* **1999**, *103*, 304-314.

[96] F. H. C. Crick. Codon-anticodon pairing: the wobble hypothesis. *J. Mol. Biol.* **1966**, 548-555.

[97] G. A. Leonard, K. McAuley-Hecht, T. Brown, W. N. Hunter. Do C-H···O hydrogen bonds contribute to the stability of nucleic acid base pairs? *Acta Cryst. D* **1995**, *51*, 136-139.

[98] S. F. Boys, F. Bernardi. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553-566.

[99] A. Asensio, N. Kobko, J. J. Dannenberg. Cooperative hydrogen-bonding in adenine-tymine and guanine-cytosine base pairs. Density functional theory and Møller-Plesset molecular orbital study. *J. Phys. Chem. A* **2003**, *107*, 6441-6443.

[100] E. Espinosa, E. Molins, C. Lecomte. Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. *Chem. Phys. Lett.* **1998**, *285*, 170-173.

[101] Google search: DNA; **2018**.

[102] C. F. Matta. Molecules as networks: A localization-delocalization matrices approach. *Theor. Comput. Chem.* **2018**, *1124*, 1-14.

[103] A. Genoni, L. Bucinský, N. Claiser, J. Contreras-Garcia, B. Dittrich, P. M. Dominiak, E. Espinosa, C. Gatti, P. Giannozzi, J.-M. Gillet, D. Jayatilaka, P. Macchi, A. Ø. Madsen, L. Massa, C. F. Matta, K. M. Merz Jr., P. Nakashima, H. Ott, U. Ryde, W.

Scherer, K. Schwarz, M. Sierka, S. Grabowsky. Quantum crystallography: Current developments and future perspectives. *Chem. Eur. J.* **2018**, in revision.

[104] S. Grabowsky, A. Genoni, H.-B. Bürgi. Quantum crystallography. *Chem. Sci.* **2017**, *8*, 4159-4176.

[105] D. Jayatilaka. "Using wavefunctions to get more information out of diffraction experiments." In: Modern Charge-Density Analysis, Carlo Gatti & Piero Macchi (Eds.); Springer: Berlin, **2012**.

[106] C. Gatti, P. Macchi (Eds.) Modern Charge-Density Analysis; Springer: Berlin, **2012**.

[107] L. Massa, C. F. Matta. Quantum cystallography (QCr): Extraction of the complete quantum mechanics from X-ray scattering data. *J. Comput. Chem.* **2018**, in press.

[108] L. Massa, C. F. Matta. Exploiting the full quantum crystallography. *Can. J. Chem.* **2018**, in press.

[109] C. F. Matta. A path through quantum crystallography: A short tribute to Professor Lou Massa. *Struct. Chem.* **2017**, *28*, 1279-1283.

[110] W. Polkosnik, L. Massa. Single determinant N-representability and the kernel energy method (KEM) applied to water clusters. *J. Comput. Chem.* **2018**, in press.

[111] L. Massa. A zigzag path through quantum crystallography. *Struct. Chem.* **2017**, *28*, 1293-1296.

[112] G. Macetti, L. Lo Presti, C. Gatti. Spin density accuracy and distribution in azido Cu(II) complexes: A Source Function analysis, *J. Comp. Chem.* **2018**, *39*, 587-603.