Published in final edited form as:
J Am Chem Soc. 2010 May 12; 132(18): 6498–6506. doi:10.1021/ja100936w.

Revealing Non-Covalent Interactions

Erin R. Johnson[1], **Shahar Keinan**[1], **Paula Mori-Sánchez**[1], **Julia Contreras-García**[1], **Aron J. Cohen**[2], and **Weitao Yang**[1]

[1] Department of Chemistry, Duke University, Durham, North Carolina, 27708

[2] Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, United Kingdom

Abstract

Molecular structure does not easily identify the intricate non-covalent interactions that govern many areas of biology and chemistry, including design of new materials and drugs. We develop an approach to detect non-covalent interactions in real space, based on the electron density and its derivatives. Our approach reveals underlying chemistry that compliments the covalent structure. It provides a rich representation of van der Waals interactions, hydrogen bonds, and steric repulsion in small molecules, molecular complexes, and solids. Most importantly, the method, requiring only knowledge of the atomic coordinates, is efficient and applicable to large systems, such as proteins or DNA. Across these applications, a view of non-bonded interactions emerges as continuous surfaces rather than close contacts between atom pairs, offering rich insight into the design of new and improved ligands.

1 Introduction

Chemical interactions between a protein and a drug, or a catalyst and its substrate, self-assembly of nanomaterials,[1,2] and even some chemical reactions,[3,4] are dominated by non-covalent interactions. This class of interactions spans a wide range of binding energies, and encompasses hydrogen bonding, dipole-dipole interactions, steric repulsion, and London dispersion.[5] Molecular structure is governed by covalent, non-covalent, and electrostatic interactions, the latter two of which are the driving force in most biochemical processes. The three-dimensional molecular structure defines covalent bonds, however, non-covalent interactions are hidden within voids in the bonding network. Although there are several ways to view and analyze covalent and electrostatic interactions, an analogous method for non-covalent interactions is conspicuously missing. Such a method would aid understanding of the complex interactions between biomolecules, and the design of self-assembled materials and drugs, among others.[6]

In this work, we present an approach to map and analyze non-covalent interactions, requiring only molecular geometry information, which compliments existing methods for covalent and electrostatic interactions. Covalent bonds are intuitively represented using conventional Lewis structures.7 They can be visualized from properties of the electron density with modern quantum-mechanical models of bonding, such as the electron localization function (ELF)8›9 and atoms-in-molecules (AIM) theory.10¯12 Also, purely electrostatic interactions can be

Corresponding author: weitao.yang@duke.edu.

Supporting Information Available. DFT results for the s22 set of bi-molecular complexes and carbon-carbon covalent bonds, MP2 and promolecular results for seclected small molecules and complexes. Additional methodology details include comparison of other density-gradient ratios, comparison with AIM and ELF quantum-chemical bonding theories, promolecular atomic-density parameters, complete reference [29], and molecular geometries for selected species. This material is available free of charge via the Internet at http://pubs.acs.org.

analyzed using electrostatic potential maps.13 Non-covalent interactions are frequently visualized using distance-dependent contacts, generally without consideration of hydrogen atoms.14–16 Hydrogen-bonds can be identified from the molecular geometry[17] and from ELF, [18] while grid-based calculations based on classical force fields are used to model other van der Waals interactions.[19] Our approach outlined below, using the density and its derivatives, allows simultaneous analysis and visualization of a wide range of non-covalent interactions types as real-space surfaces and adds an important tool to a chemist's arsenal.

2 Theory

2.1 Background

The quantum-mechanical electron density, ρ, from which all chemical properties can, in principle, be obtained[20] is the key quantity in density functional theory (DFT). The reduced density gradient, coming from the density and its first derivative, ($s=\frac{1}{2(3\pi^2)^{1/3}}\frac{|\nabla\rho|}{\rho^{4/3}}$) is a fundamental dimensionless quantity in DFT used to describe the deviation from a homogeneous electron distribution.20–22 Properties of the reduced gradient have been investigated in depth in the process of developing increasingly accurate functionals.[23] In density tails (i.e. regions far from the molecule, in which the density is decaying to zero exponentially) the reduced gradient will have very large positive values. Conversely, the reduced gradient will assume very small values, approaching zero, for regions of both covalent bonding and non-covalent interactions.

2.2 Identifying Non-Covalent Interactions

To explore the features associated with small reduced gradients, we first examine plots of s versus ρ (Figure 1). These plots were generated by evaluating the B3LYP[24,25] density and reduced gradients on cuboid grids, with a 0.1 au step size, for each molecule or dimer. To provide even more sampling of the small low-density, low-gradient regions in hydrogen-bonded complexes, additional calculations were performed for water and formic acid dimers with a much denser 0.025 au grid.

Plotting s versus ρ, as in Figure 1, reveals the basic pattern of intramolecular interactions. Methane (Fig. 1a) illustrates the typical covalent bond pattern. The top left-side points (small density and large reduced gradient) correspond to the exponentially-decaying tail regions of the density, far from the nuclei. The points on the bottom right side (density values of ca. 0.25 au and low reduced gradient) correspond to the C-H covalent bonds. Covalent bonds have a characteristic saddle point in the electron density (bond critical points10–12), corresponding to s = 0. Regions near the nuclei have larger density values and appear beyond the right edge of the plot. The plot has an overall shape of the form $a\rho^{-1/3}$ because atomic and molecular densities are piecewise exponential. The results for water are very similar, the only difference being that the covalent bonds lie at higher density values, past the edge of the plot. In Figure 1b-d, we consider six examples of chemical systems displaying various types of non-covalent interactions. Plots of s versus ρ for these systems all exhibit a new feature: one or more spikes in the low-density, low-gradient region, a signature of non-covalent interactions. The origin of this feature is made apparent by considering the formation of an intermolecular complex. The predominant change in the density-gradient profile occurs for the low-density region between the two monomers. The reduced gradient changes from very large values in the monomers to near zero upon dimer formation. This is the basis of our approach.

We also explored other ratios of density and gradient values. Indeed, the reduced gradient (regardless of the constant), the fundamental dimensionless variable in DFT, is also found to be the ratio of density and gradient values that most clearly isolates non-covalent interactions in real space (see supplementary information). In order for some ratio of the density and

gradient of the form $|\nabla\rho|/\rho^n$ to be successful, it must distinguish between non-covalent interactions and the exponentially-decaying tails of the density that occur far from the molecular system. Both types of regions are characterized by low densities. In density tails, both the density and gradient approach zero exponentially. In regions of non-covalent interactions, the gradient will again approach zero and will be identically zero at the critical point.

To this point, we have found that non-covalent interactions can be isolated as regions with low density and low reduced gradient. The density values of the low-gradient spikes also appear to be an indicator of the interaction strength. However, very different types of interactions (i.e. hydrogen-bonding and steric crowding) appear in the same region of density/reduced-gradient space. To distinguish between these interactions, we consider second derivatives of the density.

2.3 Identifying Interaction Types

Although localizing low-density, low-gradient regions enables identification of weak interactions in a molecular system, more specific interaction types cannot be determined from the density values alone. Low density regions are obviously related to the weakest interactions, such as van der Waals, whereas those with higher densities will be related to stronger (either stabilizing or de-stabilizing) interactions.[26] Density derivatives can be used to this end.

The sign of the Laplacian of the density, $\nabla^2\rho$, is a widely used tool to distinguish between different types of strong interactions.[27] To understand bonding in more detail, the Laplacian is often decomposed into a sum of contributions along the three principal axes of maximal variation. These components are the three eigenvalues λ_i of the electron-density Hessian (second derivative) matrix, such that $\nabla^2\rho = \lambda_1 + \lambda_2 + \lambda_3$, ($\lambda_1 \leq \lambda_2 \leq \lambda_3$). Analysis of these components has been widely applied to chemical bonding.[11,27,28]

At nuclei (or non-nuclear attractors), the density reaches a local maxima and all three eigenvalues are negative. Interatomic regions between bonded atoms are characterized by the presence of one positive and two negative eigenvalues ($\lambda_1 < 0$; $\lambda_2 < 0$; $\lambda_3 > 0$). In the case of covalent interactions, the negative contributions are dominant and the resultant Laplacian is negative. For weaker, non-covalent interactions, the Laplacian in the interatomic region is dominated by the positive contribution,[27] independently of whether they are bonding or non-bonding. Bonding interactions can be identified by the negative sign of λ_2, as for the hydrogen-bond in the water dimer example. Conversely, if atoms are in non-bonded contact, $\lambda_2 > 0$ in the interatomic region ($\lambda_3 > 0$ and λ_1 can be either positive or negative). An example of this situation occurs in the center of the bicyclo[2,2,2]octene cage. Cases where several atoms interact, but are not bonded, correspond to steric crowding in the context of classical chemistry.

Therefore, we can utilize the sign of λ_2 to distinguish bonded ($\lambda_2 < 0$) from non-bonded ($\lambda_2 > 0$) interactions. Analysis of the sign of λ_2 thus helps to discern between different types of non-covalent interactions, whereas the density itself provides information about their strength. This is illustrated in Figure 2, which shows a modification of our earlier reduced gradient and density plots, such that the ordinate is now sign(λ_2)ρ. The figure shows data for the hydrogen-bonded water dimer, with the low-density, low-gradient spike now lying at negative values indicative of stabilizing interactions. Conversely, the low-density, low-gradient spike for the sterically-crowded bicyclo[2,2,2]octene molecule remains at positive values indicating the lack of bonding in the central area of this moleule. Finally, the low-density, low-gradient spike for the dispersion-bound methane dimer is very near zero, with slightly negative values, indicative of weak attraction.

There have been many previous studies of λ_2 in different bonding situations and attempts to rationalise it in terms of movements (accumulation/depletion) of the density, that is often

understood as due to attractive/repulsive interactions.[11,27,28] For non-covalent interactions, the main features of the electron density and its derivatives appear clearly if the density is constructed from something as simple as a sum of atomic densities (see Section 4.3). Indeed, for all cases considered, results at the self-consistent and promolecular level are qualtiatively equivalent, which rules out any simple connection to accumulation or depletion of the density. However, when the effect of self-consistent calculations on the density and eigenvalues are analyzed, we find that some quantitative differences are introduced by density relaxation. As expected, the density-gradient peaks are shifted to more bonding regimes when comparing promolecular and self-consistent densities. Specifically, a large shift toward smaller density values is observed in the peak corresponding to non-bonded overlap, introducing less repulsion and greater stability.

To summarize, since non-covalent interactions are characterized by low density and reduced gradient values, they can be located by generating gradient isosurfaces enclosing the corresponding regions of real space. The interaction types can be further understood by the values of sign(λ_2)ρ in these regions. These isosurfaces are the basis of our non-covalent interaction method.

3 Computational Details

To obtain plots of the electron density (ρ) and reduced density gradient ($s=\frac{1}{2(3\pi^2)^{1/3}}\frac{|\nabla\rho|}{\rho^{4/3}}$), density-functional theory calculations were performed for a selected set of small molecules and dimers. Calculations on methane, water, branched octane, bicyclo[2,2,2]octene, and the homo-molecular dimers of methane, benzene, water, and formic acid, were performed with the B3LYP functional[24,25] and the 6-31G* basis set, using the Gaussian 03 program.[29] Molecular geometries of methane, water, branched octane, and bicyclo[2,2,2]octene are the same as those used in the G3X procedure.[30] The geometries of the methane, water, and formic acid dimers were obtained from Ref.[31] and the benzene dimer geometry was obtained from Ref.[32] Additional B3LYP/6-31G* calculations were performed on Hobza's set of 22 bi-molecular complexes.[33] Diamond and graphite calculations were performed using the CRYSTAL program,[34] with the BPW91 functional[35,36] and modified 6-21G* basis sets optimized for diamond or graphite.[34]

To generate approximate promolecular densities, fully-numerical, LSDA,[37] free-atomic densities were generated for the atoms H-Ar using the NUMOL program.[38,39] These densities were spherically averaged over space and summed over spins. Because atomic densities are piecewise exponential for each shell of electrons, they were then fit to one (H, He), two (Li-Ne), or three (Na-Ar) Slater-type functions of the form $\rho^{at} = \Sigma_j c_j e^{(-\mathbf{r}/\zeta j)}$, with the c and ζ parameter values given in the supporting information. The use of simple exponential functions to construct the promolecular density also allows first and second derivatives to be obtained analytically.

To demonstrate our method for the interaction between a ligand and a protein active site, we apply it to a complex of a bacterial regulatory protein of the tetR family with a tetracycline inhibitor.[40] The geometry was obtained from the protein data bank[41] (pdb file 2UXO, only chain B). All crystal waters were included in the calculations. Protonation states were determined by WhatIf.[42] Hydrogens were added using the hbuild function in CHARMM.[43] The positions of the protein and water hydrogen atoms were optimized, followed by the positions of the ligand hydrogen atoms. In both these geometry optimizations, the positions of all other atoms were frozen. The protein was described with the CHARMM27 force field, the water molecules were described with TIP3 force field, and the ligand was described with the PM3 semi-empirical Hamiltonian. Since the exponential atomic densities die off fairly quickly, only atoms lying within 6 Å of the tetracycline inhibitor were used in calculation of the

promolecular density and reduced gradient. For simplicity, all intramolecular interactions within the protein and ligand, as well as intermolecular interactions between the protein and water molecules, were omitted.

4 Results and Discussion

4.1 Small Molecules, Complexes, and Solids

Figure 3 displays low-gradient ($s = 0.5$ au) isosurfaces, subject to the constraint of low density, for branched octane, bicyclo[2,2,2]octene, and the homomolecular dimers of methane, benzene, water, and formic acid. A density cut-off of $\rho < 0.05$ au was chosen since it encapsulates the non-covalent interaction region of interest (see Fig. 1). Other isovalues of the reduced gradient could be used as a diagnostic of non-covalent interactions. The choice of cut-off can be made objective by representing all points within the s versus ρ peaks (see Figs. 1 and 2). However, this more complicated algorithm is not necessary, since any cut-off that fulfills the requirements of lying within the peak values already provides the interaction information. Appropriate bounds on the density and reduced gradient should be selected to isolate the low-density, low-gradient peaks for each chemical system of interest. The locations of these peaks are sensitive to the atom types involved, the interaction strengths, and the level of theory used to obtain the density (self-consistent versus promolecular densities).

The gradient isosurfaces are colored according to the corresponding values of $\text{sign}(\lambda_2)\rho$, which is found to be a good indicator of interaction strength. Large, negative values of $\text{sign}(\lambda_2)\rho$ are indicative of attractive interactions (such as dipole-dipole or hydrogen bonding); while if $\text{sign}(\lambda_2)\rho$ is large and positive, the interaction is non-bonding. Values near zero indicate very weak, van der Waals interactions.

The gradient isosurfaces provide a rich visualization of non-covalent interactions as broad areas of real space, rather than simple pair-wise contacts between atoms. We first consider the sterically-crowded molecules bicyclo[2,2,2]octene (Fig. 3a) and the branched octane isomer (Fig. 3b). In the first case, the low-density, low-gradient region corresponds to the center of the cage, where steric repulsion between the bridgehead carbons is expected. For the branched octane isomer, the isosurface lies between the closely-interacting methyl groups on opposite sides of the central C-C bond. The interactions are repulsive nearer the C-C bond and weakly attractive between the hydrogen atoms. Dispersion and hydrogen bonding can also be clearly detected. In the benzene dimer (Fig. 3c), there is an area of non-bonded overlap located at the center of each benzene ring, resembling the isosurface for bicyclo[2,2,2]octene. There is another lower-density surface between the overlapping portions of the benzenes, where π-stacking is expected. In the dispersion-bound methane dimer (Fig. 3d), the isosurface forms a disc between the individual monomers. For the water dimer (Fig. 3e), the isosurface lies between a hydrogen donor and oxygen acceptor, characteristic of H-bonding. Finally, the formic acid dimer (Fig. 3f) reveals stronger H-bonds than in the water dimer, and also weak van der Waals interactions between the two closely-interacting acidic hydrogens.

Crystalline solids exhibit rich and challenging bonding patterns. We consider the prototypical examples of carbon in the diamond (Fig. 3g) and graphite (Fig. 3h) phases at their equilibrium geometries. In diamond, the carbon atoms are sp^3 hybridized and are connected by strong covalent bonds that form a tridimensional, tetrahedral network. Fig. 3g shows a low-gradient, low-density isosurface for a cuboid section of the diamond crystal. The non-covalent surface extends through the voids of the structure, creating a network similar to that of the covalent bonds. Graphite in its α form (Fig. 3h) has a bidimensional, hexagonal lattice, with the carbon atoms sp^2 hybridized and covalently bonded to their three nearest neighbors. The low-gradient, low-density isosurface shows areas of non-bonded overlap at the center of the hexagonal rings,

as seen previously in benzene. $\pi - \pi$ stacking interactions between the graphene sheets are clearly manifested by the isosurfaces filling the interlayer spaces.

The use of DFT densities to visualize non-covalent interactions, including dispersion, may seem counter-intuitive. In the case of dispersion, predicted binding energies are highly sensitive to the choice of method. It is well known that most density-functional theory methods, including B3LYP, do not account for dispersion binding[44,45] and in fact predict no binding in the benzene and methane dimer cases. However, one of the main advantages of our approach is that it is based on properties of the electron density, the dominant features of which are quite insensitive to the choice of electronic structure method used in the computations, particularly for very weak interactions. As a side note, the main problem with DFT for dispersion is not in computing the density, but in extracting an accurate energy. Different computational methods give almost the same densities, but very different energetics. As two monomers form a dispersion-bound complex, the changes in density caused by the dispersion interaction are quite subtle. They vary depending on whether the method predicts dispersion binding or repulsion.[46] Nevertheless, they are quite small in magnitude in the inter-monomer region.

To test the validity of B3LYP densities for non-covalent interactions, we repeated our calculations for the test set of six small molecules and dimers with MP2/6-311++G**, which does predict dispersion binding. The MP2 results, with identical cut-offs, are given in the supporting information and are virtually indistinguishable from our B3LYP/6-31G* isosurfaces in Figure 3. Due to the reduced computational cost, DFT calculations with small basis sets are recommended over MP2 for practical visualization of non-covalent interactions.

To summarize, plotting low-gradient isosurfaces, subject to a further low-density constraint, allows real-space visualization of non-covalent interactions. In addition to the small test set of molecules considered above, the method was successfully applied to Hobza's set of 22 bi-molecular complexes;[33] these results are given in the supporting information. Our model discriminates between different types of weak interactions and allows for simultaneous visualization of all of them. Classification of the various non-covalent interaction types agrees with general chemical concepts. Surfaces with very low density values (i.e. $\rho < 0.005$ au) generally map to weaker dispersion interactions. Surfaces with slightly higher density values (i.e. $0.005 < \rho < 0.05$ au) map to stronger non-covalent interactions, including both attractive H-bonding (negative λ_2) and steric clashes (positive λ_2).

Although not the focus of this paper, our method is also applicable to covalent bonds, which obviously appear at much larger density values than those considered above. This is possible since all covalent bonds have a saddle point in the density, where $s = 0$. The low-gradient isosurfaces are also distinctly different for single, double, and triple bonds; results are shown for the ethane, ethylene, acetylene series in supporting information.

This also highlights a limitation of our method. It assumes an equilibrium or near-equilibrium geometry, which is broadly applicable to most nanoscale or biomolecular systems. As molecular geometries are compressed, the density at which the reduced gradient approaches zero will shift to higher value, approaching that of covalent bonds. Also, the second eigenvalue remains negative, despite the repulsive nature of the interaction at compressed separations. This shortcoming has already been observed for other density-based bonding analyses, such as Atoms in Molecules.[28]

4.2 Comparison with Quantum-Chemical Bonding Theories

Our method compliments existing quantum-chemical theories of bonding such as ELF and AIM. The electron localization function (ELF)[8,9] allows visualization of covalent bonds, electron lone pairs, and atomic shell structure from properties of the density. A comparison

between isosurfaces generated using our approach and ELF is shown for the benzene dimer in the supporting information. With a 0.9 au isosurface, ELF clearly shows the core electrons of the carbon atoms, the aromatic C-C bonds, and the electrons localized on the hydrogen atoms. While the ELF surface does not show any interaction between the benzenes, the ELF values at critical points have been used to construct tree-diagrams that reflect weak binding in intermolecular complexes.[18]

The AIM molecular graph showing critical points and bond paths, obtained using the AIM 2000 program[47] is also shown in the supporting information. Note that the dispersion interaction appears as two bonds. AIM predicts low-density critical points and bond paths connecting the benzene rings via the two closest pairs of carbon atoms, capturing the π-stacking interaction between the two benzenes. Similar plots are shown in the work of Waller *et al.*[48] However, AIM is not an ideal method for visualization of non-covalent interactions, particularly for large chemical systems where the calculations become prohibitively expensive. Moreover, AIM depicts both covalent and non-covalent interactions as local, pair-wise contacts along the bond paths;[48] in large systems the vast number of AIM bond paths obscures non-local van der Waals interactions. Conversely, our approach manifests non-covalent interactions as broad surfaces between entire functional groups. This allows a much clearer and more intuitive visualization of interactions in complex systems.

4.3 Biological Systems

Biological macromolecules, such as proteins or DNA, are probably the most relevant systems for a visual representation of non-covalent interactions, which are the main driving forces in biological processes. Understanding of these interactions is crucial for the comprehension of their three-dimensional structure and biological activity required for drug design. However, calculation of the electron densities of these systems becomes extremely computationally expensive. A more effcient way to obtain s and ρ is required to display non-covalent interactions. It is well-known[49,50] that important topological features of the electron density are contained in the sum of atomic densities, termed the promolecular density, $\rho^{pro} = \sum_i \rho_i^{at}$. A promolecular density obtained from simple exponential atomic pieces is able to predict low-density, low-reduced-gradient regions similar to density-functional results. The free atomic densities used in these calculations consist of one Slater-type function for each electron shell, fit to closely reproduce spherically-averaged, density-functional atomic densities.

Approximate promolecular densities were constructed by summing exponential atomic densities for branched octane, bicyclo[2,2,2]octene, and the homomolecular dimers of methane, benzene, water, and formic acid. Resultant plots of s^{pro} versus ρ^{pro} for these species show the same features seen in Fig. 1. Also, gradient isosurfaces generated from the promolecular density (see supporting information) are very similar to those obtained previously with self-consistent DFT and even MP2 densities (Fig. 3). Thus, our approach can be extended to very large systems by using an approximate promolecular density, from which the relevant isosurfaces are plotted. This allows rapid characterization of non-covalent interactions for biomolecules, with only the molecular geometry required as input.

For all cases considered, results at the self-consistent and promolecular level are qualitatively equivalent. Some quantitative differences are introduced in the promolecular calculations that, as expected, shift the s versus ρ peaks to more bonding regimes. This is illustrated in Fig. 2, which plots the reduced gradient versus sign(λ_2)ρ for both SCF and promolecular densities. The promolecular approach is expected to perform best for very weak or moderately attractive interactions, where there is minimal response of the density to the interaction. Because the density cannot adjust, the promolecular approximation gives the largest differences from SCF results for non-bonding interactions, as shown in Fig. 2. This explains why the largest

differences between SCF and promolecular isosurfaces appear at the center of rings, such as bicyclo[2,2,2]octene and benzene.

We first consider two model polypeptides: an α-helix consisting of 15 alanine residues and an anti-parallel β-sheet consisting of 17 glycine residues. Geometries of the polypeptides were obtained with the MMFF force field using the spartan program.[51] Both were capped with $COCH_3$ and $NHCH_3$ groups. Figure 4 displays low-gradient isosurfaces for cuboid regions at the center of these polypeptides, colored according to the corresponding density values. For the β-sheet, the lowest-density portions of the gradient isosurface arise from hydrophobic, dispersion-dominated interactions, primarily involving the CH_2 groups of the glycines. The higher-density regions correspond to inter-residue hydrogen-bonds and repulsive interactions between the adjacent C=O and N-H groups. For the α-helix, the isosurface has a large, low-density region within the helix and between the side-chain methyl groups. The higher-density portions of the isosurface correspond to inter-residue hydrogen-bonds along the helix and repulsive interactions between adjacent N-H groups.

We also considered the non-covalent interactions between nucleobases in the B-form of double-strand, six-base-pair (TGTGTG) DNA. The structure was obtained using the X3DNA program[52] with ideal geometry parameters.53 Figure 4c displays the low-gradient isosurface for a cuboid section in the center of the DNA helix, colored according to the sign(λ_2)ρ values. The calculated isosurface resembles that of graphite, with broad, low-density regions indicative of π-stacking between base-steps. The interactions between individual deoxyadenosine-deoxythymidine and deoxycytidine-deoxyguanosine pairs are shown in Fig. 4d,e. The isosurfaces show non-bonded overlap within the nucleobase rings, as in benzene and graphite, and hydrogen-bonding motifs similar to the formic acid dimer. The strong N-H – O and N-H – N hydrogen bonds can be clearly distinguished from the weaker, attractive C-H – O interaction by the density values, as shown in different colors.

The hydrogen-bonding surfaces in the DNA model have density values of ca. 0.065 au, compared to density values of ca. 0.035 au for the polypeptide hydrogen bonds. This is evident from the degree of blue shading for the hydrogen bonds in Fig. 4. Since density values at hydrogen-bond critical points correlate with the interaction strength,[12,54] our results indicate that the hydrogen bonds between nucleobase pairs are substantially stronger than between amino acids, in agreement with literature data.[55–57]

Let us now consider the interaction between a ligand and a protein active site. The low-gradient isosurface for a tetracycline inhibitor bound to the tetR protein in Fig. 5 shows a complex web of non-covalent interactions between the ligand and active site. Several other protein-ligands pairs have been tested, leading to the same conclusions. When analyzing non-covalent interactions in protein-ligand complexes, it is usually assumed that these interactions are due to a specific contact between two atoms.[16] However, it is clearly seen in Fig. 5 that this assumption is only partly correct. Hydrogen bonds, such as those between the tetracycline amine groups and two water molecules (shown in orange), are directional and specific. Conversely, van der Waals, dipole-dipole, and hydrophobic interactions, such as those between the tetracycline and the Leu61, Val91, Ile136, and Val166 residues (shown in yellow), are not atom-specific and occupy broader regions in space. The figure reveals some steric clashes (orange and red regions of the isosurface) that must be offset by stronger, attractive interactions to give binding in this crystal structure. A ligand “fits” the geometry of the active site, and the interaction energy between the ligand and protein is comprised of many small contributions. When trying to design a new ligand to fit a specific active site, one should consider all such interactions.

5 Summary

In conclusion, non-covalent interactions have a unique signature and their presence can be revealed solely from the electron density. Non-covalent interactions are highly non-local and manifest in real space as low-gradient isosurfaces with low densities. The sign of the second Hessian eigenvalue is used to give the type of interaction, and its strength can be derived from the density on the non-covalent interaction surface. This approach provides a rapid and rich representation of van der Waals interactions, hydrogen-bonds, and steric clashes, requiring only the atomic coordinates as input. This tool offers exciting possibilities for aiding design of ligands, self-assembled materials, catalysts, and other molecular systems. Of specific interest would be analysis of the interaction between two large bio-molecules, such as protein-protein and protein-DNA interfaces.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work is supported by the Natural Sciences and Engineering Research Council of Canada (ERJ), the National Science Foundation, and the National Institute of Health (WY). JCG thanks the Spanish MALTA-Consolider Ingenio-2010 program under project CSD2007-00045.

References

1. Fenniri H, Packiarajan M, Vidale KL, Sherman DM, Hallenga K, Wood KV, Stowell JG. J Am Chem Soc 2001;123:3854–3855. [PubMed: 11457132]
2. Kruse P, Johnson ER, DiLabio GA, Wolkow RA. Nano Lett 2002;2:807–810.
3. Sheiko SS, Sun FC, Randall A, Shirvanyants D, Rubinstein M, Lee H, Matyjaszewski K. Nature 2006;440:191–194. [PubMed: 16525468]
4. DiLabio GA, Piva PG, Kruse P, Wolkow RA. J Am Chem Soc 2004;126:16048–16050. [PubMed: 15584738]
5. Kollman PA. Chem Rev 1977;10:365–371.
6. Keinan S, Ratner MA, Marks TJ. Chem Phys Lett 2004;392:291–296.
7. Lewis GN. J Am Chem Soc 1916;38:762–785.
8. Becke AD, Edgecombe KE. J Chem Phys 1990;92:5397–5403.
9. Silvi B, Savin A. Nature 1994;371:683–686.
10. Bader RFW. Chem Rev 1991;91:893–928.
11. Bader, RFW. International Series of Monographs on Chemistry. Vol. 22. Oxford Science Publications; Oxford: 1990. Atoms in Molecules: A Quantum Theory.
12. Matta, CF.; Boyd, RJ. The quantum theory of atoms in molecules. Matta, CF.; Boyd, RJ., editors. Wiley-VCH; 2007. p. 1-34.
13. Honig B, Nicholls A. Science 1995;268:1144–1149. [PubMed: 7761829]
14. Word JM, Lovell SC, LaBean TH, Taylor HC, Zalis ME, Presley BK, Richardson JS, Richardson DC. J Mol Biol 1999;285:1711–1733. [PubMed: 9917407]
15. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB III, Snoeyink J, Richardson JS, Richardson DC. Nuc Acids Res 2007;35:W375–W383.
16. Sobolev V, Sorokine A, Prilusky J, Abola EE, Edelman M. Bioinformatics 1999;15:327–332. [PubMed: 10320401]
17. McDonald IK, Thornton JM. J Mol Biol 1994;238:777–793. [PubMed: 8182748]
18. Alikhani ME, Fuster F, Silvi B. Struct Chem 2005;16:203–210.
19. Cramer RD III, Patterson DE, Bunce JD. J Am Chem Soc 1988;110:5959–5967.
20. Hohenberg P, Kohn W. Phys Rev B 1964;136:864–871.

21. Becke, AD. Modern Electronic Structure Theory. Yarkony, DR., editor. World Scientific; 1995. p. 1022-1046.
22. Cohen AJ, Mori-Sánchez P, Yang W. Science 2008;321:792–794. [PubMed: 18687952]
23. Zupan A, Burke K, Ernzerhof M, Perdew JP. J Chem Phys 1997;106:10184–10193.
24. Becke AD. J Chem Phys 1993;98:5648–5652.
25. Lee C, Yang W, Parr RG. Phys Rev B 1988;37:785–789.
26. Gibbs GV, Cox DF, Rosso KM. J Phys Chem A 2004;108:7643–7645.
27. Bader RFW, Essén H. J Chem Phys 1984;80:1943–1960.
28. Bader RFW. J Phys Chem A 1998;102:7314–7323.
29. Frisch, MJ., et al. Gaussian 03, Revision C.02. Gaussian, Inc; Wallingford, CT: 2004.
30. Curtiss LA, Redfern PC, Raghavachari K, Pople JA. J Chem Phys 2001;114:108–117.
31. Zhao Y, Truhlar DG. J Chem Theory Comput 2005;1:415–432.
32. Sinnokrot MO, Sherrill CD. J Phys Chem A 2004;108:10200–10207.
33. Jurečka P, Šponer J, Černý J, Hobza P. Phys Chem Chem Phys 2006;8:1985–1993. [PubMed: 16633685]
34. Saunders, VR.; Dovesi, R.; Roetti, C.; Causá, M.; Harrison, NM.; Orlando, R.; Zicovich-Wilson, CM. CRYSTAL98 User's Manual. University of Torino; Torino (Italy): 1998.
35. Becke AD. Phys Rev A 1988;38:3098–3100. [PubMed: 9900728]
36. Perdew, JP. Electronic Structure of Solids. Ziesche, P.; Eschrig, H., editors. Akademie Verlag; Berlin: 1991. p. 11
37. Perdew JP, Wang Y. Phys Rev B 1992;45:13244–13249.
38. Becke AD. Int J Quantum Chem Symp 1989;23:599–609.
39. Becke AD, Dickson RM. J Chem Phys 1990;92:3610–3612.
40. Alguel Y, Meng C, Teran W, Krell T, Ramos JL, Gallegos MT, Zhang X. J Mol Biol 2007;369:829–840. [PubMed: 17466326]
41. http://www.rcsb.org
42. Vriend G. J Mol Graph 1990;8:52–56. [PubMed: 2268628]
43. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. J Comp Chem 1983;4:187–217.
44. Zhang Y, Pan W, Yang W. J Chem Phys 1997;107:7921–7925.
45. Johnson ER, Wolkow RA, DiLabio GA. Chem Phys Lett 2004;394:334–338.
46. Aleen MJ, Tozer DJ. J Chem Phys 2002;117:11113–11120.
47. Biegler-König, F.; Schönbohm, J. AIM 2000 version 2.0. University of Applied Science; Bielefeld, Germany: 2002.
48. Waller MP, Robertazzi A, Platts JA, Hibbs DE, Williams PA. J Comput Chem 2006;27:491–504. [PubMed: 16444702]
49. Spackman MA, Maslen EN. J Phys Chem 1986;90:2020–2027.
50. Pendás AM, Luaña V, Pueyo L, Francisco E, Mori-Sánchez P. J Chem Phys 2002;117:1017–1023.
51. Deppmeier, BJ.; Driessen, AJ.; Hehre, W.; Johnson, JA.; Klunzinger, PE.; Watanabe, M. Spartan ES 1.0.2. Wavefunction Inc; Irvine, CA: 2002.
52. Lu XJ, Olson WK. Nucleic Acids Res 2003;31:5108–5121. [PubMed: 12930962]
53. http://rutchem.rutgers.edu/~olson/Tsukuba/
54. Espinosa E, Molins E, Lecomte C. Chem Phys Lett 1988;285:170–173.
55. Wieczorek R, Dannenberg JJ. J Am Chem Soc 2004;126:14198–14205. [PubMed: 15506786]
56. Viswanathan R, Asensio A, Dannenberg JJ. J Phys Chem A 2004;108:9205–9212.
57. Jurecka P, Hobza P. J Am Chem Soc 2003;125:15608–15613. [PubMed: 14664608]

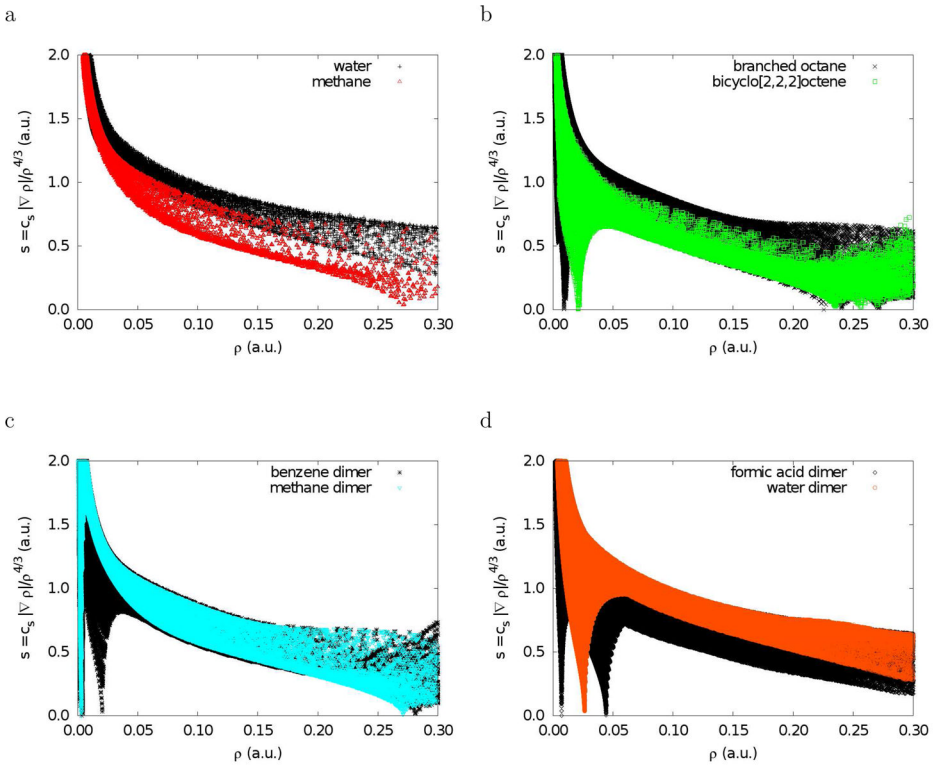
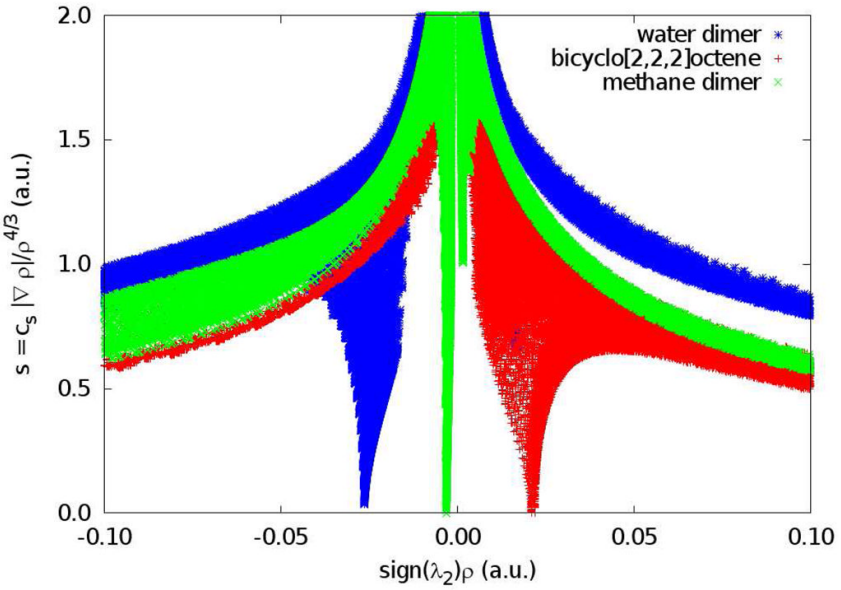


Figure 1.

Plots of the electron density and its reduced gradient for methane, water, branched octane, bicyclo[2,2,2]octene, and the homomolecular dimers of methane, benzene, water, and formic acid. The data was obtained by evaluating B3LYP/631G* density and gradient values on cuboid grids.

a



b

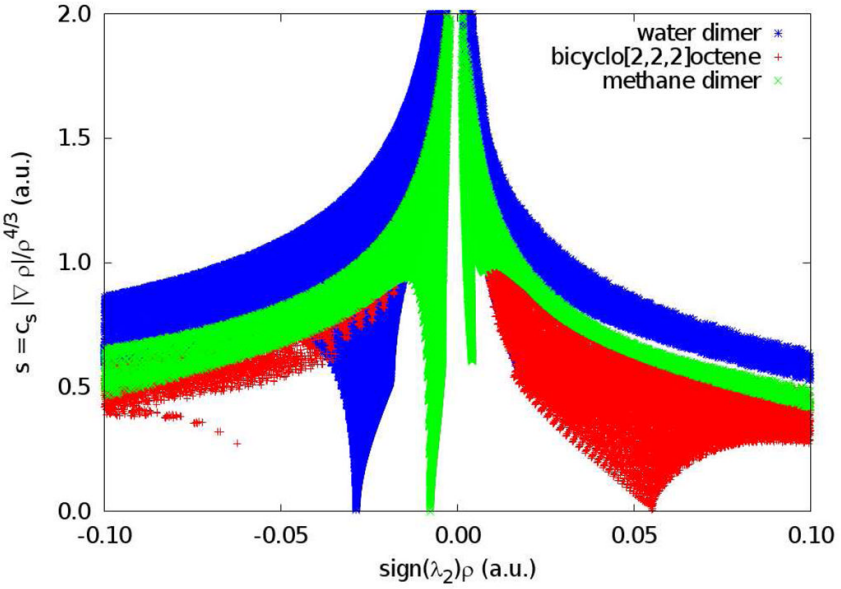


Figure 2.

Plots of the reduced density gradient versus the electron density mulitplied by the sign of the second Hessian eigenvalue. Results are shown for bicyclo[2,2,2]octene, methane dimer, and water dimer. The data was obtained by evaluating B3LYP/631G* density (a) or promolecular density (b) and gradient values on cuboid grids.

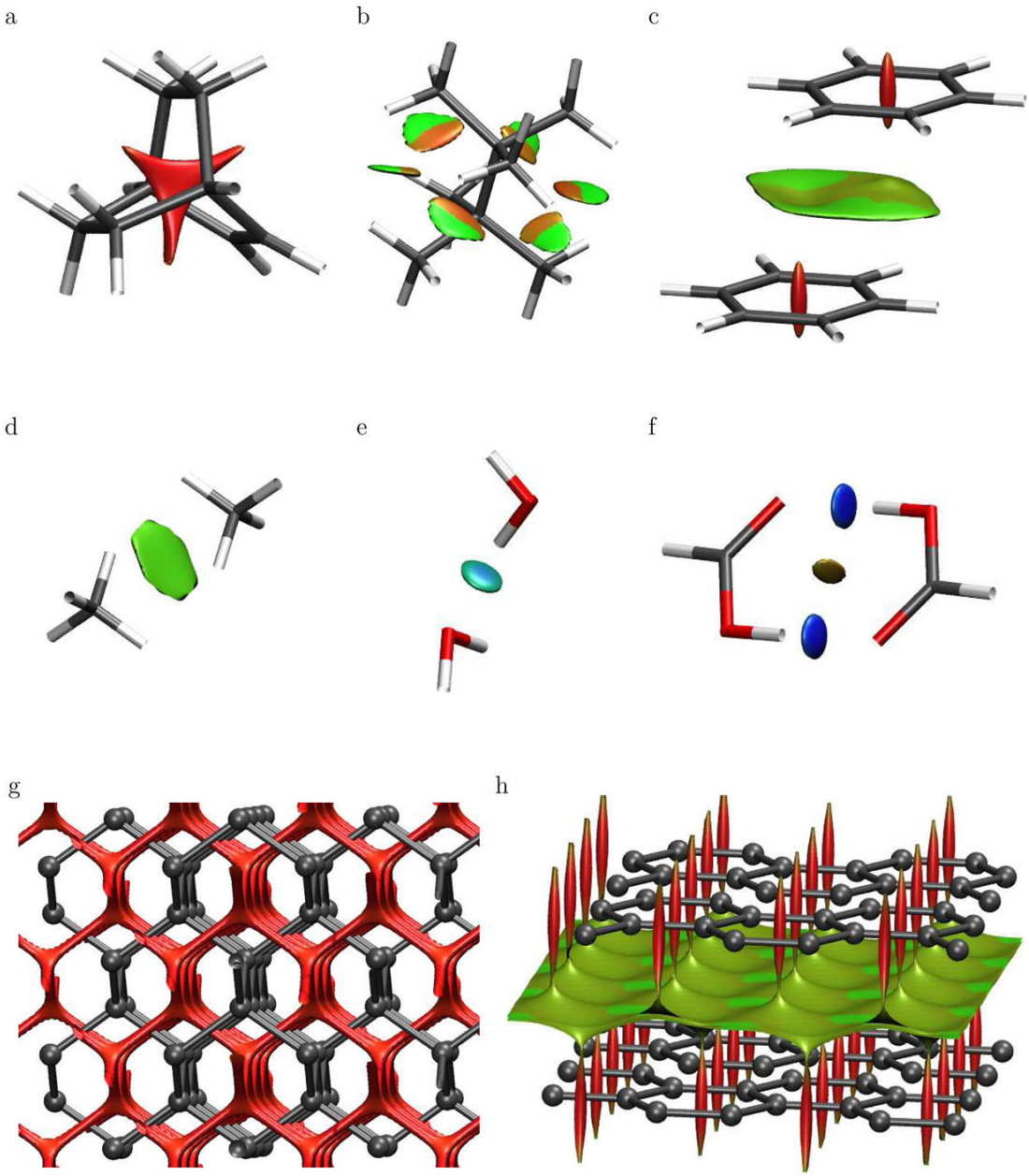


Figure 3.

Gradient isosurfaces (s = 0.5 au) for (a) bicyclo[2,2,2]octene, (b) branched octane, and the homomolecular dimers of (c) benzene, (d) methane, (e) water, and (f) formic acid. Gradient isosurfaces are also shown for cuboid sections of (g) diamond and (h) graphite. The surfaces are colored on a blue-green-red scale according to values of sign(λ_2)ρ, ranging from −0.04 to 0.02 au. Blue indicates strong attractive interactions and red indicates strong non-bonded overlap.

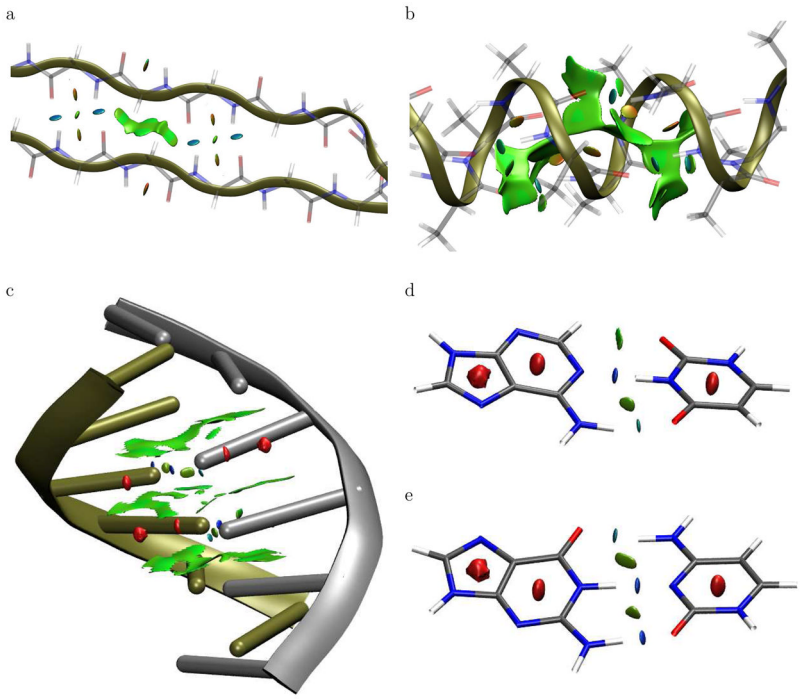


Figure 4.

Gradient isosurfaces (s^{pro} = 0.35) for cuboid sections of the (a) β-sheet and (b) α-helix polypeptides. Gradient isosurfaces (s^{pro} = 0.25) are also shown for the (c) B-form of DNA, and the (d) A-T and (e) C-G base pairs. The surfaces are colored on a blue-green-red scale according to values of sign(λ_2) ρ, ranging from −0.06 to 0.05 au. Blue indicates strong attractive interactions and red indicates strong non-bonded overlap.

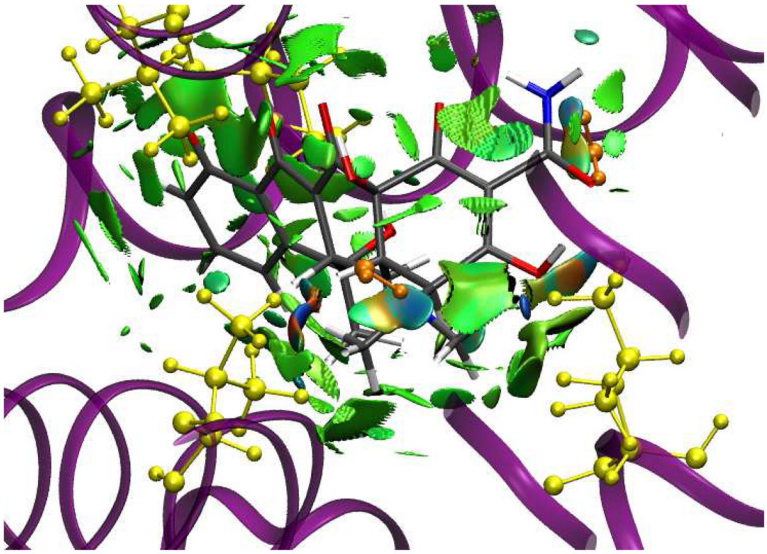


Figure 5.

Gradient isosurfaces (s^{pro} = 0.35) for interaction between the tetR protein and tetracy-cline inhibitor. The surfaces are colored on a blue-green-red scale according to values of sign(λ_2) ρ, ranging from −0.06 to 0.05 au. Blue indicates strong attractive interactions and red indicates strong non-bonded overlap.