Characterizing the Strength of Individual Hydrogen Bonds in DNA Base Pairs

Halina Szatyłowicz*,† and Nina Sadlej-Sosnowska‡

Faculty of Chemistry, Warsaw University of Technology, Noakowskiego 3, 00-664 Warsaw (Poland) and National Medicines Institute, 30/34 Chełmska Street, 00-725 Warsaw, Poland

Received July 26, 2010

The energies of individual hydrogen bonds (H-bonds) in A-T and G-C Watson−Crick base pairs were calculated according to the natural bond orbital (NBO) analysis of intermolecular interactions. The extent to which individual H-bonds are helpful in holding the two base pairs together was previously investigated quantitatively by a few different approaches, and the results of the present and previous estimations were compared. The method was validated by the determination of the H-bond strength changes in A-T and G-C pairs upon the substitution of the monomer (base) by two cationic substituents; the systems for which the changes were previously anticipated based on the modifications of the H-bonds' distances.

INTRODUCTION

Considerable theoretical attention has been focused on exploring the nature and the strength of hydrogen-bonding (H-bonding) interactions,[1,2] especially in molecules of biological significance, such as DNA and RNA.[3] H-bonding and other noncovalent interactions play a key role in the structure and the dynamics of the molecules, and of their various complexes, with endo- and egzogenic substances as well.[4,5] DNA base pairs contain multiple H-bonds between the bases of one pair of molecules. Therefore, a question arises: What is the contribution of the individual X−H···Y bonds, where X and Y refer to heavy atoms of the H-bond, to the total energy of the H-bonding in a system? To disentangle the contributions, several different philosophies have been applied.

The first approach is the evaluation of a particular H-bond strength in the absence of other intermolecular interactions.[6] To achieve this aim, one of the bases in a pair was rotated with respect to the other around the axis of each H-bond in turn, so that the planes of the individual bases become perpendicular to each other. The geometries of the systems were optimized with certain constraints: (i) the planarity of each of the bases and (ii) in the model structures (rotated species), each base was kept perpendicular to the other one with the X−H···Y angle fixed at its value in the optimized planar base pair. The difference between the sum of the individual H-bond energies and the total interaction energy of the base pairs was taken as the cooperative part of the H-bonding interactions. The calculations were made at the MP2/D95** and B3LYP/D95** levels, and the results (energies of the particular bonds) were similar at both.

The second approach was based on the calculations of inter-residue compliance constants for all possible X−H···Y contacts. The compliance constants can be used to describe inter-residual forces and to provide a unique bond strength.[7] Calculations were made at the B3LYP/6-31(d), B3LYP/6-311++G(d,p), and MP2/6-31(d,p) levels. The ordering of H-bonds from the strongest to the weakest was found the same at all applied computational levels; however, the method did not provide any explicit values of the individual bond energies.

The third method, proposed for the estimation of the individual H-bond strength in the G-C and A-T base pairs, was an atom replacement procedure.[8] To evaluate the strength of the X−H···Y interaction, the X−H group was replaced with its isoelectronic atom G, thereby suppressing the particular H-bonding and estimating the strength of the remaining H-bonds. The repulsion interaction of the close G···Y contact was evaluated by using the corresponding model system. Thus, taking into account the above-mentioned evaluations and the total interaction energy, the strength of the suppressed H-bond was estimated. The philosophy of this method, as well as of the first method mentioned above,[6] was to switch off the orbital interactions connected with other H-bonds and actually deal with it. It should be noted that this method does not take into account deformation energy[9] (also known as preparation energy),[10,11] that is the amount of energy required to deform the separate monomers (bases) from their equilibrium structures to the geometry that they acquire in the base pair.

The problem has also been attacked based on AIM theory.[12] These investigations followed the finding of Espinosa, Molins, and Lecomte[13] on the grounds of data for 83 experimentally studied H-bonds (X−H···O, X = C, N, and O; X-ray diffractions experiments). They reported a monotonic, exponential dependence between the H-bond dissociation energy, $D_e = -E_{HB}$, and the kinetic and potential energy density at the bond critical point, $G(r_{cp})$ and $V(r_{cp})$, respectively, on one hand, and the H-bond length, on the other. The relationships were derived from the calculated H-bond energies and the experimental electron density distributions. Moreover, for the H···O H-bonds a proportionality between the H-bond energy and the potential energy density at the critical point of the H-bond, $E_{HB} = 1/2\ V(r_{cp})$, was proposed. This relationship, named the EML equation (approximation), has been used to estimate the individual

* Corresponding author. E-mail: halina@ch.pw.edu.pl. Telephone: (+48) 22 234 7755.
† Warsaw University of Technology.
‡ National Medicines Institute.

10.1021/ci100288h
Published on Web 11/19/2010

H-bond energy in G-C and A-T pairs, but this evaluation was limited to H···O bonds.[14] The calculated electron density at the H-bond critical point, ρ_{CP}, has also been used for this purpose.[15] In this case, the relationship between the dissociation energy of the individual intermolecular interactions and the ρ_{CP} of the H-bonds, obtained for symmetrical complexes with two identical N−H···O H-bonds, was utilized to determine the strength of all H-bonds.

The above presented approaches will be named as the rotation, compliance constants, atom replacement, EML equation, and $E_{H\cdots B}$ vs ρ_{CP} relationship methods. The comparison of their results reveals that they substantially differ (Table 1). Therefore, we sought yet another method to evaluate the H-bonds' energies in the same base pairs, hoping that our results would reconcile with those obtained by one of the methods presented up to now. Our option was to apply the natural bond orbital (NBO) theory[16] because it deals with molecular interactions in general and with H-bonding in particular based on the charge-transfer orbital interactions.

Within the NBO theory,[16] the energy of a chemical system can be divided into two parts: one associated with the localized orbitals (E_{Lewis} or E_L) and a second originating from noncovalent contributions ($E_{nonLewis}$ or E_{nL}). Most parts of the electron density (typically more than 99%) of a single molecule or a couple of molecules bound by weak interactions are confined to the localized orbitals, which are in good agreement with the Lewis structure concepts, e.g., the one ("lone pair") and two ("bond") center ones. They are well adapted to describe the covalency effects in molecules. However, the transformation of ab initio wave functions to the NBO form leads to orbitals that are unoccupied in the formal Lewis structure and might thereby be used to describe noncovalency effects.[16] These delocalizations are routinely calculated by the perturbation theory and presented as a transfer of electron density between pairs of orbitals, the first "filled" bonding orbital (donor NBO, Lewis type) and the second "empty" antibonding or Rydberg (acceptor NBO, nonLewis type). The energy lowering resulting from all individual transfers are displayed by the NBO program.

As far as a pair of molecules is concerned, for example, connected by a H-bond, delocalization energy is grouped into contributions within both individual molecules and those resulting from delocalization from one unit to another. The bonding intermolecular interactions are sought in the latter group. The NBO analysis points to the origin of H-bonding in the charge-transfer interactions of $n \rightarrow \sigma^*$ type, involving a weak intermolecular delocalization from a lone pair (n) of the donor monomer into the proximate unfilled antibonding orbital σ^* of the acceptor monomer.[17] Such an intermolecular delocalization corresponds to a partial charge transfer (CT). It has been stated that correlations exist between the quantity of charge (q) transferred, the binding energy, and the degree of interpenetration of the van der Waals shells of each monomer.[16,17] Moreover, it has also been found that seemingly small values of the transferred charge (0.001−0.01 e) are associated with energy stabilizations of chemical significance.[17]

Contribution to the energy of CT interactions, E_{CT}, from individual orbital interactions can be estimated from the NBO Fock matrix by the second-order perturbation theory.[18] These second-order estimates, which are routinely generated by the NBO program, show that the dominant contribution to E_{CT} arises from a single $n \rightarrow \sigma^*$ interaction, in each H-bond of a complex. The interaction is enhanced by such geometry of the interacting species, which maximizes the intermolecular overlap of the n_Y lone pair and the acceptor antibond σ_{XH}^* orbitals. The idea of the method based on the NBO theory, as of that based on the compliance matrix, is to evaluate simultaneously all H-bonds in a molecular complex without switching some.

The NBO theory was perhaps the first to explain the phenomenon of H-bonding in the CT language. Yet, there are other reports where similar conclusions were drawn based on a different methodology. Here, the investigations of Bickelhaupt and collaborators have to be cited,[19,20] where the intermolecular interactions in base pairs were analyzed by applying different wave functions and energy decomposition analysis (EDA). The wave functions were obtained by the Kohn−Sham method in the framework of density functional theory (DFT); the interaction energy was decomposed into the classical electrostatic interaction, exchange (or Pauli repulsion), and attractive orbital interaction, and the calculations were performed using the Amsterdam density functional (ADF) program.[21] In a series of papers devoted to the nucleic bases structure, energetics, and interactions, it was shown that the donor−acceptor orbital interaction between a N or an O lone pair orbital on one base and a N−H antibonding σ^* orbital on the other base provides a substantial bonding contribution, which is of the same order of magnitude as the electrostatic interaction term. According to the authors, these two interactions are vital to the behavior and the stability and thereby evolution of the genetic code.

Apart from the bonding delocalizations, the proximity of atoms belonging to two molecules brings about steric repulsions between them.[16] The energy of the steric repulsion, as well as classical electrostatic effects, is included in E_L of a H-bonded system. Similar to bonding delocalizations, the repulsive interactions are also separately decomposed into those belonging to contributions from the electron pairs inside the separate units and those from the interactions of electron pairs belonging to separate noncovalently bound units. The inter-residue repulsive interactions also contribute to the energy of the H-bonding. The most important are the steric repulsions between filled n_Y and σ_{XH} orbitals of the X−H···Y interaction. The balance between steric and donor−acceptor forces is not fully equivalent to the binding energy of the H-bonded system. As a consequence of the interaction between two molecules in near proximity, the charges of all atoms differ from their values for the same isolated molecules. It ought to be emphasized that the perturbation of the electron density takes place not only in the immediate vicinity of the H and X atoms but within the entire complex.[1,22,23] Thus, all atoms and bonds of the whole complex take part in the interactions. This is illustrated by maps showing the regions of increased and decreased electron densities. They can be distinctly seen in Figures 1−6 of ref 22, depicting the alternating shifts of electron density occurring in the complexes of the acceptor (H_2O) and a number of donor molecules. Consequently, there are other contributions to the energy that might be attributed to residual classical electrostatic interaction, in the sense used in ref 16.

In the molecules possessing single H-bonds, the latter interaction might have a higher energy than that of the

interactions, in which only the X−H···Y atoms, directly involved in the H-bond, participate.[17] In these connections, for our base pairs possessing two or three H-bonds, one cannot expect that the sum of the bonding and steric interactions inside the individual X−H···Y moieties would be equal to the bonding energy of a given pair nor is it entitled to attribute the difference to some presumed cooperative effects.[6,8]

The purpose of the present study is to estimate the relative energies of the individual H-bonds, based on the NBO analysis of H-bonding. We aimed not only to evaluate the energies of the H-bonds in the essential, ubiquitous A-T and G-C DNA base pairs but also to analyze the same base pairs modified by two substituents, NH_3^+ and OH_2^+. The choice of the modified pairs was suggested by the findings that the introduction of a charged substituent to one partner of a base pair resulted in the significant alterations of the H-bond distances and the energy of the pair.[10,11] We wondered to what extent the energy of the particular bonds was different in the modified and parent base pairs and whether the energy increments correlated with the H-bond shortening or elongation.

THEORETICAL METHOD

The geometry of the Crick−Watson pairs was minimized on the basis of potential energy surface (PES) applying a wide range of theoretical levels from HF to DFT[24,25] and up to MP2[26−31] as well as various basis sets. In the case of the DFT method, the hybrid functional of Becke[32] with Lee, Yang, and Parr gradient correction,[33] B3LYP, and the gradient correction of Perdew,[34] B3P86, were applied. The triple-ζ split-valence basis set, denoted as 6-311++G(d,p) according to Pople's nomenclature,[35] as well as the D95 and D95** basis sets[36] were used to expand the orbitals and describe the geometric and electronic structural parameters. The B3P86/6-311++G(d,p) was recently used to investigate the geometric, electronic, and energetic properties of 31 canonical and wobble base pairs,[37] and this level of theory has also been used for substituted G-C and A-T derivatives. Moreover, it has been found that it is one among five functionals (out of the 44 investigated) that provides the best performance for H-bonding.[38]

For the G-C and A-T native pairs, the recently designed M05-2X functional,[39] provided a very good performance for the closed-shell organic systems[40] and the noncovalent interactions,[41] were also applied with 6-31+G(d,p), 6-311++G(d,p) and Dunning's[42] (aug-cc-pVDZ) basis sets.

The harmonic frequencies were calculated to confirm that the obtained geometry corresponds to a minimum on PES.

The calculations were carried out using the Gaussian03 series of programs.[43]

The binding energy of an A···B complex, E, was calculated as the difference between the energy of the complex and the sum of the energies of its components, taking into account the basis set superposition error[44] (BSSE). The energy can be decomposed into deformation and interaction energies.[9] Deformation energy is always positive because it determines the energy required to reshape monomers A and B from their equilibrium structure to their geometry in the A···B complex:

$$E_{def} = E_A(basis_A;opt_{A\cdots B}) - E_A(basis_A;opt_A) + E_B(basis_B;opt_{A\cdots B}) - E_B(basis_B;opt_B) \quad (1)$$

where $E_A(basis_A; opt_{A\cdots B})$ and $E_A(basis_A; opt_A)$ are the energies of the A molecule for its geometries obtained during optimization procedure of the A···B complex, $opt_{A\cdots B}$, and for the optimized monomer A, opt_A, respectively. Analogous definition stands for E_B.

Interaction energy, corrected by the BSSE,[44] was calculated as follows:

$$E_{int} = E_{A\cdots B}(basis_{A\cdots B};opt_{A\cdots B}) - E_A(basis_{A\cdots B};opt_{A\cdots B}) - E_B(basis_{A\cdots B};opt_{A\cdots B}) \quad (2)$$

where $E_A(basis_{A\cdots B}; opt_{A\cdots B})$ is the energy of molecule A, calculated using the basis of the A···.B complex, named $basis_{A\cdots B}$, and its geometry obtained during the optimization procedure of the complex, $opt_{A\cdots B}$. The other terms in eq 2 should be understood in the same way.

Subsequently, the NBO method NBO 5.G[45] was applied to study the electron distribution among the individual "natural" orbitals in a system and to estimate the individual H-bond strengths.[16]

According to the NBO analysis, X−H···Y H-bonding corresponds to an intermolecular donor−acceptor interaction between a lone pair (n) of the Lewis base (the H-bond acceptor, Y; in our case, an oxygen or nitrogen atom) and the proximate antibonding orbital (σ^*) of the Lewis acid (the H-bond donor, X−H; in our case, N−H or C−H). The strength of this interaction, E_{NBO}, can be estimated as

$$E_{NBO} = E^{(2)}_{n\to\sigma^*} + E_{\sigma\to n} \quad (3)$$

where $E^{(2)}_{n\to\sigma^*}$ means the second-order stabilization energy of the partial CT $n_Y \rightarrow \sigma_{X-H}^*$ interaction, and $E_{\sigma\to n}$ denotes the steric repulsion energy of σ_{X-H} and n_Y orbitals.

The NBO analysis was also performed at the MP2/aug-cc-pVDZ level for the molecular geometry of the G-C and A-T pairs taken from the Benchmark Energy and Geometry Data Base,[46,47] BEGDB (G···C WC1 and A···T WC from JSCH-2005 data set).

RESULTS AND DISCUSSION

The results of the estimation of the individual H-bonds strength for G-C and A-T pairs obtained by different methods substantially differed, as mentioned above. A question arises as to whether the results of this investigation are compatible with any of the previous approaches. To answer it, we confronted the sequence of the three H-bonds **a**, **b**, and **c** (see Scheme 1) from the strongest to the weakest. The outcome is shown in Table 1, which shows that the largest variety of the H-bond succession was found for the G-C pairs. In the case of the A-T complex, four methods generated the same sequence, with the central N−H···N bond (**b**) being the strongest, and all methods predicted the C−H···O (**c**) interaction as the weakest. Moreover, the NBO result suggests that this interaction is very weak or does not exist, the obtained E_{NBO} being very small. This agrees with a previous study showing an absence of this interaction,[20] and—in contrast—provided arguments in favor of the C−H···O H-bond in the A-T pair.[6−8,14,15,48]

Scheme 1. H-Bonded Complexes of Nucleic Acid Pairs

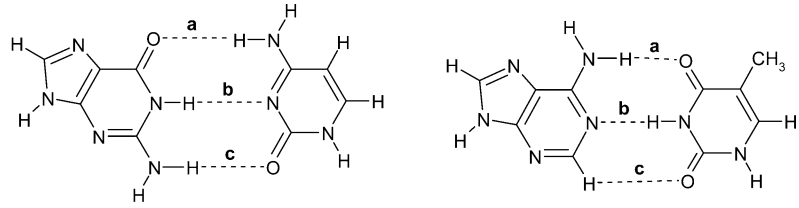
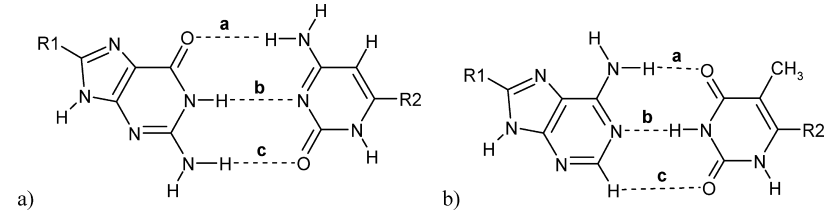


Table 1. The Sequence of the Individual H-bond Strengths in the G-C and A-T Base Pairs (Scheme 1) on the Basis of Literature Data and Our Results

method of characterizing the H-bonds	calculation level	G-C	A-T
rotation[6]	B3LYP/D95**	**c** > **b** > **a**	**a** > **b** > **c**
compliance constants[7]	B3LYP/6-311++G**	**b** > **a** > **c**	**b** > **a** > **c**
atom replacement[8]	B3LYP/6-311++G**	**a** > **c** > **b**	**b** > **a** > **c**
EML equation[14] for experimental geometry	B3LYP/6-311++G**	**a** > (**b**)[a] > **c**	**a** > (**b**)[a] > **c**
$E_{H\cdots B}$ vs ρ_{CP} relation[15]	B3LYP/6-311++G**	**a** > **c** > **b**	**b** > **a** > **c**
NBO	B3P86/6-311++G**	**a** > **b** > **c**	**b** > **a** > **c**

[a] The energy of this (N−H···N) H-bond was obtained from the difference in the total binding energy between the two monomers for the base pair and the energies of the remaining (**a** and **c**) interactions.

Scheme 2. H-Bonded Complexes of Substituted Nucleic Acid Pairs[a]



[a] (a) Guanine-cytosine (R1G-CR2) and (b) adenine-thymine (R1A-TR2); R1 = NH_3^+ or H and R2 = H or OH_2^+, respectively.

The following points can be distinguished at this point: (i) the usability of the NBO approach for the evaluation of the strength of particular H-bonds, (ii) the comparison of the energy obtained by NBO (i.e., the sum of the individual H-bonds) with the interaction energy calculated according to the supramolecular approach, and (iii) the results of the NBO approach versus previous approaches.

NBO Approach for Estimating the Strength of Particular H-Bonds. To check the ability of the NBO method to estimate the energies of the individual H-bonds, we tested whether it is adequate to account for the modification of H-bond energies due to substituents introduced to the Watson−Crick base pairs. We examined the two pairs to which one of the two substituents, NH_3^+ or OH_2^+, was introduced. They were previously examined together with the base pairs with many other substituents, as the so-called nanoswitches.[10,11] The NH_3^+ and OH_2^+ groups were chosen because they significantly modified the strength of the individual H-bonds. Schematic presentations of the studied substituted systems are shown in Scheme 2. The obtained results, at the B3P86/6-311++G** level of calculations for G-C and A-T and their substituted derivatives, are reported in Table 2 and depicted in Figure 1.

Table 2 shows that the substituents considerably changed the strength of the intermolecular interactions, both of the individual H-bonds and the binding energy and its components. What is more, the changes were consistent with what would be expected based on the modifications of the H-bond length.

It is well-known that the strength of the H-bond and its length are mutually correlated[9,49−51] and that the bond strength also depends on the types of donor and acceptor atoms. The stronger interaction means shorter H-bonds, and exactly such variability is shown in Figure 1. The shortening of the H-bond is accompanied by an increase in the energy of steric repulsion ($E_{\sigma\rightarrow n}$), a decrease in the energy of the partial CT interaction, ($E^{(2)}_{n\rightarrow\sigma^*}$), and a decrease in the H-bond energy (E_{NBO}). In Figure 1, the successive terms are marked as red, yellow, and gray symbols, respectively. Obviously, the absolute values of the $E^{(2)}_{n\rightarrow\sigma^*}$ and E_{NBO} energies increase with the shortening of the H-bond.

From the data reported in Table 2 it can be seen that in G-C, in accordance with ref 10, the substitution of G by NH_3^+ attenuates the upper (**a**, in Scheme 2a) and enhances the two lower (**b**, **c**) H-bonds, which causes net strengthening of the bonding. The substitution of the C base by OH_2^+ enhances the upper (**a**, in Scheme 2) and attenuates the two other bonds (**b**, **c**), which is accompanied by the net weakening of the complex. The corresponding sequence of the binding energies is −26.73, −34.57, and −23.90 kcal/mol and agrees with the previously obtained[10] (−26.06, −34.06, and −22.94 kcal/mol, respectively). Despite the weakest "global" intermolecular interactions in the latter case (G-C−OH_2^+ pair), the effect of the substituent on particular H-bonds is the largest, as illustrated in Figure 1a (triangles). The energy changes of G-C are congruent with the corresponding increments of the corresponding bond distances.

Table 2. Energetic Characteristics of the Individual H-bonds (**a**, **b**, **c**) in the Base Pairs[a]

		$E^{(2)}_{n\to\sigma^*}$	$E_{\sigma\to n}$	E_{NBO}	$\sum E_{NBO}$	comparison of **a**, **b**, **c** strength	E_{int}/E_{def}	E
G-C	**a**	−28.49	18.75	−9.74	−21.34	**a** > **b** > **c**	−30.70	−26.73
	b	−22.67	15.35	−7.32			3.97	(−27.86)[b]
	c	−14.98	10.70	−4.28				
NH_3^+−G-C	**a**	−16.65	11.88	−4.77	−23.74	**b** > **c** > **a**	−38.35	−34.57
	b	−28.75	18.34	−10.41			3.78	(−35.78)[b]
	c	−25.95	17.39	−8.56				
G-C−OH_2^+	**a**	−63.30	34.91	−28.39	−33.56	**a** > **b** > **c**	−30.13	−23.90
	b	−15.57	11.61	−3.96			6.23	(−25.07)[b]
	c	−4.72	3.51	−1.21				
A-T	**a**	−14.46	10.44	−4.02	−14.42	**b** > **a** > **c**	−15.05	−13.23
	b	−31.77	21.24	−10.53			1.82	(−14.10)[b]
	c	−0.43	0.56	0.13				
NH_3^+−A-T	**a**	−25.98	16.85	−9.13	−14.74	**a** > **b** > **c**	−17.81	−16.16
	b	−20.64	15.01	−5.63			1.65	(−17.07)[b]
	c	−0.24	0.26	0.02				
A-T−OH_2^+	**a**	−18.95	12.81	−6.14	−22.53	**b** > **a** > **c**	−57.63	−24.88
	b	−44.81	28.08	−16.73				(−26.17)[b]
	c	−0.47	0.81	0.34				

[a] Energy of the n → σ^* CT, $E^{(2)}_{n\to\sigma^*}$, energy of the σ → n steric repulsion, $E_{\sigma\to n}$, and their sum, E_{NBO} eq 3. $\sum E_{NBO}$ represents sum of the NBO energies for all the H-bonds in a base pair. The results are compared with the total bonding energy, E, and its components: the energy of interaction, E_{int}, eq 2, and the energy of deformation, E_{def}, eq 1. All data in kcal/mol; B3P86/6-311++G**. [b] Energies without the counterpoise correction.

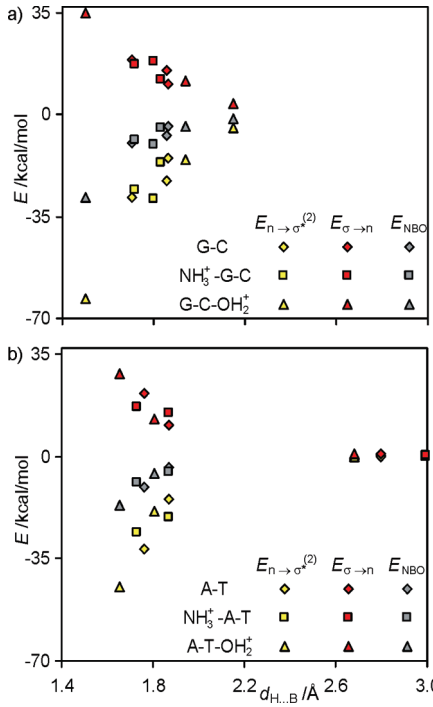
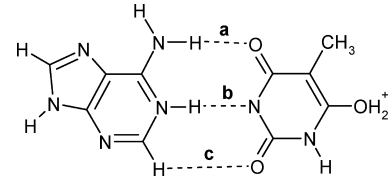


Figure 1. Relationship between the energy of the n → σ^* CT, $E^{(2)}_{n\to\sigma^*}$ (yellow symbols), energy of the σ → n steric repulsion, $E_{\sigma\to n}$ (red symbols), and their sum, E_{NBO} according to eq 3 (gray symbols) for individual H-bonds and their length, $d_{H\cdots B}$ for: (a) R1G-CR2 and (b) R1A-TR2 pairs. Diamonds indicate G-C and A-T complexes, squares mean NH_3^+−G-C and NH_3^+−A-T systems, and triangles signify G-C−OH_2^+ and A-T−OH_2^+ pairs; B3P86/6-311++G** results.

The substitution of A-T by NH_3^+ (NH_3^+ A-T complex) enhances the upper (**a**, in Scheme 2b) and attenuates the lower bond (**b**), but the substitution of T by OH_2^+ enhances both these H-bonds (Table 2). The corresponding binding energy sequence for the parent A-T and its two derivatives is −13.23, −16.16, and −24.88 kcal/mol (the previously obtained[11] −13.00, −15.88, and −25.44 kcal/mol), respectively. Obviously, the order of the interaction energy is the same but interesting are their values, particularly the last one: −15.05, −17.81, and −57.63 kcal/mol. In the A-T−OH_2^+ pair, the proton transfer takes place, that is the nitrogen atom of adenine acts as the proton donor in **b** H-bond (Scheme 3) and an electrostatic part of the interaction energy has to be particularly large. This explains the difference between the energy of H-bonds localized by NBO ($\sum E_{NBO}$ = −22.53 kcal/mol) and the interaction energy calculated according to the supramolecular approach (E_{int} = −57.63 kcal/mol).

Scheme 3. Particular H-bonds in the A-T−OH_2^+ Pair



In the case of the G-C and A-T pairs, two types of intermolecular interactions can be distinguished: X−H···O, where X = N or C, and N−H···N. Figure 2 presents a compilation of the data depicted in both parts of Figure 1, but the meaning of the symbols is different. The gray symbols indicate the N−H···O and C−H···O interactions, and the empty ones signify the N−H···N H-bonds, whereas the circles, triangles, and diamonds designate the E_{NBO}, $E^{(2)}_{n\to\sigma^*}$, and $E_{\sigma\to n}$ energies, respectively. For both types of the above-mentioned intermolecular interactions presented in Figure 2, a dependence of the energies and their components for the individual H-bonds on their lengths is nicely monotonic. Moreover, for the N−H···O and N−H···N bonds of the same energy, the latter should be longer than the former one.[52,53] This means that the observed relationship for the N−H···N interaction should be shifted toward the longer H-bonds in comparison with the N−H···O ones. The obtained dependence of the E_{NBO}, $E^{(2)}_{n\to\sigma^*}$, and $E_{\sigma\to n}$ energies on the H-bond length for both types of the interaction agree with the expectation. Figure 2 reveals that the results based

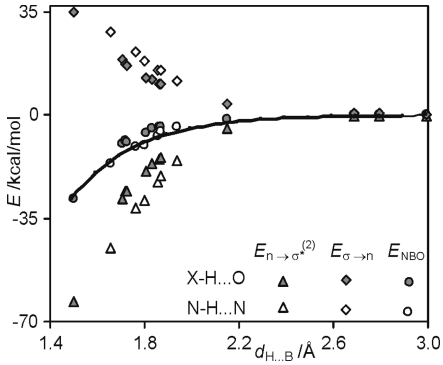


Figure 2. Relationship between the energy of the n → σ^* CT, $E_{n\to\sigma^*}^{(2)}$ (triangles), energy of the $\sigma\to$n steric repulsion, $E_{\sigma\to n}$ (diamonds), their sum, E_{NBO} eq 3 (circles) for individual H-bonds and their length, and $d_{H\cdots B}$ for G-C, A-T and their substituted derivatives R1G-CR2 and R1A-TR2 pairs. Gray symbols indicate N−H···O and C−H···O interactions, whereas empty ones signify N−H···N H-bonds. The curve represents exponential fitting for experimentally observed X−H···O (X = C, N, O) H-bonds.[13]

on the NBO approach are consistent with the conclusions attained by the comparison of the lengths of the particular H-bonds in systems with several intermolecular H-bonds. Moreover, an agreement of the obtained relative strength (NBO results) with the former qualitative prediction of the relative importance of individual H-bonds should be noted.[10,11,19,20]

Energy of the Interaction: NBO versus the Binding Energy of Complexes. The curve representing the exponential fitting of the binding energy (the dissociation energy in original paper[13]) and the length of the H-bond for 83 experimentally observed X-H···O (*X* = C, N, O) H-bonds has been added in Figure 2. As we see, the E_{NBO} energy for the N−H···O interaction is slightly lower (less negative) than predicted by the exponential fitting.

The energy of the H-bonding obtained by NBO is the sum of the strengths of the individual H-bonds, $\sum E_{NBO}$, and its value can be compared with the interaction energy, E_{int}, calculated according to the supramolecular approach. The results are reported in Table 2. It also comprises the deformation energy, E_{def}, and the binding energy, E, being the sum of the deformation and interaction energies. The latter is given with and without the counterpoise correction. The comparison of $\sum E_{NBO}$ with the E_{int} and E reveals that they substantially differ. This is because, in a polyatomic molecule, there are many "through bond" and "through space" interactions, and the values in the last two columns of Table 2 measure the total interaction of the two fragments instead of the energy confined to the H-bonds themselves. Sometimes, (e.g., for A-T) these other interactions, attractive or repulsive, cancel out each other, and the total binding energy is close to the energy of the specific H-bonds, but it is not a rule. In the Weinhold and Landis monography,[16] a difference between the two energies is denoted "residual classical electrostatic interactions energy", and its value can be compared with the energy of the H-bonding (cf. example 5.3 and Table 5.1, pp. 598−600, in ref 16). The role of the electrostatic interactions was also discussed in the studies of Reed et al.[17] and Bickelhaupt et al.[19,20]

It should be underlined that the interactions within the individual X−H···Y bonds do not constitute the total interaction energy of the two base pairs. The $n_Y \to \sigma^*_{X-H}$ interactions are a part of the total energy originating from the noncovalent contributions (E_{nL}). In a like manner, the energy of the steric interactions $\sigma_{X-H}\to n_Y$ is a part of the total energy associated with localized orbitals (E_L). The other Lewis-type interactions might be positive or negative; the analogy with the "compliance matrix" is evident, because its nondiagonal entries can also have different signs. Hence, the difference between E_{int} and the sum of the all the $E_{X-H\cdots Y}$ terms is observed.

An analogous picture emerges from the different methodological treatments, namely from the analysis of the compliance constants.[7,54] Here, the key issue is the existence of both diagonal and nondiagonal elements in the compliance matrix. The results based on the compliance matrix have been criticized because the compliance constants measure the total interaction of two fragments not the individual bond strengths because there are nonzero compliance constants for pairs of atoms that are many Angstroms apart.[55] The response to this criticism is that the interactions described by the nondiagonal entries are a real feature of polyatomic molecules, where many through-space and -bond interactions exist.[54]

This picture also emerges from the NBO analysis of all orbital interactions in a polyatomic molecule. In the language accepted in the monograph on the NBO theory and its applications,[16] all these secondary interactions (between atoms and orbitals) that are not localized within X−H···Y groupings are termed "electrostatic interactions". However, it has to be remembered that the classical electrostatic interactions alone, which sometimes were considered sufficient for interpreting the stabilities of multiply H-bonded complexes, can often be deceptive.[56] The same conclusion can be found for complexes with a single H-bond.[16]

In the NBO approach presented in this study, these electrostatic interactions can be estimated as the difference between the E_{int} of a complex and the sum of interaction energies localized in the individual H-bonds, $\sum E_{NBO}$. They seem to be positive in G-C−OH_2^+ and negative in the remainder studied pairs (Table 2), with the maximal value for A-T−OH_2^+, where the proton-transfer phenomena takes place (see above and Scheme 3).

Results of the NBO Approach versus Previous Approaches. Looking for the reasons for the differences in the sequence in G-C between all methods (Table 1), we repeated the calculations for G-C at the B3LYP/6-311++G** and B3LYP/D95** levels (applied in ref 6) to see whether differences between the results of our method and others could be ascribed to the different calculation levels. In the first case, the obtained sequence of the individual H-bonds agrees with the B3P86/6-311++G** results (Table 3). However, for the latter the results were neither in accordance with those obtained at the B3LYP/6-311++G** nor at the B3P86/6-311++G** levels. Therefore, we repeated the NBO calculations for G-C at the MP2/6-31+G*, HF/6-311++G**, MP2/6-311++G**, and B3LYP/D95, and additionally, M05-2X/aug-cc-pVDZ, M05-2X/6-31+G**, M05-2X/6-311++G** and MP2/aug-cc-pVDZ levels. The results of all calculations are displayed in Table 3. The results show that the relative energies of the interactions ascribed to the upper and central H-bond, as well as the central and lower H-bond in the G-C pair, depend on the theoretical level applied to the calculations. The obtained energies of **a** and **b**, as well as of **b** and **c**, are in some cases close to each other (differences in the

Table 3. Energetic Characteristics of the Individual H-Bonds (**a**, **b**, **c**) in the G-C Base Pair Obtained at Various Computation Levels[a]

G-C		$E^{(2)}_{n\rightarrow\sigma^*}$	$E_{\sigma\rightarrow n}$	E_{NBO}	$\sum E_{NBO}$
B3P86/6-311++G**	**a**	−28.49	18.75	−9.74	−21.34
	b	−22.67	15.35	−7.32	**a > b > c**
	c	−14.98	10.70	−4.28	
HF/6-311++G**	**a**	−12.52	8.95	−3.57	−9.93
	b	−12.92	8.67	−4.25	**b > a > c**
	b	−8.05	5.94	−2.11	
B3LYP/6-311++G**	**a**	−21.95	15.11	−6.84	−15.92
	b	−18.45	12.59	−5.86	**a > b > c**
	c	−12.01	8.79	−3.22	
B3LYP/D95	**a**	−36.10	24.27	−11.83	−21.37
	b	−26.95	22.34	−4.61	**a > c > b**
	c	−19.14	14.21	−4.93	
B3LYP/D95**	**a**	−28.0	18.95	−9.05	−16.68
	b	−20.46	17.08	−3.38	**a > c > b**
	c	−15.42	11.17	−4.25	
MP2/6-31+G*	**a**	−27.73	13.62	−14.11	−36.43
	b	−25.50	11.75	−13.75	**a > b > c**
	c	−17.16	8.59	−8.57	
MP2/6-311++G**	**a**	−23.40	16.11	−7.29	−18.17
	b	−24.34	16.31	−8.03	**b > a > c**
	c	−11.01	8.16	−2.85	
MP2/aug-cc-pVDZ[b]	**a**	−29.56	15.28	−14.28	−30.59
	b	−24.50	14.04	−10.46	**a > b > c**
	c	−15.56	9.71	−5.85	
M05-2X/aug-cc-pVDZ	**a**	−27.38	14.11	−13.27	−28.46
	b	−21.67	12.28	−9.39	**a > b > c**[c]
	c	−15.37	9.57	−5.80	

[a] Energy of the n → σ^* CT, $E^{(2)}_{n\rightarrow\sigma^*}$, energy of the $\sigma\rightarrow$n steric repulsion, $E_{\sigma\rightarrow n}$, and their sum, E_{NBO} eq 3. $\sum E_{NBO}$ represents the sum of the NBO energies for all the H-bonds in the base pair. All data are in kcal/mol. [b] $E_{int} = -32.06$ kcal/mol (ref 47), $E_{def} = 3.36$ kcal/mol (ref 47), and $E_{tot} = -28.80$ kcal/mol (refs 47 and 60). [c] The same sequence was found at the M05-2X/6-311++G** and M05-2X/6-31+G** levels.

Table 4. Energetic Characteristics of the Individual H-bonds (**a**, **b**, **c**) in the A-T Base Pair Obtained at Various Computation Levels[a]

A-T		$E^{(2)}_{n\rightarrow\sigma^*}$	$E_{\sigma\rightarrow n}$	E_{NBO}	$\sum E_{NBO}$
B3P86/6-311++G**	**a**	−14.46	10.44	−4.02	−14.42
	b	−31.77	21.24	−10.53	**b > a > c**
	c	−0.43	0.56	0.13	
HF/6-311++G**	**a**	−5.92	4.57	−1.35	−6.28
	b	−15.34	10.39	−4.95	**b > a > c**
	c	−0.18	0.20	0.02	
B3LYP/6-311++G**	**a**	−11.5	8.57	−2.93	−10.63
	b	−24.73	16.84	−7.89	**b > a > c**
	c	−0.33	0.52	0.19	
B3LYP/D95	**a**	−18.70	13.67	−5.03	−16.38
	b	−45.85	34.91	−10.94	**b > a > c**
	c	−0.60	0.19	0.41	
B3LYP/D95**	**a**	−14.83	10.82	−4.01	−9.92
	b	−27.90	22.06	−5.84	**b > a > c**
	c	−0.23	0.16	−0.07	
MP2/6-31+G*	**a**	−16.52	8.14	−8.38	−13.24
	b	−33.18	15.70	−17.48	**b > a > c**
	c	−0.68	0.22	−0.46	
MP2/6-311++G**	**a**	−11.61	8.76	−2.85	−26.32
	b	−31.87	20.84	−10.53	**b > a > c**
	c	−0.38	0.52	0.14	
MP2/aug-cc-pVDZ[b]	**a**	−14.95	8.12	−6.83	−22.48
	b	−33.22	17.57	−15.65	**b > a > c**
	c	−0.59	0.42	−0.17	
M05-2X/aug-cc-pVDZ	**a**	−13.99	7.61	−6.38	−22.24
	b	−33.04	17.18	−15.86	**b > a > c**[c]
	c	−0.74	0.57	−0.17	

[a] Energy of the n → σ^* CT, $E^{(2)}_{n\rightarrow\sigma^*}$, energy of the $\sigma\rightarrow$ n steric repulsion, $E_{\sigma\rightarrow n}$, and their sum, E_{NBO} eq 3. $\sum E_{NBO}$ represents the sum of the NBO energies for all the H-bonds in the base pair. All data are in kcal/mol. [b] $E_{int} = -16.86$ kcal/mol (ref 47), $E_{def} = 1.43$ kcal/mol (ref 47), and $E_{tot} = -15.43$ kcal/mol (refs 47 and 60). [c] The same sequence was found at the M05-2X/6-311++G** and M05-2X/6-31+G** levels.

obtained E_{NBO} energies are less than 1 kcal/mol); hence, the differences in their proper ordering. The data in the Table 3 show three possible sequences of the individual bond strengths: (i) **b > a > c**, obtained at HF/6-311++G** and MP2/6-311++G** levels, and reported by Grunenberg;[7] (ii) **a > c > a**, B3LYP/D95 and B3LYP/D95** results, and produced by the atom-replacement method;[8] and (iii) **a > b > c**, suggested by Matta et al.[14] and the NBO approach at seven calculation levels: B3P86/6-311++G**, B3LYP/6-311++G**, M05-2X/aug-cc-pVDZ, M05-2X/6-31+G**, M05-2X/6-311++G**, MP2/6-31+G*, and MP2/aug-cc-pVDZ.

The instability of the electronic structure of the G-C pair towards the applied computational level can be tentatively explained by the instability of guanine geometry and by necessity its electronic structure.[57] The energy difference between the two guanine tautomers, GUA-7 and GUA-9 (according to the nomenclature used),[57] is very small (e.g., 0.09 kcal/mol) but strongly varies at different computational levels.

It has to be mentioned that calculations using the HF and DFT methods rendered the G-C complex of planar geometry. However, the optimization at the MP2/6-31+G* and MP2/6-311++G** levels rendered the complex not exactly planar. At the first level, the two molecules had a dihedral angle between four heavy atoms belonging to **a** and **b** bonds equal to 8.61° and the same angle for **b** and **c** bonds equal to 9.09°. (This dihedral was introduced as a parameter characterizing local planarity of a base pair, within the H-bonding region.)[37] At the MP2/6-311++G** level, the dihedrals were 8.78 and 9.28°, respectively. This experiment showed that the two bases in the G-C pair are not coplanar; they are twisted about the H-bond like the blades of a propeller. This twist was about 7°.[58,59]

The incongruous results for G-C led us to repeat the calculations for the other base pairs at the above-mentioned levels to see whether the discrepancies observed for the G-C pair would also be manifested for these complexes. The obtained results are reported in Tables 4−8.

The NBO method allowed an unequivocal determination of the energy sequence in the parent A-T pair (Table 4) and in the pair substituted with the NH_3^+ (Table 5) and OH_2^+ (Table 6) substituents, at all computational levels. The energy sequence was **b > a > c** for A-T and A-T−OH_2^+, but in the latter, the strength of both **a** and **b** bonds was enhanced compared with those in A-T. By contrast, the sequence in NH_3^+A-T changed to **a > b > c**.

However, for the native and substituted G-C pairs the results generated at different levels were less equivocal. For G-C (Table 3), as mentioned above, the sequence was **a > b > c**, **b > a > c**, or **a > c > b**. Regarding the two substituted G-C pairs, the sequences were steadily **b > c > a** for NH_3^+−G-C (Table 7) and **a > b > c** for G-C−OH_2^+ (Table 8) at the five levels applied; the results were different only at the B3LYP/D95 and B3LYP/D95** levels.

Table 5. Energetic Characteristics of the Individual H-Bonds (**a**, **b**, **c**) in the NH_3^+−A-T Base Pair Obtained at Various Computation Levels[a]

NH_3^+−A-T		$E^{(2)}_{n\to\sigma^*}$	$E_{\sigma\to n}$	E_{NBO}	$\sum E_{NBO}$
B3P86/6-311++G**	**a**	−25.98	16.85	−9.13	−14.74
	b	−20.64	15.01	−5.63	**a > b > c**
	c	−0.24	0.26	0.02	
HF/6-311++G**	**a**	−11.04	7.75	−3.29	−5.72
	b	−9.68	7.24	−2.44	**a > b > c**
	c	−0.05	0.06	0.01	
B3LYP/6-311++G**	**a**	−20.87	14.6	−6.81	−10.88
	b	−15.72	11.59	−4.13	**a > b > c**
	c	−0.17	0.11	0.06	
B3LYP/D95	**a**	−30.38	20.47	−9.91	−14.79
	b	−30.58	25.94	−4.64	**a > b > c**
	c	−0.38	0.62	−0.24	
B3LYP/D95**	**a**	−26.15	17.65	−8.50	−11.01
	b	−18.47	15.94	−2.53	**a > b > c**
	c	−0.11	0.13	0.02	
MP2/6-31+G*	**a**	−24.95	11.93	−13.02	−26.03
	b	−24.79	12.17	−12.62	**a > b > c**
	c	−0.55	0.16	−0.39	
MP2/6-311++G**	**a**	−21.14	14.42	6.72	−12.33
	b	−21.43	15.76	5.67	**a > b > c**
	c	−0.28	0.34	−0.06	

[a] Energy of the n → σ^* CT, $E^{(2)}_{n\to\sigma^*}$, energy of the σ → n steric repulsion, $E_{\sigma\to n}$, and their sum, E_{NBO} eq 3. $\sum E_{NBO}$ represents the sum of the NBO energies for all the H-bonds in the base pair. All data are in kcal/mol.

Table 6. Energetic Characteristics of the Individual H-bonds (**a**, **b**, **c**) in the A-T−OH_2^+ Base Pair Obtained at Various Computation Levels[a]

A-T−OH_2^+		$E^{(2)}_{n\to\sigma^*}$	$E_{\sigma\to n}$	E_{NBO}	$\sum E_{NBO}$
B3P86/6-311++G**	**a**	−18.95	12.81	−6.14	−22.53
	b	−44.81	28.08	−16.73	**b > a > c**
	c	−0.47	0.81	0.34	
HF/6-311++G**	**a**	−25.01	17.36	−7.65	−54.4
	b	−98.91	51.38	−47.53	**b > a > c**
	c	−1.26	2.04	0.78	
B3LYP/6-311++G**	**a**	−16.22	11.19	−5.03	−18.59
	b	−37.75	24.13	−13.62	**b > a > c**
	c	−0.52	0.58	0.06	
B3LYP/D95	**a**	−20.66	14.93	−5.73	−21.87
	b	−59.39	42.63	−16.76	**b > a > c**
	c	−0.78	−1.40	0.62	
B3LYP/D95**	**a**	−19.41	13.64	−5.77	−16.11
	b	−42.11	31.41	−10.77	**b > a > c**
	c	−0.47	−0.90	0.43	
MP2/6-31+G*	**a**	−22.86	10.84	−12.02	−37.75
	b	−49.16	23.96	−25.02	**b > a > c**
	c	−1.27	0.56	−0.71	
MP2/6-311++G**	**a**	−17.88	12.42	−5.46	−27.86
	b	−55.42	32.69	−22.73	**b > a > c**
	c	−0.71	1.04	0.33	

[a] Energy of the n → σ^* CT, $E^{(2)}_{n\to\sigma^*}$, energy of the σ → n steric repulsion, $E_{\sigma\to n}$, and their sum, E_{NBO} eq 3. $\sum E_{NBO}$ represents the sum of the NBO energies for all the H-bonds in the base pair. All data are in kcal/mol.

Table 7. Energetic Characteristics of the Individual H-bonds (**a**, **b**, **c**) in the NH_3^+−G-C Base Pair Obtained at Various Computation Levels[a]

NH_3^+−G-C		$E^{(2)}_{n\to\sigma^*}$	$E_{\sigma\to n}$	E_{NBO}	$\sum E_{NBO}$
B3P86/6-311++G**	**a**	−16.65	11.88	−4.77	−23.74
	b	−28.75	18.34	−10.41	**b > c > a**
	c	−25.95	17.39	−8.56	
HF/6-311++G**	**a**	−6.62	5.09	−1.53	−11.92
	b	−16.11	10.11	−6.00	**b > c > a**
	c	−14.82	10.43	−4.39	
B3LYP/6-311++G**	**a**	−12.69	9.39	−3.30	−18.31
	b	−23.47	15.12	−8.35	**b > c > a**
	c	−21.46	14.80	−6.66	
B3LYP/D95	**a**	−36.00	7.74	−28.26	−102.2
	b	−47.75	11.87	−35.88	**c > b > a**
	c	−48.19	10.15	−38.04	
B3LYP/D95**	**a**	−27.10	5.63	−21.47	−85.95
	b	−37.99	9.03	−28.96	**c > b > a**
	c	−43.37	8.85	−34.52	
MP2/6-31+G*	**a**	−19.63	9.34	−10.29	−41.75
	b	−30.13	13.30	−16.83	**b > c > a**
	c	−28.8	14.70	−14.63	
MP2/6-311++G**	**a**	−13.26	9.97	−3.29	−21.17
	b	−30.49	19.18	−11.31	**b > c > a**
	c	−22.19	15.62	−6.57	

[a] Energy of the n → σ^* CT, $E^{(2)}_{n\to\sigma^*}$, energy of the σ → n steric repulsion, $E_{\sigma\to n}$, and their sum, E_{NBO} eq 3. $\sum E_{NBO}$ represents the sum of the NBO energies for all the H-bonds in the base pair. All data are in kcal/mol.

Table 8. Energetic Characteristics of the Individual H-bonds (**a**, **b**, **c**) in the G-C−OH_2^+ Base Pair Obtained at Various Computation Levels[a]

G-C−OH_2^+		$E^{(2)}_{n\to\sigma^*}$	$E_{\sigma\to n}$	E_{NBO}	ΣE_{NBO}
B3P86/6-311++G**	**a**	−63.30	34.91	−28.39	−33.56
	b	−15.57	11.61	−3.96	**a > b > c**
	c	−4.72	3.51	−1.21	
HF/6-311++G**	**a**	−26.86	17.3	−9.56	−11.6
	b	−7.11	5.45	−1.66	**a > b > c**
	c	−1.69	1.31	−0.38	
B3LYP/6-311++G**	**a**	−49.05	28.85	−20.2	−23.7
	b	−11.64	8.88	−2.84	**a > b > c**
	c	−2.98	2.32	−0.66	
B3LYP/D95	**a**	−68.88	40.78	−28.09	−31.29
	b	−18.78	17.26	−1.52	**a > c > b**
	c	−7.58	5.90	−1.68	
B3LYP/D95**	**a**	−50.41	30.38	−20.03	−21.94
	b	−11.52	10.66	−0.85	**a > c > b**
	c	−3.68	2.63	−1.05	
MP2/6-31+G*	**a**	−50.45	23.24	−27.21	−39.93
	b	−17.38	8.37	−9.01	**a > b > c**
	c	−6.71	3.00	−3.71	
MP2/6-311++G**	**a**	−55.82	31.73	−24.09	−28.36
	b	−15.48	11.73	−3.75	**a > b > c**
	c	−2.81	2.29	−0.52	

[a] Energy of the n → σ^* CT, $E^{(2)}_{n\to\sigma^*}$, energy of the σ → n steric repulsion, $E_{\sigma\to n}$, and their sum, E_{NBO} eq 3. $\sum E_{NBO}$ represents the sum of the NBO energies for all the H-bonds in the base pair. All data are in kcal/mol.

This result suggests that the latter two basis sets are less convenient as a starting point for the calculation of the localized natural atomic orbitals as a first step in the NBO calculations. It is worth remembering that the data obtained by the rotation method[6] at the B3LYP/D95** level differed significantly from the rest of the results, in line with this suggestion. This inference will be tested in our next study by the confrontation of the present NBO results with what would be expected based on the AIM analysis.

Finally, the difference in G-C, **a > b > c** or **b > a > c** (the latter in agreement with the compliance constants method,[7] might be accounted for by the finding that the **a** and **b** bonds are very close in energy (except that calculated using the D95 and D95** basis sets) and that even a slight difference in the calculated geometry can impact the H-bonding sequence.

Table 9. Energies of the Individual H-bonds (**a**, **b**, **c**) in G-C and A-T Base Pairs[a]

	E_{NBO} kcal/mol	
	G-C	A-T
a	−7.96	−1.93
a*	−7.31	
b	−6.97	−9.45
c	−4.07	
c*	−4.07	0.43

[a] E_{NBO} eq 3, for the rotated configurations allowing the isolation of one H-bond, calculated by the NBO method at the B3P86/6-311++G** level.

As mentioned above, the Dannenberg et al.[6] data (at the B3LYP/D95** level) strongly depart from the results obtained by other methods (Table 1). Therefore, we decided to check what would be the relative strengths of the intermolecular interactions in G-C and A-T if the H-bond strengths had been calculated for rotated base pairs using the same NBO methodology to that applied for the objects in their optimized planar geometry. The calculations were performed at the B3P86/6-311++G** level, and the results are displayed in Table 9.

For both base pairs, the energies of the **a** and **c** bonds were calculated in the same configuration as that applied by Dannenberg[6] (results for **a**, **b**, and **c**). For bonds in G-C, we also calculated a different configuration, namely, one of the bases was rotated in the common base plane, where the rotation axis, perpendicular to the plane, passed through the hydrogen atom belonging to the **a** bond (configuration marked as **a***) or **c** bond (configuration marked as **c***). In this way, the two bases of a pair formed a scissors-like arrangement. Obviously, it was impossible to form such a configuration for the central H-bond.

It can be seen that the results obtained for both pairs are similar to those in Table 2, calculated for the actual base pairs geometry at the same level. The results for **a** and **a*** are similar and the same goes for **c** and **c***, and so the energy of these bonds was practically invariable toward the way of isolating a bond to make it unpolluted by the others. Moreover, the sum of the individual H-bonds obtained by the "rotation NBO" approach (Table 9) for G-C and A-T amounted to −19.00 and −10.95 kcal/mol, respectively. The results of the comparison of these values with those obtained in the planar molecules (Table 2, −21.34 and −14.42 kcal/mol, respectively) suggest the cooperative contributions −2.34 and −3.47 kcal/mol to the H-bonding interactions in the G-C and A-T base pairs, respectively.

CONCLUSION

It was found for G-C, A-T and their substituted derivatives that the dependences of the obtained hydrogen bond (H-bond) energy, E_{NBO}, and their components, the CT, $E^{(2)}_{n\rightarrow\sigma^*}$, and steric, $E_{\sigma\rightarrow n}$, energies on the H-bond length, are monotonic (showing exponential characteristics). Moreover, the H-bonds of two types, H···O and H···N, are nicely distinguished (Figure 2), according to expectations.[52,53] Therefore, the natural bond orbital (NBO) approach allows for the calculation of strength of every particular H-bond in complexes formed by multiple H-bonding.

For A-T and its substituted derivatives, the obtained energy sequence of the individual H-bonds does not depend on the applied computational level (see Tables 4−6), and the energy sequence for A-T agrees with that generated by the compliance constant,[7] atom replacement[8] and $E_{H\cdots B}$ vs ρ_{CP} relationship[15] methods (Table 1).

The obtained results for G-C and its two derivatives were significantly different only at the B3LYP/D95 and B3LYP/D95** levels, therefore these levels seems to be inappropriate for the analysis of the H-bond strength. All other results for the substituted NH_3^+−G-C pair generate the **a** > **b** > **c** sequence and for the G-C−OH_2^+ the **b** > **c** > **a** ordering.

There should also be noted an agreement of the presented NBO and the former energy decomposition analysis (EDA)[10,11] results of the substituent effect on strength of interactions for the native and substituted Watson−Crick pairs.

For the native G-C pair, the **a** > **b** > **c** sequence was found at seven computational levels (Table 3), including results obtained for the high-quality G-C geometry[47,60] and using the M05-2X funtional. The sequence changed to **b** > **a** > **c** only at HF/6-311++G** and MP2/6-311++G** computational levels, but the difference between the strength of **a** and **b** H-bonds is less than 1 kcal/mol. This difference can be accounted for by the instability of geometry of the G-C pair toward the computational level. The above-mentioned noncoplanarity of the both bases may serve as an example. One can imagine that small modifications of the G-C geometry in varying biochemical environment may also impact relative energies of the individual H-bonds.

ACKNOWLEDGMENT

H.S. and N.S.-S. gratefully acknowledge the Interdisciplinary Center for Mathematical and Computational Modeling (Warsaw, Poland) for providing computer time and facilities. H.S. thanks the Ministry of Science and Higher Education of Poland for supporting this work under the grant no. N N204 127338, and N.S.-S. thanks the National Medicines Institute for financial support.

REFERENCES AND NOTES

(1) Scheiner, S. *Hydrogen Bonding, a Theoretical Perspective*, Oxford University Press: New York, 1997.
(2) *Hydrogen Bonding - New Insights*; Grabowski, S. J., Ed.; Springer: Dordrecht, The Netherlands: 2006.
(3) Hobza, P.; Sponer, J. Structure, Energetics, and Dynamics of the Nucleic Acid Base Pairs: Nonempirical Ab Initio Calculations. *Chem. Rev.* **1999**, *99*, 3247–3276.
(4) Atkins, P. W. *Physical Chemistry*. Oxford University Press: Oxford, U.K., 1998.
(5) Jeffrey, G. A.; Saenger, W. Hydrogen Bonding in Biological Structures. Springer: Berlin, Germany, 1991.
(6) Asensio, A.; Kobko, N.; Dannenberg, J. J. Cooperative hydrogen-bonding in adenine-thymine and guanine-cytosine base pairs. Density functional theory and møller-plesset molecular orbital study. *J. Phys. Chem. A* **2003**, *107*, 6441–6443.
(7) Grunenberg, J. Direct assessment of interresidue forces in Watson-Crick base pairs using theoretical compliance constants. *J. Am. Chem. Soc.* **2004**, *126*, 16310–16311.
(8) Dong, H.; Hua, W.; Li, S. Estimation on the individual hydrogen-bond strength in molecules with multiple hydrogen bonds. *J. Phys. Chem. A* **2007**, *111*, 2941–2945.
(9) Szatylowicz, H. Structural aspects of the intermolecular hydrogen bond strength: H-bonded complexes of aniline, phenol and pyridine derivatives. *J. Phys. Org. Chem.* **2008**, *21*, 897–914.
(10) Fonseca Guerra, C.; van der Wijst, T.; Bickelhaupt, F. M. Supramolecular switches based on the Guanine-Cytosine (GC) watson-crick

pair: Effect of neutral and ionic substituents. *Chem.—Eur. J.* **2006**, *12*, 3032–3042.
(11) Fonseca Guerra, C.; van der Wijst, T.; Bickelhaupt, F. M. Nanoswitches based on DNA base pairs: Why adenine-thymine is less suitable than guanine-cytosine. *ChemPhysChem* **2006**, *7*, 1971–1979.
(12) Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*. Oxford University Press: Oxford, U.K., 1990.
(13) Espinosa, E.; Molins, E.; Lecomte, C. Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. *Chem. Phys. Lett.* **1998**, *285*, 170–173.
(14) Matta, C. F.; Castillo, N.; Boyd, R. J. Extended weak bonding interactions in DNA: π-stacking (base-base), base-backbone, and backbone-backbone interactions. *J. Phys. Chem. B* **2006**, *110*, 563–578.
(15) Ebrahimi, A.; Habibi Khorassani, S. M.; Delarami, H. Estimation of individual binding energies in some dimers involving multiple hydrogen bonds using topological properties of electron charge density. *Chem. Phys.* **2009**, *365*, 18–23.
(16) Weinhold, F.; Landis, C. R. *Valency and Bonding. A Natural Bond Orbital Donor-Acceptor Perspective*; Cambridge University Press; Cambridge, U.K., 2005.
(17) Reed, A. E.; Weinhold, F.; Curtiss, L. A.; Pochatko, D. J. Natural bond orbital analysis of molecular interactions: Theoretical studies of binary complexes of HF, H2O, NH3, N 2, O2, F2, CO and CO2 with HF, H2O, and NH3. *J. Chem. Phys.* **1985**, *84*, 5687–5705.
(18) Curtiss, L. A.; Pochatko, D. J.; Reed, A. E.; Weinhold, F. Investigation of the differences in stability of the OC. .HF and CO. .HF complexes. *J. Chem. Phys.* **1985**, *82*, 2679–2687.
(19) Fonseca Guerra, C.; Bickelhaupt, F. M. Charge transfer and environment effects responsible for characteristics of DNA base pairing. *Angew. Chem., Int. Ed.* **1999**, *38*, 2942–2945.
(20) Fonseca Guerra, C.; Bickelhaupt, F. M.; Snijders, J. G.; Baerends, E. J. The Nature of the Hydrogen Bond in DNA Base Pairs: The Role of Charge Transfer and Resonance Assistance. *Chem.—Eur. J.* **1999**, *5*, 3581–3594.
(21) Baerends, E. J.; Autschbach, J.; Berces, A.; Bickelhaupt, F. M.; Bo, C.; Boerrigter, P. M.; Cavallo, L.; Chong, D. P.; Deng, L.; Dickson, R. M.; Ellis, D. E.; van Faassen, M.; Fan, L.; Fischer, T. H.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Gotz, A. W.; Groeneveld, J. A.; Gritsenko, O. V.; Gruning, M.; Harris, F. E.; van den Hoek, P.; Jacob, C. R.; Jacobsen, H.; Jensen, L.; van Kessel, G.; Kootstra, F.; Krykunov, M. V.; van Lenthe, E.; McCormack, D. A.; Michalak, A.; Neugebauer, J.; Nicu, V. P.; Osinga, V. P.; Patchkovskii, S.; Philipsen. P. H. T.; Post, D.; Pye, C. C.; Ravenek, W.; Rodriguez, J. I.; Ros, P.; Schipper, P. R. T.; Schreckenbach, G.; Snijders, J. G.; Sola, M.; Swart, M.; Swerhone, D.; te Velde, G.; Vernooijs, P.; Versluis, L.; Visscher, L.; Visser, O.; Wang, F.; Wesolowski, T. A.; van Wezenbeek, E. M.; Wiesenekker, G.; Wolff, S. K.; Woo, T. K.; Yakovlev, A. L.; Ziegler, T. *Amsterdam Density Functional Software, ADF 2008.01*; SCM, Vrije Universiteit: Amsterdam, The Netherlands, 2008; http://www.scm.com.
(22) Scheiner, S.; Kar, T. Red- versus blue-shifting hydrogen bonds: Are there fundamental distinctions. *J. Phys. Chem. A* **2002**, *106*, 1784–1789.
(23) Carroll, M. T.; Bader, R. F. W. An analysis of the hydrogen bond in BASE-HF complexes using the theory of atoms in molecules. *Mol. Phys.* **1988**, *65*, 695–722.
(24) Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864–B871.
(25) Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.
(26) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618–622.
(27) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 energy evaluation by direct methods. *Chem. Phys. Lett.* **1988**, *153*, 503–506.
(28) Saebo, S.; Almlof, J. Avoiding the integral storage bottleneck in LCAO calculations of electron correlation. *Chem. Phys. Lett.* **1989**, *154*, 83–89.
(29) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. Semi-direct algorithms for the MP2 energy and gradient. *Chem. Phys. Lett.* **1990**, *166*, 281–289.
(30) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. A direct MP2 gradient method. *Chem. Phys. Lett.* **1990**, *166*, 275–280.
(31) Head-Gordon, M.; Head-Gordon, T. Analytic MP2 frequencies without fifth-order storage. Theory and application to bifurcated hydrogen bonds in the water hexamer. *Chem. Phys. Lett.* **1994**, *220*, 122–128.
(32) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098–3100.
(33) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785–789.
(34) Perdew, J. P. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *33*, 8822–8824.
(35) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **1980**, *72*, 650–654.
(36) Dunning, T. H., Jr.; Hay, P. J. In *Modern Theoretical Chemistry*; Schaefer, H. F., III, Ed.; Plenum: New York, 1976; Vol. 3
(37) Sadlej-Sosnowska, N. Quantum chemical considerations on degeneracy of the genetic code: Anticodon-codon wobble base pairing. *J. Mol. Struct. THEOCHEM* **2009**, *913*, 270–276.
(38) Zhao, Y.; Truhlar, D. G. Benchmark Databases for Nonbonded Interactions and Their Use To Test Density Functional Theory. *J. Chem. Theory Comput.* **2005**, *1*, 415–432.
(39) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalnet interactions. *J. Chem. Theory Comput.* **2006**, *2*, 364–382.
(40) Zhao, Y.; Truhlar, D. G. Assessment of Density Functional for π Systems: Energy differences between cumulenes and poly-ynes; proton affinities, bond length alternation, and torsional potentials of conjugated polyenes; and proton affinities of conjugated shiff bases. *J. Phys. Chem. A* **2006**, *110*, 10478–10486.
(41) Zhao, Y.; Truhlar, D. G. Density functionals with broad applicability in chemistry. *Acc. Chem. Res.* **2008**, *41*, 157–167.
(42) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
(43) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(44) Boys, S. B.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Theor. Chim. Acta* **1970**, *19*, 553–566.
(45) Glendening, E. D.; Badenhoop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Weinhold, F. *NBO 5.G*, Theoretical Chemistry Institute: University of Wisconsin, Madison, WI, 2004; http://www.chem.wisc.edu/~nbo5.
(46) Rezac, J.; Jurecka, P.; Riley, K. E.; Cerny, J.; Valdes, H.; Pluchackova, K.; Berka, K.; Rezac, T.; Pitonak, M.; Vondrasek, J.; Hobza, P. Quantum chemical benchmark energy and geometry database for molecular clusters and complex molecular systems (www.begdb.com): a user manual and examples. *Collect. Czech. Chem. Commun.* **2008**, *10*, 1261–1270.
(47) Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complex, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.
(48) Parthasarathi, R.; Amutha, R.; Subramanian, V.; Nair, B. U.; Ramasami, T. Bader's and reactivity descriptors' analysis of DNA base pairs. *J. Phys. Chem. A* **2004**, *108*, 3817–3828.
(49) Steiner, T. The hydrogen bond in the solid state. *Angew. Chem., Int. Ed.* **2002**, *41*, 48–76.
(50) Boyd, R. J.; Choi, S. C. A bond-length-bond-order relationship for intermolecular interactions based on the topological properties of molecular charge distributions. *Chem. Phys. Lett.* **1985**, *120*, 80–85.
(51) Knop, O.; Rankin, K. N.; Boyd, R. J. Coming to grips N-H···N bonds. 2. Homocorrelations between parameters deriving from the electron density at the bond critical point. *J. Phys. Chem. A.* **2003**, *107*, 272–284.
(52) Steiner, T. Lengthening of the covalent X-H bond in heteronuclear hydrogen bonds quantified from organic and organometallic neutron crystal structures. *J. Phys. Chem. A* **1998**, *102*, 7041–7052.
(53) Gilli, P.; Pretto, L.; Bertolasi, V.; Gilli, G. Predicting Hydrogen-Bond strengths from Acid-Base molecular properties. the pKa slide rule: Toward the solution of a Long-Lasting problem. *Acc. Chem. Res.* **2009**, *42*, 33–44.
(54) Brandhorst, K.; Grunenberg, J. Characterizing Chemical Bond Strengths Using Generalized Compliance Constants. *ChemPhysChem* **2007**, *8*, 1151–1156.

(55) Barker, J.; Pulay, P. The interpretation of compalians constants and their suitability for characterizing hydrogen bonds and other weak interactions. *J. Am. Chem. Soc.* **2006**, *128*, 11324–11325.
(56) Lukin, O.; Leszczynski, J. Rationalizing the strength of hydrogen-bonded complexes. Ab initio HF and DFT studies. *J. Phys. Chem. A* **2002**, *106*, 6775–6782.
(57) Leszczynski, J. The potential energy surface of guanine is not flat: An ab initio study with large basis sets and higher order electron correlation contributions. *J. Phys. Chem. A* **1998**, *102*, 2357–2362.
(58) Saenger, W. *Principles of Nucleic Acid Structure*. Springer-Verlag: New York, 1983, p.125.
(59) Sobell, H. M.; Tomita, K.-I.; Rich, A. The crystal structure of an intermolecular complex containing a guanine and a cytosine derivative. *Proc. Natl. Acad. Sci. U.S.A.* **1963**, *49*, 885–892.
(60) Jurecka, P.; Hobza, P. True stabilization energies for the optimal planar hydrogen-bonded and stacked structures of Guanine···Cytosine, Adenine···Thymine, and their 9- and 1-methyl derivatives: complete basis set calculations at the MP2 and CCSD(T) levels and comparison with experiment. *J. Am. Chem. Soc.* **2003**, *125*, 15608–15613.

CI100288H