CHEMISTRY
A EUROPEAN JOURNAL

DOI: 10.1002/chem.200501301

Supramolecular Switches Based on the Guanine–Cytosine (GC) Watson–Crick Pair: Effect of Neutral and Ionic Substituents

Célia Fonseca Guerra, Tushar van der Wijst, and F. Matthias Bickelhaupt*[a]

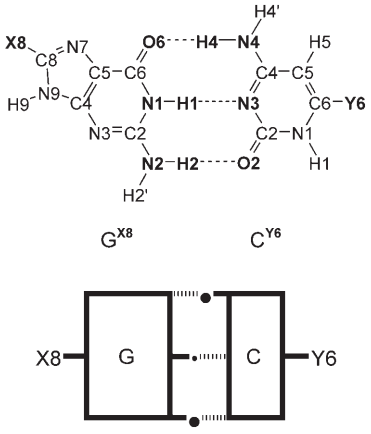
Dedicated to Professor Evert Jan Baerends on the occasion of his 60th birthday

Abstract: We have theoretically analyzed Watson–Crick guanine–cytosine (GC) base pairs in which purine-C8 and/or pyrimidine-C6 positions carry a substituent X = NH^-, NH_2, NH_3^+ (N series), O^-, OH, or OH_2^+ (O series), using the generalized gradient approximation (GGA) of density functional theory at the BP86/TZ2P level. The purpose is to study the effects on structure and hydrogen-bond strength if X = H is substituted by an anionic, neutral, or cationic substituent. We found that replacing X = H by a neutral substituent has relatively small effects. Introducing a charged substituent, on the other hand, led to substantial and characteristic changes in hydrogen-bond lengths, strengths, and hydrogen-bonding mechanism. In general, introducing an anionic substituent reduces the hydrogen-bond-donating and increases the hydrogen-bond-accepting capabilities of a DNA base, and vice versa for a cationic substituent. Thus, along both the N and O series of substituents, the geometric shape and bond strength of our DNA base pair can be chemically switched between three states, thus yielding a chemically controlled supramolecular switch. Interestingly, the orbital-interaction component in some of these hydrogen bonds was found to contribute to more than 49% of the attractive interactions and is thus virtually equal in magnitude to the electrostatic component, which provides the other (somewhat less than) 51% of the attraction.

Keywords: density functional calculations · DNA structures · hydrogen bonds · substituent effects · supramolecular switch

Introduction

Hydrogen bonding plays an important role in the formation of DNA base pairs[1] and has, therefore, been the subject of several theoretical studies.[2,3] Previously, we have shown that the generalized gradient approximation (GGA) of density functional theory (DFT) is an efficient alternative to conventional ab initio theory for accurately describing the hydrogen bonds involved in Watson–Crick base pairs (adenine–thymine (AT) and guanine–cytosine (GC), see Scheme 1).[2] Quantitative bond analyses in the framework of Kohn–Sham DFT[4] revealed that the contribution of occupied–virtual orbital interactions to the Watson–Crick hydrogen bonds is of the same order of magnitude as electrostatic interactions.[2a,c,e–g] Based on an atoms-in-molecules (AIM) analysis, Poater et al. arrived at a similar result for singly hydrogen-bonded model complexes.[3i] Our analyses contrast, however, with any approach in which a bond-



Scheme 1. Structure and simplified model of substituted Watson–Crick pairs $G^{X8}C^{Y6}$.

[a] Dr. C. Fonseca Guerra, T. van der Wijst, Dr. F. M. Bickelhaupt
Theoretische Chemie
Scheikundig Laboratorium der Vrije Universiteit
De Boelelaan 1083, 1081 HV Amsterdam (The Netherlands)
Fax: (+31) 20-598-7629
E-mail: FM.Bickelhaupt@few.vu.nl

Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author.

energy decomposition is carried out at the Hartree–Fock level of ab initio theory.[3j] The larger electrostatic attraction in Hartree–Fock is an artifact caused by the too diffuse electron density of the DNA bases (and the base pair) due to the complete absence of Coulomb correlation at this level of theory.[4] The importance of including electron correlation in ab initio studies of DNA base pairs was recently pointed out in a detailed study by Sponer et al.[3k]

The orbital-interaction component that we find, mostly originates from donor–acceptor orbital interactions of lone pairs on nitrogen and oxygen atoms of one DNA base with empty N–H σ* orbitals of the other base. Replacing N–H···O and N···H–N hydrogen bonds by N–H···F and N···H–C, respectively, leads to an elongation and weakening of these hydrogen bonds, because the fluorine atom and H–C bond are less polar, and they have lower-energy lone-pair and higher-energy σ* orbitals, respectively, than the carbonylic oxygen and H–N bond, which results in less stabilizing electrostatic and orbital interactions.[2e,f]

In the present study, we investigate the effect on the hydrogen bonding in GC if anionic, neutral, or cationic substituents are introduced at the X8 and Y6 positions. This is illustrated in Scheme 1, which shows the structure of our model systems as well as a simplified representation that will be used later on in this paper. In particular, we analyze how the hydrogen-bond lengths, strength, and bonding mechanism is affected by replacing the hydrogen atoms H8 and/or H6 in the natural guanine and cytosine bases, respectively, by the substituents NH^-, NH_2, NH_3^+ (N series), O^-, OH, or OH_2^+ (O series), see Scheme 1. Our computations are carried out with the Amsterdam density functional (ADF) program at the BP86/TZ2P level of theory.[4,5]

The above serves to lay a foundation for the design of artificial DNA bases with the purpose of tuning the bonding capabilities, that is, to make them stronger or weaker as desired. In the long term, this is of relevance for applications in, for example, supramolecular chemistry[6] and antisense technology.[7] We also compare our results with those obtained by Kawahara et al.[8] at the MP2//HF level. It will be demonstrated that, in principle, a chemically controlled, supramolecular switch can be built based on the DNA pair GC. This switch can be moved between three different states of hydrogen-bond strength and geometrical shape, simply by protonation or deprotonation of substituents.

Abstract in Portuguese: *Foram analisados teoricamente os pares de bases GC de Watson e Crick, em cujas posições C8 da purina e/ou C6 da pirimidina se encontra o substituinte X = NH^-, NH_2, NH_3^+ (série de N), O^-, OH ou OH_2^+ (série de O), utilizando a aproximação generalizada com gradientes (AGG) da teoria do funcional da densidade ao nível BP86/TZ2P. Este estudo tem como objectivo analisar os efeitos na estrutura e na força de ligação das pontes de hidrogénio, quando X = H é substituído por um substituinte aniónico, neutro ou catiónico. Foi observado que substituindo X = H por um substituinte neutro os efeitos são relativamente pequenos. Por outro lado, introduzindo um substituinte carregado induz mudanças substanciais e características no comprimento, na força e no mecanismo de ligação das pontes de hidrogénio. Em geral, introduzir um substituinte aniónico reduz, na base de ADN, a capacidade de doação e aumenta a capacidade de aceitação de pontes de hidrogénio, e vice versa no caso de um substituinte catiónico. Assim, ao longo de ambas as séries de substituintes de N e O, a estrutura e a força de ligação do nosso par de base de ADN podem ser quimicamente alternados entre três estados, obtendo-se assim um interruptor supramolecular controlado quimicamente. Foi observado que, curiosamente, em algumas destas pontes de hidrogénio, a componente de interacção das orbitais contribui mais de 49 % para as interacções atractivas, sendo, portanto, virtualmente igual em magnitude à componente electrostática que fornece os outros (um pouco menos de) 51 % da atracção.*

Computational Methods

General procedure: All calculations were performed using the Amsterdam density functional (ADF) program developed by Baerends and others.[5] The numerical integration was performed using the procedure developed by te Velde et al.[5g,h] The molecular orbitals (MOs) were expanded in a large uncontracted set of Slater-type orbitals (STOs) containing diffuse functions: TZ2P (no Gaussian functions are involved).[5i] The basis set is of triple-ζ quality for all atoms and has been augmented with two sets of polarization functions, that is, 3d and 4f on C, N, and O atoms, and 2p and 3d on H atoms. The 1s core shells of carbon, nitrogen, and oxygen atoms were treated with the frozen-core approximation.[5c] An auxiliary set of s, p, d, f, and g STOs was used to fit the molecular density and to represent the Coulomb and exchange potentials accurately in each self-consistent field cycle.[5j]

Equilibrium structures were optimized by using analytical gradient techniques.[5k] Geometries and energies were calculated at the BP86 level of the generalized gradient approximation (GGA): exchange is described by the Slater Xα potential[5l] with nonlocal corrections due to Becke[5m,n] added self-consistently, and correlation is treated in the Vosko–Wilk–Nusair (VWN) parameterization,[5o] with nonlocal corrections due to Perdew[5p] added, again, self-consistently (BP86).[5q]

Bond-energy analysis: The overall bond energy ΔE is made up of two major components [Eq. (1)]:

$$\Delta E = \Delta E_{\text{prep}} + \Delta E_{\text{int}} \quad (1)$$

In this formula, the preparation energy ΔE_{prep} is the amount of energy required to deform the separate bases from their equilibrium structure to the geometry that they acquire in

the pair. The interaction energy ΔE_{int} corresponds to the actual energy change when the prepared bases are combined to form the base pair. It is analyzed in the hydrogen-bonded model systems in the framework of the Kohn–Sham MO model using a decomposition of the bond into electrostatic interaction, exchange repulsion (or Pauli repulsion), and (attractive) orbital interactions [Eq. (2)].[4,9,10]

$$\Delta E_{int} = \Delta V_{elstat} + \Delta E_{Pauli} + \Delta E_{oi} \quad (2)$$

The term ΔV_{elstat} corresponds to the classical electrostatic interaction between the unperturbed charge distributions of the prepared (i.e., deformed) bases and is usually attractive. The Pauli repulsion ΔE_{Pauli} comprises the destabilizing interactions between occupied orbitals and is responsible for the steric repulsion. The orbital interaction ΔE_{oi} in any MO model, and therefore also in Kohn–Sham theory, accounts for charge transfer (i.e., donor–acceptor interactions between occupied orbitals on one moiety with unoccupied orbitals of the other, including the HOMO–LUMO interactions) and polarization (empty/occupied orbital mixing on one fragment due to the presence of another fragment).

Because the Kohn–Sham MO method of DFT in principle yields exact energies and, in practice, with the available density functionals for exchange and correlation, rather accurate energies, we have the special situation that a seemingly one-particle model (an MO method) in principle completely accounts for the bonding energy. In particular, the orbital-interaction term of the Kohn–Sham theory comprises the often-distinguished attractive contributions, namely, charge transfer, induction (polarization), and dispersion. One could in the Kohn–Sham MO method try to separate polarization and charge transfer, as has been done by Morokuma in the Hartree–Fock model, but this distinction is not sharp. In fact, contributions such as induction and charge transfer, and also dispersion, can be given an intuitive meaning, but whether, or with what precision, they can be quantified, remains a controversial subject. In view of the conceptual difficulties, we refrain from further decomposing the KS orbital-interaction term, except by symmetry (see below). We have observed that the orbital interactions are mostly of the donor–acceptor type (N or O lone pair on one moiety with N–H σ* orbital of the other), and we feel it is therefore justified to denote the full orbital-interaction term for brevity just as "charge-transfer" or "covalent" contribution, as opposed to the electrostatic and Pauli repulsion contributions. However, the straightforward denotation "orbital interaction" avoids confusion with the charge-transfer energy, which features in other elaborate decomposition schemes[11] that also give rise to induction and dispersion contributions. We do not attempt to quantify such contributions, but combine them in the Kohn–Sham orbital interaction.

The orbital-interaction energy can be decomposed into the contributions from each irreducible representation Γ of the interacting system [Eq. (3)] using the extended transition-state (ETS) scheme developed by Ziegler and Rauk:[10]

$$\Delta E_{oi} = \sum_{\Gamma} \Delta E_{\Gamma} \quad (3)$$

Note that our approach differs in this respect from the Morokuma scheme,[9] which instead attempts a decomposition of the orbital interactions into polarization and charge transfer). In systems with a clear σ, π, or A′, A″ separation (such as our DNA base pairs), the above symmetry partitioning proves to be most informative.

Analysis of the charge distribution: The electron-density distribution was analyzed using the Voronoi deformation density (VDD) method.[12] The VDD charge Q_A was computed as the (numerical) integral of the deformation density $\Delta\rho(\mathbf{r}) = \rho(\mathbf{r}) - \sum_B \rho_B(\mathbf{r})$ associated with the formation of the molecule from its atoms over the volume of the Voronoi cell of atom A [Eq. (4)]. The Voronoi cell of atom A is defined as the compartment of space bound by the bond midplane on and perpendicular to all bond axes between nucleus A and its neighboring nuclei (cf. Wigner–Seitz cells in crystals).[13]

$$Q_A = -\int_{\text{Voronoi cell A}} (\rho(\mathbf{r}) - \sum_B \rho_B(\mathbf{r}))d\mathbf{r} \quad (4)$$

Here, $\rho(\mathbf{r})$ is the electron density of the molecule and $\sum_B \rho_B(\mathbf{r})$ the superposition of atomic densities ρ_B of a fictitious promolecule without chemical interactions, which is associated with the situation in which all atoms are neutral. The interpretation of the VDD charge Q_A is rather straightforward and transparent. Instead of measuring the amount of charge associated with a particular atom A, Q_A directly monitors how much charge flows, due to chemical interactions, out of ($Q_A > 0$) or into ($Q_A < 0$) the Voronoi cell of atom A, that is, the region of space that is closer to nucleus A than to any other nucleus.

Results and Discussion

Assessment of the approach: The results of our BP86/TZ2P study on the formation of the natural and substituted Watson–Crick pairs $G^{X8}C^{Y6}$ are summarized in Table 1 (geometries and bond energies), Table 2 (bond-energy decomposition), and Figures 1 (geometry), 2 and 3 (atomic charges), and 4 (orbital electronic structure). Cartesian coordinates of the equilibrium geometries of all species that occur in this work are provided in the Supporting Information. The nomenclature that we use is illustrated in Scheme 1. We use the letters G and C for the natural DNA bases guanine and cytosine and indicate with a superscript the substitutions of the atoms X8 and Y6 in the bases. Thus, $G^{H8}C^{H6}$ represents the natural Watson–Crick pair GC whereas $G^{(NH_2)8}C^{(NH_2)6}$ refers to a guanine–cytosine complex in which guanine H8 and cytosine H6 have been replaced by amino groups.

Table 1. Hydrogen-bond lengths (in Å) and energies ΔE (in kcal mol^{-1}) in $G^{X8}C^{Y6}$.[a]

X8	Y6	O6···H4–N4	N1–H1···N3	N2–H2···O2	ΔE
H	H	2.73	2.88	2.87	−26.06
NH_2	NH_2	2.73	2.87	2.87	−25.65[b] (−23.45[a])
OH	OH	2.73	2.88	2.85	−25.57
NH^-	H	2.60	2.90	3.00	−22.70
NH_2	H	2.70	2.88	2.90	−25.64[b] (−23.96[a])
NH_3^+	H	2.82	2.83	2.75	−34.06
O^-	H	2.58	2.91	3.01	−22.04
OH	H	2.72	2.88	2.87	−25.94
OH_2^+	H	2.84	2.84	2.73	−34.60
H	NH^-	2.88	2.82	2.67	−37.02
H	NH_2	2.75	2.87	2.85	−26.15[b] (−25.76[a])
H	NH_3^+	2.58	2.91	3.05	−23.62
H	O^-	2.87	2.83	2.67	−36.21
H	OH	2.73	2.88	2.85	−25.75
H	OH_2^+	2.56	2.92	3.08	−22.94

[a] Computed at the BP86/TZ2P level with bases in C_1 symmetry and base pairs in C_s symmetry unless stated otherwise. [b] Bases and base pairs computed in C_1 symmetry. This full optimization, without symmetry constraints, yields (within our quantum chemical approach) the exact bond energies for systems involving amino substituents at guanine C8 and/or cytosine C6.

The choice for the BP86 density functional[5m–p] and the TZ2P basis set is based on our previous investigation[2b] of the performance of various GGA density functionals and basis sets for the AT and GC Watson–Crick base pairs, in which it was shown that BP86/TZ2P agrees excellently with experimental data. This was also confirmed in a more recent study on mismatches[2d] of DNA bases and artificial mimics[2e–g] of Watson–Crick pairs involving N–H···F and N···H–C hydrogen bonds. In the present study, we have optimized all isolated DNA bases in C_1 symmetry without any geometry restriction, whereas all DNA base pairs have been optimized and analyzed in C_s symmetry. We have verified that it is valid to impose this symmetry: optimization of the base pairs in C_1 symmetry, with amino groups substituted at the X8 position of guanine and/or the Y6 position of cytosine, starting from structures with slightly pyramidal amino groups, yields base pairs with amino groups involved in hydrogen bonds being planar and amino substituents being pyramidal. The hydrogen-bond lengths in C_1-symmetric base pairs differ by less than 0.01 Å and the hydrogen-bond energies are slightly stronger, that is, by 1.68, 0.39, and 2.20 kcal mol^{-1} for $G^{(NH_2)8}C$, $GC^{(NH_2)6}$, and $G^{(NH_2)8}C^{(NH_2)6}$, respectively, if compared with the corresponding base pairs optimized in C_s symmetry (compare the ΔE values in Tables 1 and 2). The slight strengthening in hydrogen-bond energies, if one goes from C_s- to C_1-optimized base pairs, originates from the stabilization that is associated with the geometric relaxation of the amino substituents from a planar to a pyramidal structure, and not to differences in the Watson–Crick hydrogen bonding between the DNA bases. For the separate bases $G^{(NH_2)8}$ and $C^{(NH_2)6}$, the energy gain associated with allowing the amino substituents to relax from a planar to a pyramidal geometry amounts to 1.36 and 0.33 kcal mol^{-1}, respectively. Note that this stabilization associated with pyramidalization of the amino substituents in the separate DNA bases is not very different from that for the corresponding singly-substituted Watson–Crick pairs $G^{(NH_2)8}C$ and $GC^{(NH_2)6}$. For the hydroxyl and the charged substituents, optimization of the base pairs in C_1 symmetry, starting from structures with slightly pyramidal amino groups, again yields, within the numerical precision, the C_s-symmetric structures mentioned above (hydrogen-bond distances and energies differ by less than 0.01 Å and 0.1 kcal mol^{-1}, respectively). It is therefore justified to carry out the bond-energy analysis in C_s symmetry, which enables us to quantitatively separate the orbital interactions occurring in the σ- and π-electron systems [see Eq. (3)].

In our previous work on natural Watson–Crick pairs,[2b] mismatches of DNA bases,[2d] and the artificial mimic AF,[2e–g] we have shown that hydrogen-bond lengths optimized at the Hartree–Fock (HF) level are up to 0.25 Å longer than those optimized at the BP86/TZ2P level, while geometry optimization at the MP2 level agrees significantly better with our DFT approach. The latter is nicely reconfirmed by a comparison of our BP86/TZ2P values for the hydrogen-bond distances O6–N4, N1–N3, and N2–O2 in GC (2.73, 2.88, and 2.87 Å, respectively) and those obtained recently by Sponer et al.[3k] at the RI-MP2/cc-pVTZ level (2.75 Å, 2.90 Å, and 2.89 Å, respectively): again, the two approaches agree remarkably well. Here we arrive at a similar result for $G^{(NH_2)8}C$ and $GC^{(NH_2)6}$ which, to the best of our knowledge, are the only of our model base pairs that have been investigated in an earlier theoretical study. For $G^{(NH_2)8}C$, for example, Kawahara et al.[8] found O6–H4, N1–H3, and H2–O2 distances of 1.90, 2.05, and 2.01 Å, respectively, at HF/6-31G(d,p), which are up to 0.26 Å longer than our BP86/TZ2P values of 1.64, 1.84, and 1.87 Å, respectively (not shown in Table 1). Note that the model systems of Kawahara et al.[8] and ours differ: Kawahara et al. use DNA bases that are methylated at purine-N9 and pyrimidine-N1 positions whereas our DNA bases carry a hydrogen atom at these positions. We have shown,[2a,b] however, that introducing methyl groups at purine-N9, pyrimidine-N1, and at pyrimidine-C5 (i.e., replacing uracil (U) by T) affects hydrogen-bond distances by 0.01 Å or less, that is, by much less than the effect of 0.26 Å discussed above. This effect must therefore be attributed to the difference in quantum chemical methods, in particular, to the notorious tendency of Hartree–Fock theory to underestimate hydrogen-bond strengths and to overestimate hydrogen-bond distances.

Our BP86 bond energies agree excellently with the BSSE-corrected ab initio bond energies obtained by Kawahara et al.[8a] at MP2/6-31+G(2d′,p′)//HF/6-31G(d,p) (see Table 3 in ref. [8a]). For GC, $G^{(NH_2)8}C$, and $GC^{(NH_2)6}$ our DFT results

(bases and pair both optimized in C_1 symmetry) are −26.06, −25.64, and −26.15, respectively, and the MP2 energies (bases and base pair both optimized in C_s symmetry) are −26.08, −25.79, and −26.21 kcal mol^{-1}, respectively. There is also satisfactory agreement for the GC bond energy between our BP86/TZ2P value of 26.1 kcal mol^{-1} and the complete-basis-set extrapolated RI-MP2 CBS value of −28.2 kcal mol^{-1} for GC, computed by Sponer et al.[3k] The latter authors also estimated the correction for correlation effects at the CCSD(T) level to be only −0.6 kcal mol^{-1} using the 6-31G* basis set.

Substituent effects on geometries and energies: Substituting hydrogen by a neutral substituent (NH_2 or OH) at X8 and Y6 in the Watson–Crick pair $G^{X8}C^{Y6}$ causes relatively small changes in hydrogen-bond distances and energies (see Table 1), which is in agreement with earlier work[2h] on neutral halogen (F, Cl, and Br)-substituted base pairs. The hydrogen-bond distances O6···H4−N4, N1−H1···N3, and N2−H2···O2 in the natural Watson–Crick pair GC are 2.73, 2.88, and 2.87 Å, respectively, and provide a hydrogen-bond energy ΔE of −26.06 kcal mol^{-1} (Table 2; see also ref. [2a,b]). Introducing one or two neutral substituents at guanine X8 and/or cytosine X6 has hardly any effect on hydrogen-bond distances and the hydrogen-bond energies: the effects are 0.03 Å and 0.5 kcal mol^{-1} or less, as can be seen in Table 1. On the other hand, introducing a charged substituent at guanine X8 or cytosine X6 has quite pronounced effects: hydrogen bonds contract or expand by up to 0.21 Å and Watson–Crick hydrogen bonds are stabilized or destabilized by as much as 11 kcal mol^{-1}. In the following, we examine in more detail the substituent effects and trends therein along X = NH^-, NH_2, and NH_3^+ (N series) and along X = OH^-, OH, and OH_2^+ (O series). This corresponds to three different states of protonation connected through elementary protonation (from left to right) or deprotonation steps (from right to left).

First, we inspect the trends along the N and O series in $G^{X8}C$, that is, if only guanine carries a substituent X8 while cytosine is unsubstituted (see Scheme 1 and Table 1). Introducing a negatively charged amide substituent NH^- at X8, that is, going from GC to $G^{(NH^-)8}C$, causes the Watson–Crick hydrogen-bond strength ΔE to decrease from −26.06 to −22.70 kcal mol^{-1} while, simultaneously, the O6···H4−N4 bond contracts from 2.73 to 2.60 Å and the N1−H1···N3 and N2−H2···O2 bonds expand from 2.88 to 2.90 Å and from 2.87 to 3.00 Å, respectively. Next, two successive protonation steps, that is, proceeding along $G^{(NH^-)8}C$, $G^{(NH_2)8}C$ and $G^{(NH_3^+)8}C$, have the effect of switching the Watson–Crick hydrogen-bond strength from −22.70 ("weak") to −25.64 ("intermediate") to −34.06 kcal mol^{-1} ("strong"). This switching in bond strength is accompanied by a characteristic change in the geometric shape of the substituted guanine–cytosine base pair: in the orientation that we use in our illustrations, keeping guanine fixed, this corresponds to cytosine being "bent up", "not bent", and "bent down" relative to natural GC. This is schematically illustrated in Figure 1a–c, in which bases are represented with bold lines, natural hydrogen bonds by broad dashes, weakened hydrogen bonds by narrow dashes, and strengthened hydrogen bonds by plain lines (see also Scheme 1).

The origin of this switching of the geometrical shape is the step-wise weakening and elongation of the upper hydrogen bond O6···H4−N4 together with the step-wise stabilization and contraction of the middle and lower hydrogen bonds N1−H1···N3 and N2−H2···O2, every time a proton is added to the substituent N8 atom at guanine as we go along the series $G^{(NH^-)8}C$, $G^{(NH_2)8}C$, and $G^{(NH_3^+)8}C$ (see Table 1). Note that the strengthening of the two lower hydrogen bonds along this series outweighs the weakening of the upper one as follows from the net increase in Watson–Crick hydrogen-bond strength. The mechanism and features in the electronic structure that are behind this substituent-induced switching will be discussed below.

Table 2. Analysis of the Watson–Crick hydrogen-bond energy ΔE (in kcal mol^{-1}) in $G^{X8}C^{Y6}$ base pairs.[a]

X8	H	NH_2	OH	NH^-	NH_2	NH_3^+	O^-	OH	OH_2^+	H	H	H	H	H	H
Y6	H	NH_2	OH	H	H	H	H	H	H	NH^-	NH_2	NH_3^+	O^-	OH	OH_2^+
							Orbital-interaction decomposition								
ΔE_σ	−29.29	−29.70	−29.71	−37.08	−30.11	−33.15	−40.52	−29.60	−33.03	−41.23	−29.18	−36.12	−40.28	−29.44	−39.65
ΔE_π	−4.74	−4.58	−4.71	−8.61	−4.84	−5.73	−7.48	−4.77	−5.75	−6.92	−4.53	−6.87	−6.72	−4.69	−7.48
ΔE_{oi}	−34.04	−34.27	−34.42	−45.69	−34.95	−38.88	−47.99	−34.37	−38.77	−48.15	−33.71	−42.99	−47.00	−34.13	−47.13
							Bond-energy decomposition								
ΔE_{Pauli}	51.93	52.98	52.92	63.97	53.45	54.93	68.31	52.50	54.30	67.16	51.86	59.04	66.04	52.35	62.85
ΔV_{elstat}	−48.49	−48.87	−49.01	−51.30	−48.88	−54.45	−53.76	−48.72	−54.43	−64.79	−48.89	−47.36	−63.77	−48.82	−48.34
$\Delta E_{Pauli}+\Delta V_{elstat}$	3.43	4.11	3.91	12.67	4.57	0.48	14.54	3.78	-0.12	2.37	2.98	11.68	2.27	3.52	14.52
ΔE_{oi}	−34.04	−34.27	−34.42	−45.69	−34.95	−38.88	−47.99	−34.77	−38.77	−48.15	−33.71	−42.99	−47.00	−34.13	−47.13
ΔE_{int}	−30.61	−30.17	−30.51	−33.02	−30.38	−38.40	−33.45	−30.58	−38.90	−45.77	−30.73	−31.31	−44.73	−30.61	−32.62
ΔE_{prep}	4.54	6.72	4.94	10.32	6.42	4.34	11.41	4.64	4.30	8.75	4.97	7.69	8.52	4.86	9.68
ΔE	−26.06	−23.45	−25.57	−22.70	−23.96	−34.06	−22.04	−25.94	−34.60	−37.02	−25.76	−23.67	−36.21	−25.75	−22.94
$\%\Delta E_{oi}$[b]	41.2	41.2	41.3	47.1	41.7	41.7	47.2	41.4	41.6	42.6	40.8	47.6	42.4	41.1	49.4

[a] Computed at BP86/TZ2P with bases in C_1 symmetry and base pairs in C_s symmetry. See also Table 1 for ΔE of $G^{X8}C^{Y6}$ with X8, Y6 = NH_2. [b] Percentage ΔE_{oi} of all attractive forces (i.e., $\Delta V_{elstat}+\Delta E_{oi}$).

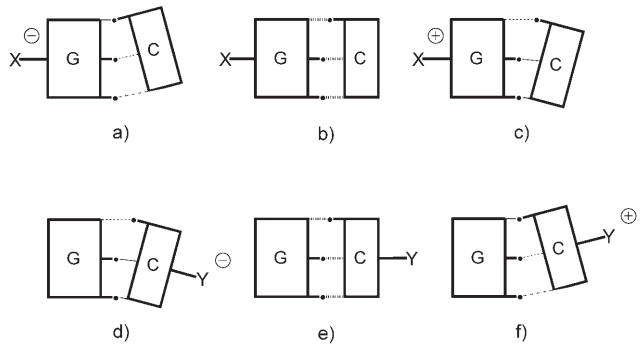


Figure 1. Schematic representation of substituent effects on the Watson–Crick hydrogen bonds in GC. Note that weak, intermediate, and strong hydrogen bonding goes with bending cytosine "up", "not", and "down" along a)–c) and d)–f), respectively (see also Scheme 1).

The same trends in hydrogen-bond strength and molecular shape emerge in an even more pronounced manner along the O series, from X8 = OH^- to OH to OH_2^+ (see Table 1). Thus, going in two successive protonation steps from $G^{(O^-)8}C$ to $G^{(OH)8}C$ to $G^{(OH_2^+)8}C$, the Watson–Crick hydrogen-bond energy ΔE is strengthened from −22.04 ("weak") to −25.94 ("intermediate") to −34.60 kcal mol^{-1} ("strong") while simultaneously, in the orientation we use in our illustrations, the cytosine base is bending downwards from "bent up" to "not bent" to "bent down", as shown schematically in Figure 1a–c.

On the other hand, introduction of the same N and O series of substituents at the other DNA base, that is, at cytosine C6 in GC^{Y6}, induces substituent effects and trends therein that run counter to the substituent effects at guanine C8 (see Scheme 1 and Table 1). Thus, along $GC^{(NH^-)6}$, $GC^{(NH_2)6}$ and $GC^{(NH_3^+)6}$, the Watson–Crick hydrogen-bond energy ΔE is destabilized from −37.02 to −26.15 to −23.62 kcal mol^{-1} and also, along $GC^{(O^-)6}$, $GC^{(OH)6}$ and $GC^{(OH_2^+)6}$, ΔE is destabilized from −36.21 to −25.75 to −22.94 kcal mol^{-1} (see Table 1). In both series this switching in bond strength from "strong" to "intermediate" to "weak" is accompanied by a characteristic change in the geometric shape of the substituted guanine–cytosine base pair: in the orientation that we use in our illustrations, keeping guanine fixed, this corresponds to cytosine being "bent down", "not bent", and "bent up" relative to natural GC. This is schematically illustrated in Figure 1d–e (see also Scheme 1).

Apart from the aspect of having a chemically controlled supramolecular switch, the above results are also interesting regarding the common idea that ionic hydrogen bonds are in general stronger than neutral ones.[1b,c,14] Here, we see that this picture is not true. For example, the positively charged NH_3^+ and OH_2^+ substituents behave in conformance with textbook knowledge if they are introduced at guanine C8 in $G^{X8}C$. Thus, the positively charged ionic Watson–Crick pairs $G^{(NH_3^+)8}C$ (ΔE = −34.06 kcal mol^{-1}) and $G^{(OH_2^+)8}C$ (ΔE = −34.60 kcal mol^{-1}) are indeed approximately 8 kcal mol^{-1} more strongly bound than natural GC (ΔE = −26.06 kcal mol^{-1}). However, if the same positively charged substituents are introduced at cytosine C6, the textbook rules are violated: thus, the positively charged ionic Watson–Crick pairs $GC^{(NH_3^+)6}$ (ΔE = −23.62 kcal mol^{-1}) and $GC^{(OH_2^+)6}$ (ΔE = −22.94 kcal mol^{-1}) are about 3 kcal mol^{-1} more weakly bound than natural GC. Likewise, the negatively charged $GC^{(NH^-)6}$ (ΔE = −37.02 kcal mol^{-1}) and $GC^{(O^-)6}$ (ΔE = −36.21 kcal mol^{-1}) that are some 11 kcal mol^{-1} more strongly bound than natural GC, conform to the textbook idea that ionic hydrogen bonds are stronger than neutral ones, but this rule is again violated in $G^{(NH^-)8}C$ (ΔE = −22.70 kcal mol^{-1}) and $G^{(O^-)8}C$ (ΔE = −22.04 kcal mol^{-1}), which are 3–4 kcal mol^{-1} more weakly bound than natural GC.

Origin of the substituent effects: The substituent effects described above, in particular the switching of Watson–Crick hydrogen-bond strength and geometric shape along the N and O series of substituents, can be understood on the basis of how the substituents interfere with the electron density distribution and the orbital electronic structure of the DNA bases and, thus, how they modify the DNA bases' hydrogen-bonding capabilities (see Figures 2–4). The changes in the bonding can furthermore be monitored through the quantitative decomposition of the hydrogen-bond energy described in the section about bond energy analysis (see Table 2). There are two major mechanisms that reinforce each other: 1) the effect of the substituents on the electron-density distribution of the DNA base, in particular, the charges of the front atoms (i.e., those atoms that are involved in hydrogen bonding), and 2) the effect of the substituent on the energies of the molecular orbitals of the DNA base. The former affects the electrostatic attraction whereas the latter influences the donor–acceptor orbital interactions associated with the hydrogen bonds.

First, we inspect the substituent effects on the front atomic charges. Figures 2 and 3 show the VDD atomic charges of isolated (i.e., noninteracting) DNA bases G^{X8} and C^{Y6} with neutral and charged substituents. As can be seen, replacing the hydrogen atom at guanine C8 or cytosine C6 by a neutral substituent has little effect on the atomic charges of the front atoms. However, charged substituents have a much larger effect: the negatively charged substituents (NH^- and O^-) inject an excess negative charge into the DNA bases and cause the hydrogen front atoms (i.e., the ones that are involved in hydrogen bonding) to become less positive by up to 62 mau (mau = milli atomic units), which weakens the electrostatic attraction in the corresponding hydrogen bond. Nitrogen or oxygen front atoms, on the other hand, become more negative by up to 109 mau, which strengthens the electrostatic attraction in the corresponding hydrogen bond (see Figures 2 and 3). Positively charged substituents (NH_3^+ and OH_2^+) cause the hydrogen front atoms to become more positive by up to 57 mau, which strengthens the electrostatic attraction in the corresponding hydrogen bond. Nitrogen or oxygen front atoms become less negative by up to 114 mau, which weakens the electrostatic attraction in the corresponding hydrogen bond (see Figures 2 and 3).

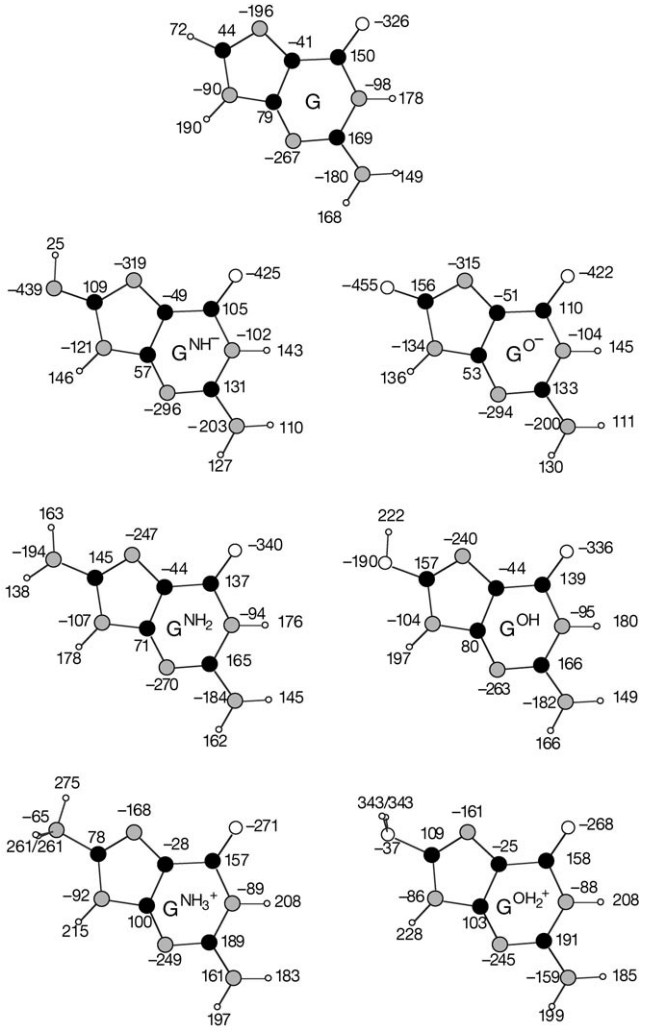


Figure 2. VDD atomic charges Q_A (in mau) in isolated G^{X8} bases with the geometries they adopt in GC, $G^{(NH_2)8}C^{(NH_2)6}$, $G^{(NH_3^+)8}C$, $G^{(NH^-)8}C$, $G^{(OH)8}C^{(OH)6}$, $G^{(OH_2^+)8}C$, and $G^{(O^-)8}C$ (see also Scheme 1).

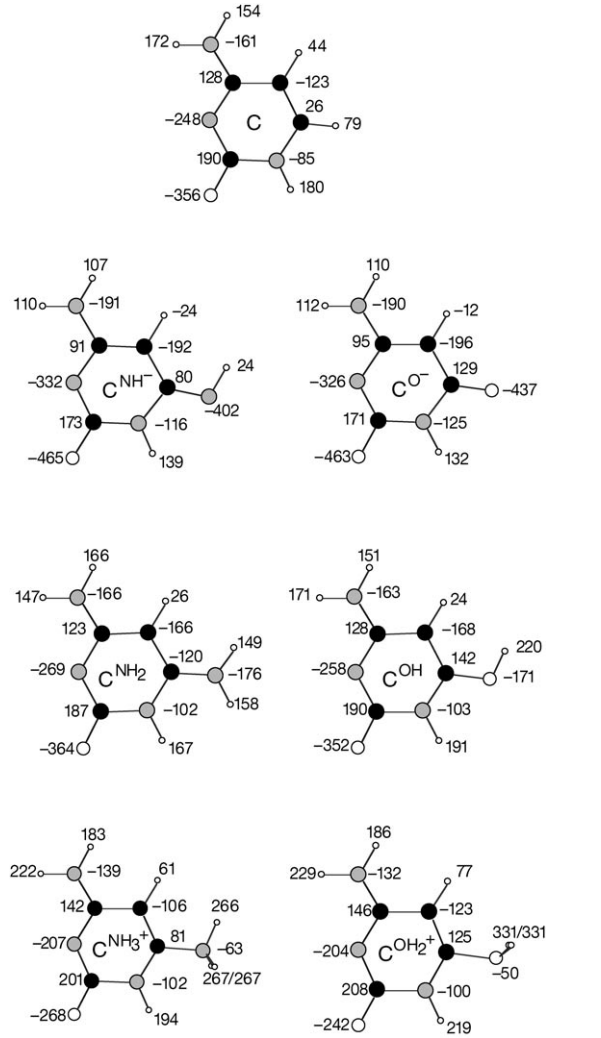


Figure 3. VDD atomic charges Q_A (in mau) in isolated C^{Y6} bases with the geometries they adopt in GC, $G^{(NH_2)8}C^{(NH_2)6}$, $GC^{(NH_3^+)6}$, $GC^{(NH^-)6}$, $G^{(OH)8}C^{(OH)6}$, $GC^{(OH_2^+)6}$, and $GC^{(O^-)6}$ (see also Scheme 1).

Thus, along the N series in, for example, the guanine-substituted GC base pairs $G^{X8}C$, the electrostatic attraction in the upper O6···H4–N4 hydrogen bond of the Watson–Crick pair is weakened whereas the electrostatic attraction in the middle N1–H1···H3 and lower N2–H2···O2 hydrogen bonds is strengthened if one goes from the negative NH^- via the neutral NH_2 to the positive NH_3^+ substituent (this is not immediately clear from the energy decomposition in Table 2; an explanation for this is provided shortly).

Next, we examine the substituent effects on the orbital electronic structure of the bases and the consequences thereof for the donor–acceptor orbital interactions. As will be seen, the latter play a key role amongst the stabilizing forces in the hydrogen bonds as well as in the trends in substituent effects thereon, in line with previous studies.[2c,d,f–h] Figure 4 shows a simplified MO interaction diagram for natural GC, in which the repulsive interactions are left out and only the donor–acceptor interactions are represented (for a detailed discussion of the orbital interactions in GC see

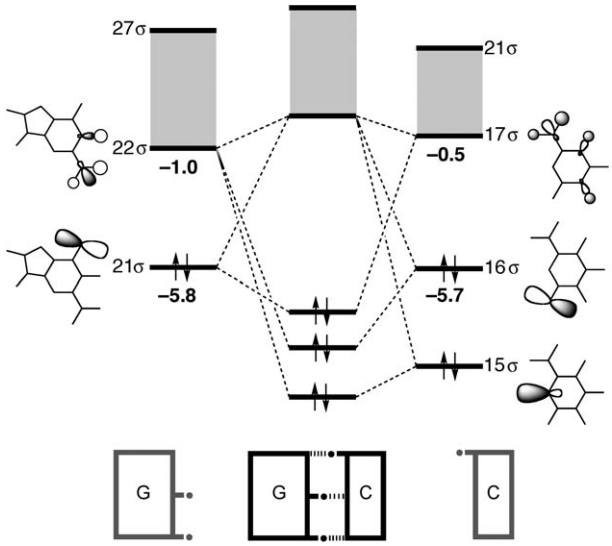


Figure 4. MO interaction diagram for GC Watson–Crick hydrogen bonds (repulsive interactions left out for clarity, see ref. [2c] for a more complete picture; see also Scheme 1).

ref. [2c]). Our analyses of the Kohn–Sham orbital electronic structure of the DNA bases reveal that introducing a neutral amino or hydroxy substituent hardly affects the orbital energies: they are shifted by only a few tenths of an electron volt, similar to what we have found previously for neutral halogen substituents. The picture completely changes if one goes to anionic and cationic substituents, which shift the entire orbital-energy spectrum up and down, respectively, by 3 to 4 eV. Thus, a positively charged substituent causes a sizeable stabilization of the orbitals of the substituted DNA. The stabilization of σ^*_{N-H} acceptor orbitals strengthens the donor–acceptor orbital interaction that occurs in the corresponding hydrogen bond (because this reduces the orbital-energy gap with lone-pair donor orbitals of the other base, see Figure 4). The stabilization of N or O lone-pair orbitals weakens the orbital interactions that occur in the corresponding hydrogen bond (because this increases the orbital-energy gap with σ^*_{N-H} acceptor orbitals of the other base, see Figure 4). Thus, a substitution with NH_3^+ or OH_2^+ at guanine X8 promotes weakening and elongation of O6···H4–N4 and strengthening and contraction of both N1–H1···N3 and N2–H2···O2. The opposite happens in the case of a substitution with NH_3^+ or OH_2^+ at cytosine Y6, which promotes strengthening and contraction of O6···H4–N4 and weakening and elongation of both N1–H1···N3 and N2–H2···O2. A negatively charged substituent causes a destabilization of the orbitals of the substituted DNA base: the destabilization of σ^*_{N-H} acceptor orbitals weakens the donor–acceptor orbital interaction that occurs in the corresponding hydrogen bond (because this increases the orbital-energy gap with lone-pair donor orbitals of the other base, see Figure 4) whereas the destabilization of N or O lone-pair orbitals strengthens the orbital interactions that occur in the corresponding hydrogen bond (because this decreases the orbital-energy gap with σ^*_{N-H} acceptor orbitals of the other base, see Figure 4). Thus, a substitution with NH^- or O^- at guanine X8 promotes strengthening and contraction of O6···H4–N4 and weakening and elongation of both N1–H1···N3 and N2–H2···O2. The opposite happens in the case of a substitution with NH^- or O^- at cytosine Y6, which promotes weakening and elongation of O6···H4–N4 and strengthening and contraction of both N1–H1···N3 and N2–H2···O2.

The changes, described above, in the atomic charges of the front atoms and orbital energies of the DNA bases along the anionic, neutral, and cationic substituents in both N and O series, nicely agree with and explain the observed trends in hydrogen-bond strengths and geometries. But, at first sight, there seems to be a discrepancy with the results of the bond-energy decomposition in Table 2. In particular, in those instances in which the upper O6···H4–N4 hydrogen bond is strengthened and the two other ones (N1–H1···N3 and N2–H2···O2) are weakened, we do not always see reduced electrostatic attraction ΔV_{elstat} and reduced orbital interaction ΔE_{oi}. Instead, these bonding energy terms are often even stabilized and the reduction in overall hydrogen-bond energy ΔE comes from both increased Pauli repulsion ΔE_{Pauli} and preparation (or deformation) energy ΔE_{prep} (see the section on bond-energy analysis for a brief explanation of these terms and ref. [4] for a more extensive survey). For example, from X8 = OH to OH_2^+, the Watson–Crick hydrogen-bond strength ΔE in $G^{X8}C$ increases from −25.94 to −34.60 kcal mol^{-1} which, as expected, originates from a substantial strengthening in both the electrostatic attraction ΔV_{elstat} (from −48.72 to −54.43 kcal mol^{-1}) and the orbital interactions ΔE_{oi} (from −34.77 to −38.77 kcal mol^{-1}, see Table 2). On the other hand, on going from X8 = OH to O^-, the Watson–Crick hydrogen-bond strength ΔE in $G^{X8}C$ decreases from −25.94 to −22.04 kcal mol^{-1}. Here this does not originate from a similar trend in either the electrostatic attraction ΔV_{elstat} or the orbital interactions ΔE_{oi} which, in both cases, is in fact opposite: ΔV_{elstat} is stabilized (from −48.72 to −53.76 kcal mol^{-1}) and so is ΔE_{oi} (from −30.58 to −33.45 kcal mol^{-1}). The destabilization of the overall ΔE is now entirely contained in the strong increase in Pauli repulsion ΔE_{Pauli} (from 52.50 to 68.31 kcal mol^{-1}) and destabilization of the preparation (or geometric deformation) energy ΔE_{prep} (from 4.64 to 11.41 kcal mol^{-1}).

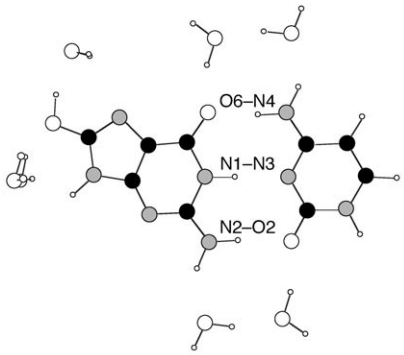
This seemingly counterintuitive result can be understood if one realizes that the effects of introducing ionic substituents are relatively large if compared with the intrinsic strength of the individual hydrogen bonds in the neutral Watson–Crick pair (they are much larger than, for example, the changes that occur along the series X = H, Br, Cl, and F described previously[2h]). This has two consequences. First, the stabilization of an individual hydrogen bond due to the introduction of an ionic substituent is larger than the destabilization of another one. For example, the stabilization of O6···H4–N4, if one goes from $G^{(OH)8}C$ to $G^{(O^-)8}C$, nearly outweighs the combined weakening of N1–H1···N3 and N2–H2···O2, and the overall weakening in ΔE is smaller than the overall strengthening in the case of going from $G^{(OH)8}C$ to $G^{(OH_2^+)8}C$ (see Table 2). Secondly, the substantial changes that the ionic substituents initially cause in the individual components of the hydrogen-bond energy ΔE also induce substantial geometric changes. These changes in the geometry are the result of achieving a new balance between the stabilizing and destabilizing forces and they can and do, in turn, modify the individual components of ΔE. For example, if one goes from $G^{(OH)8}C$ to $G^{(O^-)8}C$, the initial increase in donor–acceptor orbital interactions between the guanine O6 lone-pair orbital 21σ and the cytosine σ^*_{N-H} acceptor orbitals 17σ to 21σ causes the corresponding O6···H4–N4 hydrogen bond to contract and, in addition (see Table 1), it leads to a substantial elongation of the cytosine H4–N4 bond from 1.06 to 1.11 Å (values not shown in Table 1). Note that the latter is indicative of an increased tendency towards proton transfer from cytosine N4 to guanine O6 (which however does not occur in our model systems). Both these geometry-relaxation effects further increase ΔE_{oi} owing to 1) the increase in overlap integrals that goes with hydrogen-bond contraction and 2) the reduction in the HOMO–LUMO gap as the N–H antibonding σ^*_{N-H} acceptor orbitals on cytosine go down in energy due to H4–N4 bond lengthening (values not shown in Table 2). The contraction of the O6···H4–N4

hydrogen bond from $G^{(OH)8}C$ to $G^{(O^-)8}C$ also yields an increase in Pauli repulsion ΔE_{Pauli} (see Table 2). The substantial elongation of the cytosine H4–N4 bond, finally, is associated with (or, occurs at the cost of) an increased preparation or geometric deformation energy ΔE_{prep} (see Table 2). The importance, in general, of accounting for geometric relaxation and the role of achieving balance between repulsive and attractive forces in a bond, was recently pointed out and exemplified in the context of simple organic and inorganic molecules (e.g., C_2H_6, NH_4^+).[15]

The above features in the bonding mechanism lead to yet another interesting phenomenon: an unusually high extent of covalency or orbital interactions in the weakest of a series of Watson–Crick hydrogen bonds. As discussed in the introduction, we have already previously shown that orbital interactions ΔE_{oi} contribute to approximately 40% of all bonding forces (i.e., $\Delta V_{elstat} + \Delta E_{oi}$) and are therefore in the same order of magnitude as electrostatic attraction. Here, we find that the percentage of orbital interactions increases to nearly 50% and, thus, virtually achieves the same magnitude as electrostatic attraction. Interestingly, this unusually high contribution of orbital interactions always occurs in those Watson–Crick pairs that have an anionic substituent at guanine-C8 or a cationic one at cytosine-C6, that is, in those DNA pairs that have a reduced Watson–Crick complexation energy. This finding of a stronger orbital-interaction component in the weaker Watson–Crick pairs falsifies the general idea[1b,c,14] that stronger hydrogen bonds have a larger covalent contribution to the bonding than weaker hydrogen bonds. However, as explained above, in the equilibrium structures of these ionically substituted Watson–Crick pairs, the actual interaction energy ΔE_{int} has been strengthened at the cost of a larger geometric deformation energy ΔE_{prep}. Thus, although the hydrogen-bond strength ΔE in, for example, $G^{(NH^-)8}C$ ($\Delta E = -22.70\,kcal\,mol^{-1}$) is slightly weaker than that in $G^{(NH_2)8}C$ ($\Delta E = -23.96\,kcal\,mol^{-1}$), the actual interaction ΔE_{int} between the deformed DNA bases is, in fact, stronger in the former ($\Delta E_{int} = -33.02\,kcal\,mol^{-1}$) than in the latter ($\Delta E_{int} = -30.38\,kcal\,mol^{-1}$). In this particular example, that is, in $G^{(NH^-)8}C$, the orbital interactions ΔE_{oi} contribute to 47.1% of all bonding forces. The largest percentage-wise contribution of orbital interactions is achieved in $GC^{(OH_2^+)6}$, namely 49.4%.

Force-field approaches that are based on simple pair additive potentials may perform quite satisfactorily over wide ranges of various DNA base pairs, as long as the orbital-interaction term adopts a more or less constant proportion relative to the electrostatic attraction term. Under such circumstances, the effect of orbital interactions can be implicitly described through other terms, in particular, electrostatic terms. Indeed, we have previously found that neutral DNA base pairs (Watson–Crick pairs, mismatches, artificial mimics) usually show an orbital-interaction component of approximately 40% and electrostatic attraction of approximately 60% of all bonding forces.[2] This is also in line with the observation by Sponer et al. that the Cornell et al. force field performs well for various neutral DNA base pairs.[3k] However, if large fluctuations in the relative proportions of orbital and electrostatic interactions occur (for example, from about 40:60 to 50:50%, as in the present series of model systems), one may envisage problems with such force fields.

Solvation effects on the supramolecular switch: The above is a proof of the principle that chemically controlled supramolecular switches can be constructed on the basis of a DNA base pair, GC. Of course, any experimental implementation of this finding will be associated with many (unforeseen) practical problems. Although such an experimental implementation is beyond the scope of this study, we wish to address one major issue, namely, the question whether the effects observed here, computationally, in the gas phase persist under more realistic condensed-phase conditions (for the importance of solvent effects, see for example, ref. [16]). To this end, we have made an attempt to simulate the main features of a $G^{(OH)8}C$ pair being built-in into a DNA sequence under physiological conditions. Under these conditions, the first solvation shell occurs at those positions that are exposed to the solvent: the minor groove and the major groove; an essential position for solvation is the –OH substituent at guanine C8 that, along the O series, changes from $-O^-$ via –OH to $-OH_2^+$. We have included this first solvation shell into our simulation by microsolvating the $G^{X8}C$ switch with seven water molecules, as shown for the neutral $G^{(OH)8}C$ system in Figure 5.



	$G^{O^-}C{\cdot}7H_2O$	$G^{OH}C{\cdot}7H_2O$	$G^{OH_2^+}C{\cdot}7H_2O$
O6–N4	2.67 (2.58)	2.79 (2.72)	2.86 (2.84)
N1–N3	2.99 (2.91)	2.95 (2.88)	2.92 (2.84)
N2–O2	3.09 (3.01)	2.93 (2.87)	2.85 (2.73)

Figure 5. Hydrogen-bond distances (in Å) in microsolvated $G^{X8}C$ pairs (gas-phase values in parentheses). The illustration shows $G^{(OH)8}C{\cdot}7H_2O$.

Three water molecules bind to the –OH substituent in $G^{(OH)8}C$ (and to the $-O^-$ and $-OH_2^+$ substituents in $G^{(O^-)8}C$ and $G^{(OH_2^+)8}C$, not graphically shown in Figure 5), two bind near the O6–N4 hydrogen bond at the major groove, and two bind near the N2–O2 hydrogen bond at the minor groove. The discrete nature of the solvent molecules is fully conserved.

Next, we examine the geometry of the microsolvated GC pair as we proceed along $G^{(O^-)8}C{\cdot}7H_2O$, $G^{(OH)8}C{\cdot}7H_2O$ and $G^{(OH_2^+)8}C{\cdot}7H_2O$. The O6–N4, N1–N3, and N2–O2 hydrogen-bond distances are displayed below the illustration in Figure 5. Note that, for comparison, the gas-phase values are also shown in parentheses. The main difference with respect to the gas-phase values is that the three Watson–Crick bonds are roughly up to one tenth of an angstrom longer in the microsolvated systems. Strikingly, however, the switching is fully conserved under microsolvation conditions. Along $G^{(O^-)8}C{\cdot}7H_2O$, $G^{(OH)8}C{\cdot}7H_2O$, and $G^{(OH_2^+)8}C{\cdot}7H_2O$, the O6–N4 bond expands from 2.67 to 2.79 to 2.86 Å, the N1–N3 bond only slightly contracts from 2.99 to 2.95 to 2.92 Å, and the N2–O2 bond contracts from 3.09 to 2.93 to 2.85 Å (see Figure 5). Thus, the changes in the O6–N4, N1–N3, and N2–O2 hydrogen-bond distances, from anionic to cationic DNA base pair, are almost as large under microsolvation conditions (i.e., +0.19, −0.07, −0.24 Å, respectively) as in the pure gas phase (i.e., +0.26, −0.07, −0.28 Å, respectively). These results suggest that the switching persists in the condensed phase and that supramolecular switches based on DNA base pairs are likely a feasible target in experimental studies.

Conclusions

Our DFT computations show that it is possible to build a supramolecular switch, based on the DNA base pair GC, that can be chemically switched between three states that differ in hydrogen-bond strength (weak, intermediate, strong) and geometrical shape (e.g., GC bending in-plane "up", "not", and "down"). The chemical switching involves deprotonation or protonation of hydroxy and amino substituents at guanine C8 and cytosine C6. This behavior has been shown to persist under microsolvation conditions.

The substituent effects can be understood in terms of how they modify the electronic structure of the DNA bases and thus the electrostatic and donor–acceptor orbital interactions that provide Watson–Crick hydrogen bonding. Introducing a neutral substituent (OH or NH_2) has relatively small effects. A charged substituent, on the other hand, leads to substantial and characteristic changes in hydrogen-bond lengths, strengths, and bonding mechanism: an anionic substituent (O^-, NH^-) reduces hydrogen-bond-donating and increases the hydrogen-bond-accepting capabilities of a DNA base (amongst others, by pushing up all orbital energies), and vice versa for a cationic substituent (OH_2^+, NH_3^+).

The orbital-interaction component in some of these hydrogen bonds is found to contribute to more than 49% of the attractive interactions and is thus virtually equal in magnitude as the electrostatic component, which provides the other somewhat less than 51% of the attraction. Interestingly, such increased covalent character occurs in modified GC base pairs that are less stable than natural (unsubstituted) GC. This contradicts the textbook knowledge[1b,c] that weaker hydrogen bonds have less covalent character than stronger ones.

Acknowledgements

We thank the National Research School Combination for Catalysis (NRSC-C) for a postdoctoral fellowship for C.F.G. and the Netherlands organization for Scientific Research (NWO-CW and NWO-NCF) for financial support. We also thank the referees for valuable suggestions.

[1] a) L. Stryer, *Biochemistry*, W. H. Freeman and Company, New York, **1988**; b) G. A. Jeffrey, W. Saenger, *Hydrogen Bonding in Biological Structures*, Springer, Berlin, **1991**; c) G. A. Jeffrey, *An Introduction to Hydrogen Bonding*, Oxford University Press, New York, **1997**.

[2] a) C. Fonseca Guerra, F. M. Bickelhaupt, *Angew. Chem.* **1999**, *111*, 3120; *Angew. Chem. Int. Ed.* **1999**, *38*, 2942; b) C. Fonseca Guerra, F. M. Bickelhaupt, J. G. Snijders, E. J. Baerends, *J. Am. Chem. Soc.* **2000**, *122*, 4117; c) C. Fonseca Guerra, F. M. Bickelhaupt, J. G. Snijders, E. J. Baerends, *Chem. Eur. J.* **1999**, *5*, 3581; d) C. Fonseca Guerra, E. J. Baerends, F. M Bickelhaupt, *Cryst. Growth Des.* **2002**, *2*, 239; e) C. Fonseca Guerra, F. M. Bickelhaupt, *Angew. Chem.* **2002**, *114*, 2194; *Angew. Chem. Int. Ed.* **2002**, *41*, 2092; f) C. Fonseca Guerra, F. M. Bickelhaupt, *J. Chem. Phys.* **2003**, *119*, 4262; g) C. Fonseca Guerra, F. M. Bickelhaupt, E. J. Baerends, *ChemPhysChem* **2004**, *5*, 481; h) C. Fonseca Guerra, T. van der Wijst, F. M. Bickelhaupt, *Struct. Chem.* **2005**, *16*, 211.

[3] a) P. Hobza, J. Sponer, *Chem. Rev.* **1999**, *99*, 3247; b) J. Bertran, A. Oliva, L. Rodríguez-Santiago, M. Sodupe, *J. Am. Chem. Soc.* **1998**, *120*, 8159; c) K. Brameld, S. Dasgupta, W. A. Goddard III, *J. Phys. Chem. B* **1997**, *101*, 4851; d) J. Sponer, J. Leszczynski, P. Hobza, *J. Phys. Chem.* **1996**, *100*, 1965; e) I. R. Gould, P. A. Kollman, *J. Am. Chem. Soc.* **1994**, *116*, 2493; f) R. Santamaria, A. Vázquez, *J. Comput. Chem.* **1994**, *15*, 981; g) J. Sponer, P. Hobza, *J. Phys. Chem. A* **2000**, *104*, 4592; h) P. Hobza, J. Sponer, E. Cubero, M. Orozco, F. J. Luque, *J. Phys. Chem. B* **2000**, *104*, 6286; i) J. Poater, X. Fradera, M. Solà, M. Duran, S. Simon, *Chem. Phys. Lett.* **2003**, *369*, 248; j) V. I. Danilov, V. M. Anisimov, *J. Biomol. Struct. Dyn.* **2005**, *22*, 471; k) J. Sponer, P. Jurecka, P. Hobza, *J. Am. Chem. Soc.* **2004**, *126*, 10142.

[4] F. M. Bickelhaupt, E. J. Baerends in *Rev. Comput. Chem., Vol. 15* (Eds.: K. B. Lipkowitz, D. B. Boyd) Wiley, New York, **2000**, pp. 1–86.

[5] a) G. te Velde, F. M. Bickelhaupt, S. J. A. van Gisbergen, C. Fonseca Guerra, E. J. Baerends, J. G. Snijders, T. Ziegler, *J. Comput. Chem.* **2001**, *22*, 931; b) C. Fonseca Guerra, O. Visser, J. G. Snijders, G. te Velde, E. J. Baerends in *Methods and Techniques for Computational Chemistry* (Eds.: E. Clementi, G. Corongiu), STEF, Cagliari, **1995**, pp. 305–395; c) E. J. Baerends, D. E. Ellis, P. Ros, *Chem. Phys.* **1973**, *2*, 41; d) E. J. Baerends, P. Ros, *Chem. Phys.* **1975**, *8*, 412; e) E. J. Baerends, P. Ros, *Int. J. Quantum Chem., Quantum Chem. Symp.* **1978**, *12*, 169; f) C. Fonseca Guerra, J. G. Snijders, G. te Velde, E. J. Baerends, *Theor. Chem. Acc.* **1998**, *99*, 391; g) P. M. Boerrigter, G. te Velde, E. J. Baerends, *Int. J. Quantum Chem.* **1988**, *33*, 87; h) G. te Velde, E. J. Baerends, *J. Comp. Phys.* **1992**, *99*, 84; i) J. G. Snijders, E. J. Baerends, P. Vernooijs, *At. Data Nucl. Data Tables* **1982**, *26*, 483; j) J. Krijn, E. J. Baerends, *Fit-Functions in the HFS-Method; Internal Report (in Dutch)*, Vrije Universiteit, Amsterdam, **1984**; k) L. Versluis, T. Ziegler, *J. Chem. Phys.* **1988**, *88*, 322; l) J. C. Slater, *Quantum Theory of Molecules and Solids, Vol. 4*, McGraw-Hill, New York, **1974**; m) A. D. Becke, *J. Chem. Phys.* **1986**, *84*, 4524; n) A. Becke, *Phys. Rev. A* **1988**, *38*, 3098; o) S. H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **1980**, *58*, 1200; p) J. P. Perdew, *Phys. Rev. B* **1986**, *33*, 8822 (erratum: *Phys. Rev. B* **1986**, *34*, 7406); q) L. Fan, T. Ziegler, *J. Chem. Phys.* **1991**, *94*, 6057.

[6] See, for example: F. Vögtle, *Supramolecular Chemistry*, Wiley, Chichester, **1993**.

[7] See, for example: *Antisense Research and Application* (Ed.: S. T. Crooke), Springer, Berlin **1998**.
[8] a) S.-I. Kawahara, T. Uchimaru, *Eur. J. Org. Chem.* **2003**, 2577; b) S.-I. Kawahara, T. Uchimaru, K. Taira, M. Sekine, *J. Phys. Chem. A* **2002**, *106*, 3207; c) S. I. Kawahara, A. Kobori, M. Sekine, K. Taira, T. Uchimaru, *J. Phys. Chem. A* **2001**, *105*, 10596; d) S.-I. Kawahara, T. Uchimaru, K. Taira, M. Sekine, *J. Phys. Chem. A* **2001**, *105*, 3894; e) S.-I. Kawahara, T. Wada, S. Kawauchi, T. Uchimaru, M. Sekine, *J. Phys. Chem. A* **1999**, *103*, 8516.
[9] a) K. Morokuma, *J. Chem. Phys.* **1971**, *55*, 1236; b) K. Kitaura, K. Morokuma, *Int. J. Quantum. Chem.* **1976**, *10*, 325.
[10] a) T. Ziegler, A. Rauk, *Inorg. Chem.* **1979**, *18*, 1755; b) T. Ziegler, A. Rauk, *Inorg. Chem.* **1979**, *18*, 1558; c) T. Ziegler, A. Rauk, *Theor. Chim. Acta* **1977**, *46*, 1.
[11] a) A. J. Stone, *The Theory of Intermolecular Forces*, Clarendon Press, Oxford, **1996**; b) A. J. Stone, *Chem. Phys. Lett.* **1993**, *211*, 101.
[12] a) C. Fonseca Guerra, J.-W. Handgraaf, E. J. Baerends, F. M. Bickelhaupt, *J. Comput. Chem.* **2004**, *25*, 189; b) F. M. Bickelhaupt, N. J. R. van Eikema Hommes, C. Fonseca Guerra, E. J. Baerends, *Organometallics* **1996**, *15*, 2923.
[13] a) G. F. Voronoi, *Z. Reine Angew. Math.* **1908**, *134*, 198; b) C. Kittel, *Introduction to Solid State Physics*, Wiley, New York, **1986**.
[14] H. Umeyama, K. Morokuma, *J. Am. Chem. Soc.* **1977**, *99*, 1316.
[15] a) F. M. Bickelhaupt, E. J. Baerends, *Angew. Chem.* **2003**, *115*, 4315; *Angew. Chem. Int. Ed.* **2003**, *42*, 4183; b) F. M. Bickelhaupt, R. L. DeKock, E. J. Baerends, *J. Am. Chem. Soc.* **2002**, *124*, 1500.
[16] a) J. E. Sponer, J. Leszczynski, F. Glahé, B. Lippert, J. Sponer, *Inorg. Chem.* **2001**, *40*, 3269; b) J. V. Burda, J. Sponer, J. Hrabakova, M. Zeizinger, J. Leszczynski, *J. Phys. Chem. B* **2003**, *107*, 5349; c) F. M. Bickelhaupt, E. J. Baerends, N. M. M. Nibbering, *Chem. Eur. J.* **1996**, *2*, 196.

Received: October 20, 2005
Published online: February 2, 2006

