Adenine Tautomers: Relative Stabilities, Ionization Energies, and Mismatch with Cytosine

C. Fonseca Guerra,*,† F. M. Bickelhaupt,*,† S. Saha,‡ and F. Wang‡

Theoretische Chemie, Scheikundig Laboratorium der Vrije Universiteit, De Boelelaan 1083, NL-1081 HV Amsterdam, The Netherlands, and Centre for Molecular Simulation, Swinburne University of Technology, P.O. Box 218, Hawthorn, Melbourne, Victoria, 3122, Australia

Received: December 13, 2005; In Final Form: January 24, 2006

In this study, we have investigated 12 tautomers of the DNA base adenine at the BP86/TZ2P and BP86/QZ4P levels of density functional theory. The vertical and adiabatic ionization energies of all tautomers were determined as the difference in energy between the radical cation and the corresponding neutral system. Furthermore, an evaluation is made for the eigenvalue spectra calculated with the SAOP functional, which is shown to lead to substantial improvements for orbital energies compared to BP86. We have also explored the correlations between the Kohn−Sham orbitals of the different tautomers at the BP86/QZ4P and SAOP/QZ4P levels. Finally, we discuss implications of the existence of the tautomeric forms of adenine for the DNA replication.

1. Introduction

Rare tautomeric forms of DNA bases are suspected to play a key role in the formation of spontaneous mutations in DNA.[1−3] A DNA base can hydrogen bond with a complementary tautomeric form of the incorrect partner base giving a proper Watson−Crick shape. This mispairing can, therefore, be incorporated into the DNA strand without disturbing the geometrical and energetic demands of the DNA double helix on the base pair formation.[1c,4,5] This mutagenic process can be aided by metal ions, which stabilize certain tautomeric forms of the DNA bases.[6] The coexistence of these different tautomeric forms of DNA bases has been established in spectroscopic experiments in the gas phase [7] and in solution with NMR experiments.[8]

In this work, we focus on the DNA base adenine (ade_n9) which can exist in twelve different tautomeric forms by shifting one or two protons to a different nitrogen atom (see Scheme 1). We have computed the geometries and relative energies of the DNA base adenine and eleven tautomers thereof using the generalized gradient approximation of density functional theory (DFT) at the BP86/TZ2P and BP86/QZ4P levels of theory. Recently, Fonseca Guerra et al.[9] have shown that the BP86/TZ2P level of DFT leads to excellent agreement of computed structures and bond energies of DNA base pairs with experimental values.

Furthermore, we have computed (again at BP86/TZ2P and BP86/QZ4P) the vertical and adiabatic first ionization energies of all twelve tautomers as the difference in energy between the radical cation and the corresponding neutral system. These vertical ionization energies are compared with the orbital eigenvalue spectra of each of the four neutral tautomers (ade_n1, ade_n3, ade_n7, and ade_n9) which, in a one-electron model, are to be interpreted also as vertical ionization energies. The orbital energies are computed with the BP86 potential, which is known to yield too weakly bound electrons, as well as the SAOP potential, which is know to lead to substantial improvements for orbital energies compared to BP86.[10] Finally, we have also explored the correlations between the Kohn−Sham orbitals of the different tautomers (both at BP86/QZ4P and at SAOP/QZ4P), i.e., the relationships in terms of spatial character of the one-electron wave functions if one goes from one tautomer to another one.[11]

The last issue addressed in this paper is the possible implication for DNA replication when the tautomeric form imi_1_n1_n9 pairs with cytosine. The hydrogen bond lengths and strengths are calculated and compared to the former values obtained for the natural Watson−Crick pair AT.[12]

2. Methods

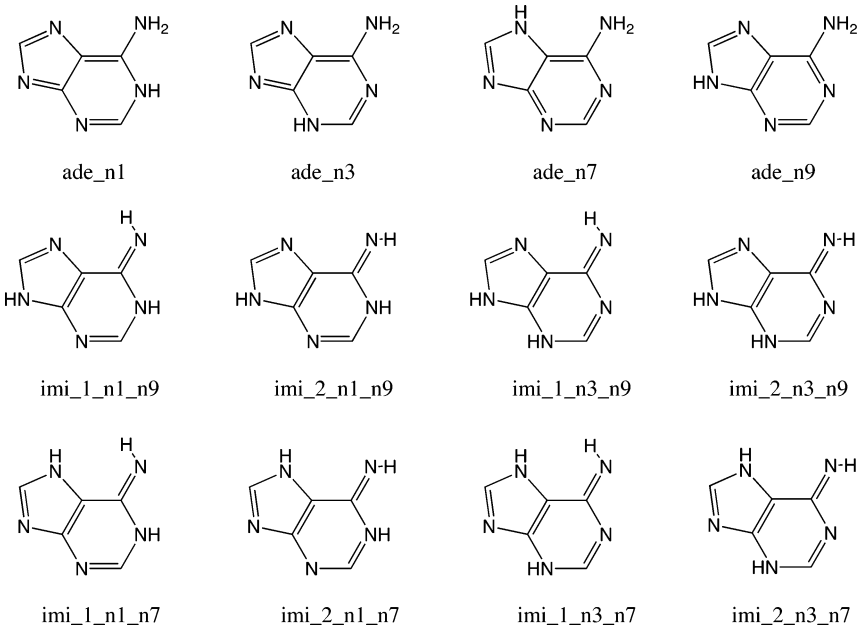
2.1. General Procedure. All calculations were performed using the Amsterdam Density Functional (ADF) program,[13] except for the OVGF/TZVP calculations,[14] which employ a Gaussian-type basis set and the *Gaussian 03* program.[15] The MOs in the ADF calculations were expanded using two different uncontracted sets of Slater-type orbitals (STOs) containing diffuse functions (no Gaussian functions are involved): (i) the TZ2P, which is of triple-ζ quality for all atoms and has been augmented with two sets of polarization functions on each atom, i.e., 3d and 4f on C and N, and 2p and 3d on H; and (ii) the QZ4P basis, which is of quadruple-ζ quality for all atoms and has been augmented with four sets of polarization functions on each atom, i.e., two 3d and two 4f STOs on C and N, and two 2p and two 3d STOs on H. The 1s core shells of carbon and nitrogen were treated by the frozen-core approximation. 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-correlation potentials accurately in each self-consistent field cycle.

Geometries and energies were calculated using the generalized gradient approximation (GGA) to DFT at the Becke88−Perdew86 (BP86) level of theory.[13m,p] Open-shell species were treated with spin-unrestricted formalism. The C_1 geometries of the adenine (ade_n9) and the tautomers ade_n1, ade_n3, and ade_n7, which have an amino group that pyramidalizes, and the C_s geometries of the imino tautomers (see Scheme 1) have

* Fax: +31-20-59 87629. Email addresses: c.fonseca.guerra @few.vu.nl; fm.bickelhaupt@few.vu.nl.
† Scheikundig Laboratorium der Vrije Universiteit.
‡ Swinburne University of Technology.

10.1021/jp057275r CCC: $33.50
Published on Web 03/02/2006

SCHEME 1: Adenine Tautomers Including Canonical Adenine (ade_n9)



been verified to be equilibrium structures (zero imaginary frequencies) through a vibrational analysis at the BP86/TZ2P level.

2.2. Bond Energy Analysis. The overall bond energy ΔE is made up of two major components (eq 1). In this formula, the

$$\Delta E = \Delta E_{\text{prep}} + \Delta E_{\text{int}} \tag{1}$$

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 molecular orbital (MO) model using a decomposition of the bond into electrostatic interaction, exchange repulsion (or Pauli repulsion), and (attractive) orbital interactions (eq 2).[4,16,17] The term ΔV_{elstat} corresponds to the classical electro-

$$\Delta E_{\text{int}} = \Delta V_{\text{elstat}} + \Delta E_{\text{Pauli}} + \Delta E_{\text{oi}} \tag{2}$$

static 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). Since 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 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,[16] 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[18] that also give rise to induction and dispersion contributions, which we do not attempt to quantify but which are all lumped together 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.[17] Note that our

$$\Delta E_{\text{oi}} = \Sigma_{\Gamma} \Delta E_{\Gamma} \tag{3}$$

approach differs in this respect from the Morokuma scheme,[16] 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.

3. Results and Discussion

The results of our computations are collected in Tables 1 (relative energies), 2 (dipole moments), 3 (dihedral angle of the amino group), and 4 and 5 (ionization energies), and in Figures 1 (structures), 2 (orbital-level correlation), and 3−5 (3D orbital plots). Furthermore, the Supporting Information provides Cartesian coordinates of the equilibrium structures of all species involved in Tables S1−S4 and orbital energy values of selected species in Tables S5 and S6.

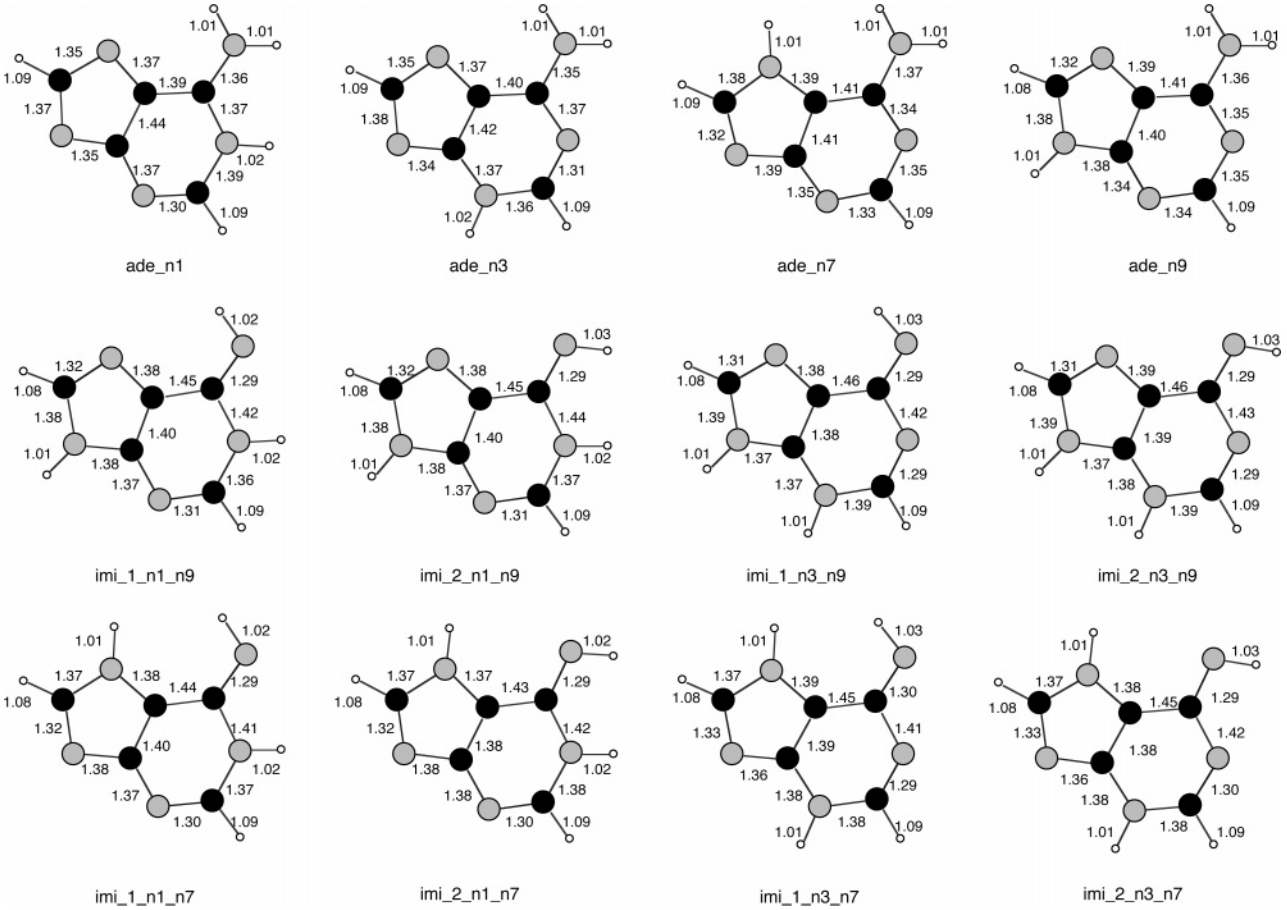


Figure 1. Geometries (in Å) of C_s symmetric adenine tautomers at BP86/QZ4P.

TABLE 1: Relative Energies (kcal/mol) Without and With Zero-Point Energy Corrections of Adenine Tautomers with Respect to C_1-Symmetric Natural Adenine (ade_n9) Computed at BP86/TZ2P and BP86/QZ4P

	BP86/TZ2P				BP86/QZ4P	
	$E(C_s)^a$	$E_{ZPE}(C_s)^a$	$E(C_1)^b$	$E_{ZPE}(C_1)^b$	$E(C_s)^a$	$E(C_1)^b$
ade_n1	17.7		17.4	17.3	17.3	16.9
ade_n3	7.0		7.0	7.1	6.8	6.8
ade_n7	8.3		7.8	7.7	7.9	7.5
ade_n9	0.0		0.0[c]	0.0	0.0	0.0[d]
imi_1_n1_n7	14.9	15.2			14.3	
imi_2_n1_n7	14.7	15.0			14.2	
imi_1_n1_n9	10.9	11.4			10.6	
imi_2_n1_n9	17.0	17.1			16.5	
imi_1_n3_n7	22.2	22.2			21.5	
imi_2_n3_n7	15.7	15.5			15.3	
imi_1_n3_n9	29.8	29.3			29.1	
imi_2_n3_n9	29.8	29.2			29.2	

[a] Geometry optimized with C_s symmetry imposed. [b] Geometry optimized in C_1 symmetry, i.e., without any symmetry restrictions. [c] Anchor point for relative energies at BP86/TZ2P. [d] Anchor point for relative energies at BP86/QZ4P.

3.1. Structures and Relative Stabilities of Adenine Tautomers. The natural adenine (ade_n9) is the most stable tautomer of the twelve systems studied (see Table 1). The structures of the tautomers is given in Figure 1. The following order from low to higher energy (and also ZPE-corrected energy) is found both at BP86/TZ2P and BP86/QZ4P: ade_n9 < ade_n3 < ade_n7 < imi_1_n1_n9 < imi_2_n1_n7 < imi_1_n1_n7 < imi_2_n3_n7 < imi_2_n1_n9 < ade_n1 < imi_1_n3_n7 < imi_1_n3_n9/imi_2_n3_n9. This ordering corresponds well with the ordering found in the interesting study by Hanus et al.[2a] at the RI-MP2/TZVPP level. The differences in the ordering encountered are due to the very small differences in energy between the tautomers themselves (for instance, the energy difference between ade_n3 and ade_n7 is only 0.7 kcal/mol at the BP86/QZ4P level of theory). Plützer and Kleinermanns[7a] showed that the conventional IR absorption spectrum of adenine in the gas phase (with a cell temperature of 280 °C) is a superposition of the IR spectra of the ade_n9 and ade_n7 tautomers. Under such experimental conditions, the coexistence of the ade_n3 tautomer (and maybe also of imi_1_n1_n9) might be feasible as far as the thermodynamics are concerned.

TABLE 2: Absolute Values of the Dihedral Angles (deg) of the Amino Group of the Adenine Tautomers in C_1 Symmetry (BP86/QZ4P)

	∠H6N6C6N1[a]	∠H6′N6C6C5[a]
ade_n1	32.5	11.1
ade_n3	0.2	0.1
ade_n7	13.7	30.8
ade_n9	6.1	6.3

[a] See also Scheme 2.

Furthermore, we can see that, for ade_n1 and ade_n7, there is a difference in energy between the C_s and the C_1 symmetric structures (see Table 1). This energy difference can be ascribed to the larger pyramidalization of the amino group in these tautomers: for ade_n1, ∠H6N6C6N1 is 32.5°, and for ade_n7, ∠H6′N6C6C5 is 30.8°, whereas in adenine (ade_n9), these angles are, respectively, 6.1° and 6.3° (see Table 2). The close contact between the hydrogen atoms of the amino group and shifted proton at the N7 and the N1 positions causes the amino group to pyramidalize (see Scheme 2 which shows canonical adenine, i.e., ade_n9, with complete atom numbering). This phenomenon has already been observed in our previous work[9b]

SCHEME 2: Canonical Adenine (ade_n9) with Atom Numbering

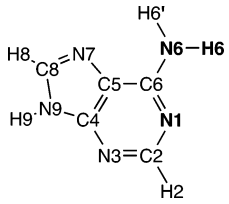


TABLE 3: Dipole Moments of Adenine Tautomers in C_1 and C_s Symmetry Computed at BP86/TZ2P and BP86/QZ4P

	BP86/TZ2P		BP86/QZ4P	
	C_s [a]	C_1 [b]	C_s [a]	C_1 [b]
ade_n1	9.004	8.437	8.849	8.377
ade_n3	3.968	3.966	3.937	3.937
ade_n7	7.503	7.001	7.319	6.913
ade_n9	2.427	2.454	2.332	2.347
imi_1_n1_n7	3.139		3.000	
imi_2_n1_n7	3.465		3.290	
imi_1_n1_n9	3.800		3.803	
imi_2_n1_n9	4.505		4.448	
imi_1_n3_n7	4.661		4.578	
imi_2_n3_n7	2.848		2.901	
imi_1_n3_n9	9.882		9.775	
imi_2_n3_n9	9.404		9.328	

[a] Geometry optimized with C_s symmetry imposed. [b] Geometry optimized in C_1 symmetry, i.e., without any symmetry restrictions.

on the natural DNA base pairs. In guanine, the pyramidalization of the amino group is larger due to the presence of a neighboring N−H group than in cytosine and canonical adenine, where the amino group has no neighboring N−H group. The hydrogen atom of the amino group (in guanine, ade_n1 and ade_n7) that is closer to such an N−H bond bends down more than the hydrogen atom that is further away from this bond. Poater et al.[19] have shown that, in biphenyl, hydrogen atoms avoid such close contacts due to a repulsive interaction.

The dipole moments are collected in Table 3. The canonical base has the smallest dipole moment, and imi_1_n3_n9 has the largest dipole moments. This is consistent with the results obtained by Hanus et al.[2b] and Preuss et al.[20]

3.2. Ionization Energies of Adenine Tautomers. We have also computed (at BP86/TZ2P and BP86/QZ4P) the vertical and adiabatic first ionization energies of twelve tautomers as the difference in energy between the radical cation and the corresponding neutral system (see Table 4). These vertical ionization energies are compared with the orbital eigenvalue spectra at the BP86/QZ4P and SAOP/QZ4P of each of the four neutral amino tautomers which, in a one-electron model, are to be interpreted also as vertical ionization energies (see Table 5).

The calculated values for the vertical and adiabatic first ionization energies for ade_n9 are 8.27 and 8.11 eV at the BP86/TZ2P level and 8.18 and 8.03 eV at the BP86/QZ4P level (see Table 4). These values deviate no more than 0.3 eV from the experimental values, which are 8.48 eV for the vertical ionization energy of ade_n9 and 8.26 eV for the adiabatic ionization energy.[21] They are also in good agreement with the MP2/6-31++G(d,p) results of 8.62 and 8.23 eV by Crespo-Hernández et al.[22a]

Furthermore, inspection of the vertical ionization energies calculated as the orbital energies (see Table 5) shows that the BP86 results deviate about 3 eV from the vertical first ionization energy calculated as the difference in energy between the radical cation and the corresponding neutral system (see Table 4). The

TABLE 4: The First Vertical and Adiabatic Ionization Energies (eV) of Adenine Tautomers

	BP86/TZ2P[a]		BP86/QZ4P[b]		OVGF/TZVP[c]
	vertical	adiabatic	vertical	adiabatic	vertical
ade_n1	8.2950	8.0732	8.2100	7.9943	8.062 (0.897)
ade_n3	8.2787	8.0852	8.1968	8.0085	8.045 (0.897)
ade_n7	8.4777	8.3046	8.3881	8.2203	8.255 (0.898)
ade_n9	8.2689	8.1075	8.1846	8.0274	8.197 (0.898)
imi_1_n1_n7	8.1366	7.9998	8.4714	7.9258	7.800 (0.893)
imi_2_n1_n7	8.1094	7.9732	8.0636	7.8977	7.915 (0.894)
imi_1_n1_n9	7.9570	7.8306	7.8830	7.7607	7.757 (0.894)
imi_2_n1_n9	7.9439	7.8096	7.8704	7.7403	7.746 (0.895)
imi_1_n3_n7	8.0348	7.9143	7.9618	7.8455	7.677 (0.891)
imi_2_n3_n7	7.9709	7.8515	7.8973	7.7830	7.610 (0.892)
imi_1_n3_n9	7.8907	7.7621	7.8188	7.6953	7.556 (0.893)
imi_2_n3_n9	7.8437	7.7094	7.7727	7.6447	7.496 (0.893)

[a] Computed at BP86/TZ2P//BP86/TZ2P. [b] Computed at BP86/QZ4P//BP86/QZ4P. [c] Computed at OVGF/TZVP//BP86/QZ4P (values in parentheses are the pole strengths).

TABLE 5: Vertical Ionization Energies (in eV) for the Outer Valence Space of the Amino Adenine Tautomers at BP86/QZ4P and SAOP/QZ4P[a]

ade_n1		ade_n3		ade_n7		ade_n9		
BP86	SAOP	BP86	SAOP	BP86	SAOP	BP86	SAOP	exp[21a]
9.52	13.47	9.74	13.68	10.05	14.03	(10.00)	(14.05)	(13.21)
(9.41)	(13.45)	(9.60)	(13.64)	(8.95)	(13.05)	(8.86)	(12.93)	(12.10)
7.70	11.90	(7.33)	(11.44)	(7.40)	(11.56)	7.61	11.81	11.39
(7.61)	(11.73)	7.12	11.28	7.26	11.49	(7.32)	(11.46)	(10.5)
(6.13)	(10.23)	6.96	11.09	6.66	10.88	6.77	10.94	10.5
6.06	10.21	(6.21)	(10.33)	(6.63)	(10.71)	(6.70)	(10.78)	(9.6)
5.78	9.96	6.02	10.20	5.85	10.05	5.91	10.08	9.6
(5.55)	(9.58)	(5.54)	(9.58)	(5.73)	(9.86)	(5.55)	(9.65)	(8.48)

[a] π vertical ionization energies in parentheses and σ vertical ionization energies without.

orbital energies calculated at SAOP/QZ4P (Table 5) agree much better with the vertical first ionization energy calculated as the difference in energy between the radical cation and the corresponding neutral system (Table 4). Comparison of the first eight vertical ionization energies of ade_n9 at the BP86/QZ4P and SAOP/QZ4P with the experimental values[21a] shows for BP86 a discrepancy of about 3 eV and for SAOP about 0.5−1.5 eV: somewhat more than the OVGF vertical ionization energies by Dolgounitcheva et al. [22b] However, both BP86 and SAOP give the same ordering of σ and π orbitals as proposed by Lin et al.[21a] For the first ionization energies calculated with the orbital energies at the SAOP level, Chong et al.[10a] also found deviations on the order of 1 eV for related systems as pyridine, but they suggested that improvements can be made with a refined approximation to the exact Kohn−Sham potential.

3.3. Resemblances and Differences of Molecular Orbitals. The energies of the molecular orbitals have been calculated for the tautomers ade_n1, ade_3, ade_n7, and ade_n9 and the imino forms imi_1_n1_n9 and imi_1_n3_n9 at the BP86/QZ4P and SAOP/QZ4P. The tautomeric form imi_1_n1_n9 is a possible candidate to pair with cytosine in a DNA strand, giving possibilities for mutations (see section 3.4). In the discussion above, we had established a better agreement for the SAOP/QZ4P orbitals with the first ionization energies. However, as the BP86 functional has been widely used, it is interesting to see in which respect the orbitals and their energies calculated with the BP86 and SAOP functional differ. Figure 2 represents the energies of the molecular orbitals at the two levels (in Figure 2a BP86/QZ4P and in Figure 2b SAOP/QZ4P). The main difference is that, in comparison to the BP86 orbitals, the SAOP orbitals are consistently shifted down in energy by about 4 eV.

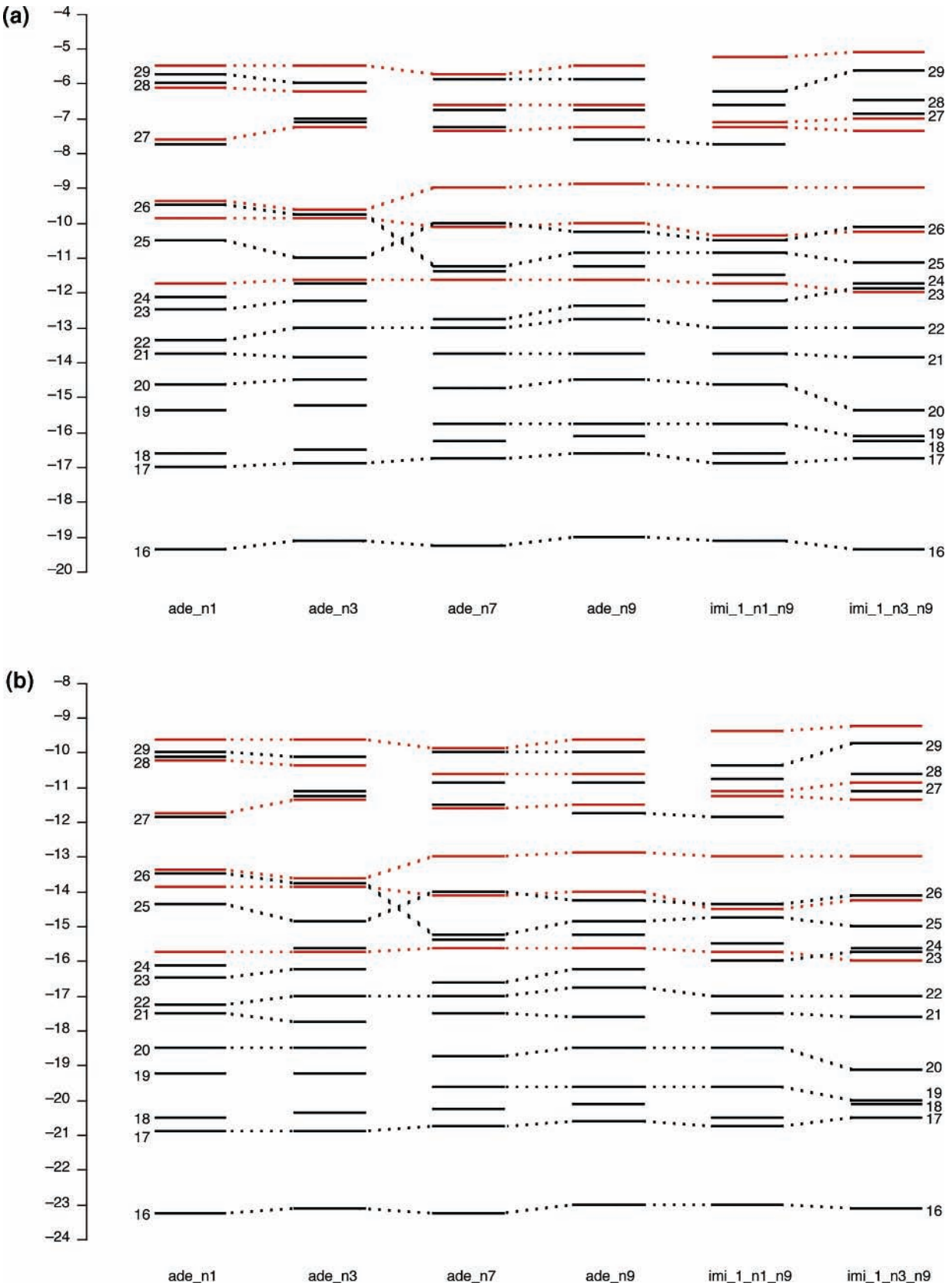


Figure 2. (a) Energies (in eV) of the molecular orbitals of adenine tautomers at BP86/QZ4P: σ orbitals in black, π orbitals in red. Dotted lines have only been drawn between orbitals that resemble each other. (b) Energies (in eV) of the molecular orbitals of adenine tautomers at SAOP/QZ4P//BP86/QZ4P: σ orbitals in black, π orbitals in red. Dotted lines have only been drawn between orbitals that resemble each other.

Furthermore, when we analyze the shape and spatial character of the orbitals (see Figures 3−5), we see negligible differences between the BP86 orbitals and the SAOP orbitals (see Supporting Information for the SAOP orbitals). Therefore, in the discussion below, when we refer to a certain orbital, it applies to the SAOP orbitals as well as to the BP86 orbitals.

Next, we examine the spatial shape of the σ and π orbitals of the various adenine tautomers. The π orbitals that, due to

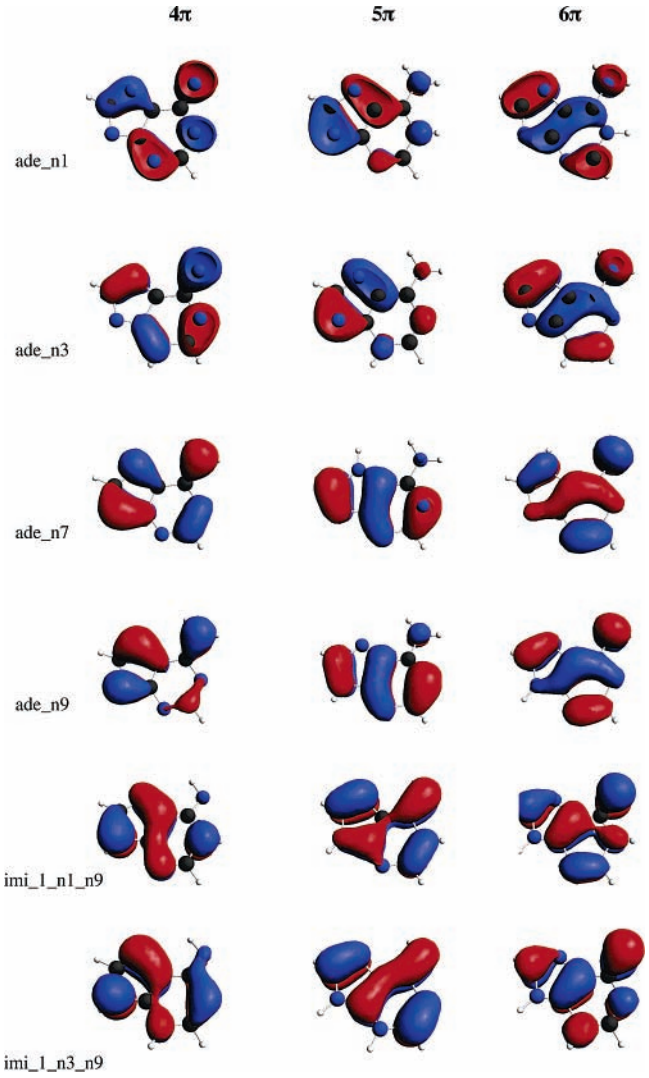


Figure 3. The molecular orbitals 4π, 5π, and 6π of the tautomers ade_n1, ade_n3, ade_n7, ade_n9, imi_1_n1_n9, and imi_1_n3_n9 calculated at the BP86/QZ4P level.

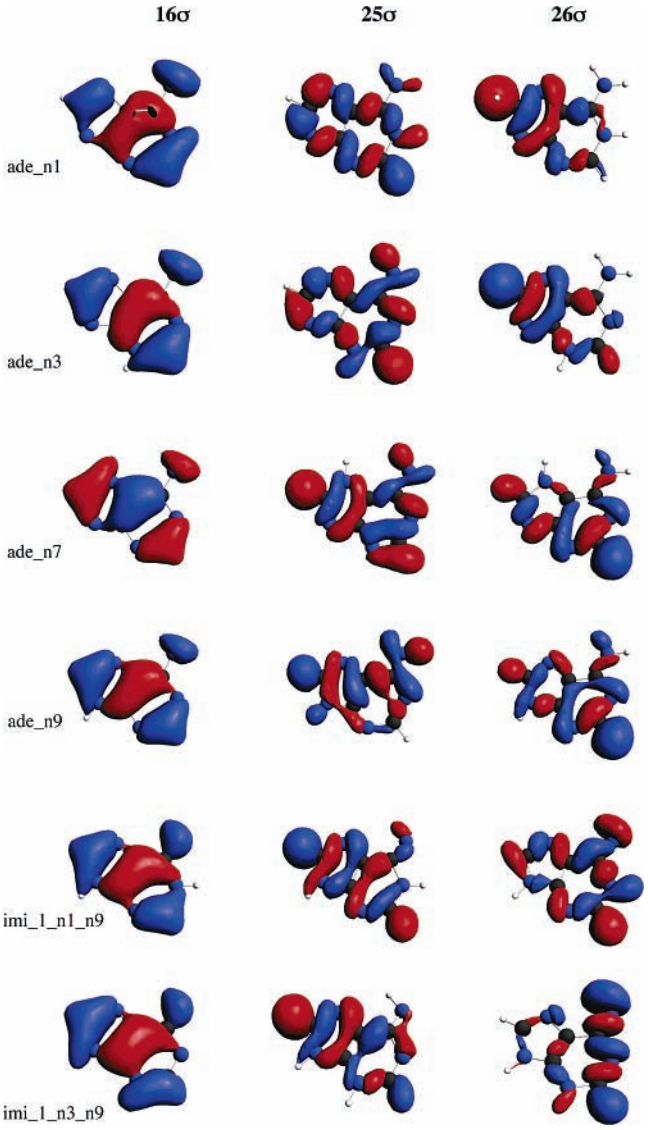


Figure 4. The molecular orbitals 16σ, 25σ, and 26σ of the tautomers ade_n1, ade_n3, ade_n7, ade_n9, imi_1_n1_n9, and imi_1_n3_n9 calculated at the BP86/QZ4P level.

the planar structure of the tautomers, are essentially linear combinations of the $2p_z$ atomic orbitals on the carbon and nitrogen atoms, have large resemblances between the different tautomers (z-axis defined as the axis perpendicular to the molecular plane). The energetically lowest π orbital (1π) has for all tautomers only one nodal plane through the molecular plane and no nodal plane perpendicular to the plane defined by the atoms. 2π and 3π orbitals have two nodal planes: one through the plane of the molecule and one perpendicular to it. The differences between the amino tautomers of adenine (ade_n1, ade_n3, ade_n7, and ade_n9) and the imino tautomers (see Scheme 1) become visible in orbitals 4π and 5π. All tautomers have three nodal planes for these orbitals: one through the plane of the molecule and two perpendicular to it. However, for the tautomers ade_n1, ade_n3, ade_n7, and ade_n9, these two nodal planes cross each other, whereas for the imino form, the nodal planes are parallel to each other. The resemblances of the 4π orbitals divides the tautomers into three groups: the systems with the proton on the six-membered ring (ade_n1 and ade_n3), the systems with the proton on the five-membered ring (ade_n7 and ade_n9), and the imino form. The next orbital, 5π, has again three nodal planes, one of which coincides with the molecular plane. This 5π orbital divides molecules into three groups also. The systems ade_n1 and ade_n3, the tautomers with the proton on the six-membered ring, have the two nodal planes of 5π perpendicular to the molecular plane crossing each other, whereas for ade_n7 and ade_n9, these nodal planes lie parallel to each other. The imino forms also have the nodal planes of 5π almost parallel to each other, but the orbital extends over some other atoms than the 5π of ade_n7 and ade_n9 (see Figure 3). The HOMO (6π) is very similar for the systems ade_n1, ade_n3, ade_n7, and ade_n9 but differs somewhat from the HOMO of the imino forms (see Figure 3).

After having established a large resemblance between the π orbitals of the different tautomers of adenine, the σ orbitals are investigated. As these molecular orbitals also directly involve the atomic orbitals on the proton that is transferred to different positions, larger differences are found, in line with a previous study.[23] Not all details of Figure 2 will be examined as for the π orbitals, but some striking observations are discussed. The lowest-lying σ orbital investigated is the 16σ. This orbital has the same shape for all six systems: It is largely σ bonding but with antibonding character, among others, between the purine system and the amino group (see Figure 4). Resemblances can be found for other orbitals also. However, when we analyze orbitals 25σ and 26σ, we see that the 26σ orbitals of ade_n1

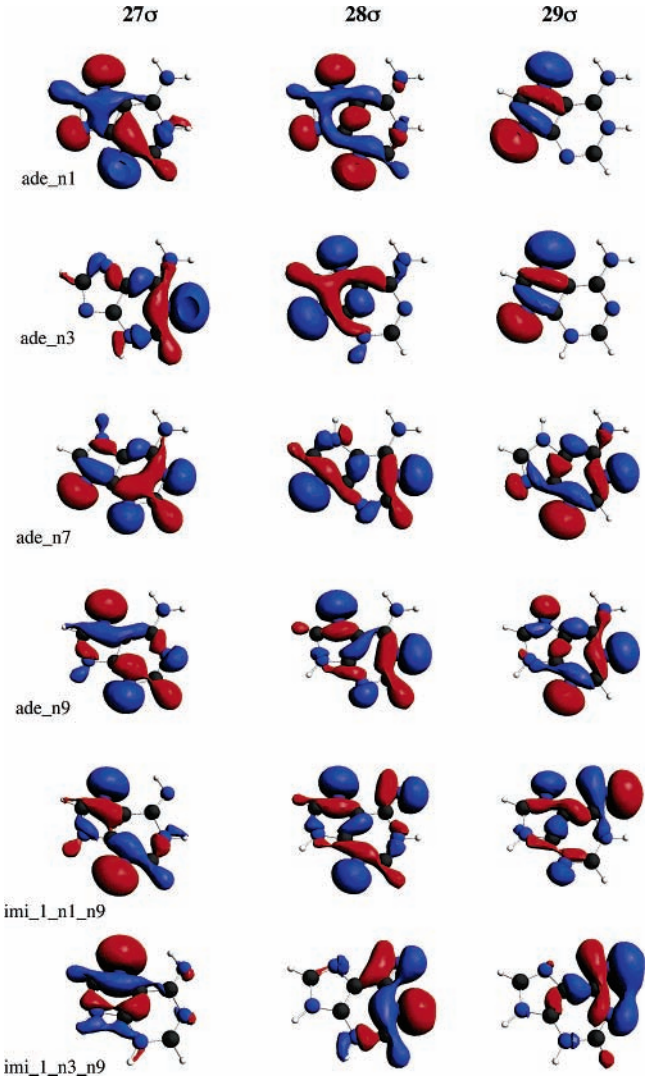


Figure 5. The molecular orbitals 27σ, 28σ, and 29σ of the tautomers ade_n1, ade_n3, ade_n7, ade_n9, imi_1_n1_n9, and imi_1_n3_n9 calculated at the BP86/QZ4P level.

and ade_n3 resemble the 25σ orbitals of ade_n7, ade_n9, imi_1_n1_n9, and imi_1_n3_n9. And the other way around, the 25σ of ade_n1 and ade_n3 resembles the 26σ of ade_n7, ade_n9, imi_1_n1_n9, and imi_1_n3_n9 (see Figure 4). So, the energetic order of the orbitals is changed. Furthermore, the three highest occupied orbitals (27σ, 28σ, and 29σ) are all linear combinations of the lone-pair orbitals on the nitrogen atoms (see also ref 10). The shift of the proton between nitrogen atoms of course strongly affects the lone-pair orbitals involved. As a consequence, they mix, and there is no clear one-on-one mapping. Note that the HOMO of imi_1_n1_n9 has lone-pair character on the front side of the base, just as ade_n9, only on a different nitrogen atom.

3.4. Mismatch with Cytosine. As pointed out in the Introduction, rare tautomeric forms of DNA bases are suspected to play a key role in the formation of spontaneous mutations in DNA.[1−3] For example, if adenine tautomerizes to imi_1_n1_n9 (from hereon designated A*), the latter will pair with cytosine instead of thymine. To obtain more insight into the nature of the resulting tautomer base pair A*C, we have analyzed its hydrogen-bonding mechanism and have compared this with that in the natural Watson−Crick AT pair. Previously, we have shown that hydrogen bonds in natural Watson−Crick base pairs[12] and mimics[5] thereof have substantial covalent character

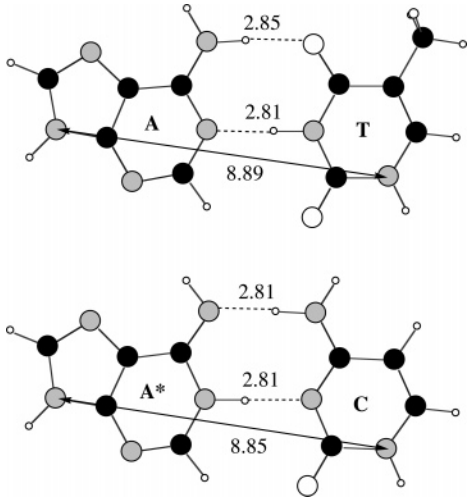


Figure 6. Hydrogen bond lengths and distances between glycosidic nitrogen atoms (in Å) in AT and A*C at BP86/TZ2P.

TABLE 6: Hydrogen Bond Energy Decomposition (in kcal/mol) for the Watson−Crick AT Pair and the Tautomer A*C Pair[a]

	AT	A*C
Orbital Interaction Decomposition		
ΔE_σ	−20.7	−27.1
ΔE_π	1.7	−3.6
ΔE_{oi}	−22.4	−30.7
Bond Energy Decomposition		
ΔE_{Pauli}	39.2	48.9
ΔV_{elstat}	−32.1	−42.5
$\Delta E_{Pauli} + \Delta V_{elstat}$	7.1	6.4
ΔE_{oi}	−22.4	−30.7
ΔE_{int}	−15.3	−24.3
ΔE_{prep}	2.3	6.0
ΔE	−13.0	−18.3

[a] A* = imi_1_n1_n9; see Scheme 1.

that stems from donor−acceptor interactions between nitrogen or oxygen lone pairs and N−H σ* acceptor orbitals. The present analyses show that the base-pairing interaction in the tautomer base pair A*C also gains substantial stabilization from orbital interactions in addition to electrostatic attraction. Figure 6 shows the geometries of AT and A*C and Table 6 summarizes the results of the bond-energy decomposition.

As shown in the experiments by Kool et al.[4] on DNA replication, an incoming nucleotide must be able to form, with its partner in the template, a base pair which sterically resembles a natural AT or GC Watson−Crick pair. In addition, we have previously shown that proceeding from Kool's steric model, also the ability of the incoming base to form hydrogen bonds is of importance.[5] The steric resemblance between the AT pair and the A*C pair is considerable: The distance between the nitrogen atoms attached to the glycosidic bond is for AT 8.89 Å and for imi_1_n1_n9−C 8.85 Å, and the hydrogen bond energy is even more stabilizing for A*C (−18.3 kcal/mol) than for AT (−13.0 kcal/mol). This confirms that A*C is a suitable candidate for the incorporation into DNA. However, one should not forget that it costs 10.9 kcal/mol to go from A to A* (i.e., imi_1_n1_n9) without taking into account the reaction barrier. In this hypothetical case where solvation, stacking interactions, and other environmental influences are left out, it is still 5.6 kcal/mol less favorable to incorporate A*C into DNA than AT. Nevertheless, once an A base has tautomerized to A*, the formation of the mispair A*C, and thus the introduction of a mutation, is geometrically and energetically perfectly feasible.

The electrostatic interaction and the orbital interaction are of the same order of magnitude in both AT and A*C: The electrostatic interactions make up 59% and 58% of all bonding interactions, while the orbital interactions contribute the remaining 41% and 42%, respectively. These values consolidate our previous finding that orbital interactions are as important as electrostatic interactions for providing bonding interaction in hydrogen bonds.[5,9,12]

4. Conclusions

We have calculated the energies and structures of adenine and eleven tautomers thereof. We have established, both at BP86/TZ2P and at BP86/QZ4P, the following order in relative energies: ade_n9 < ade_n3 < ade_n7 < imi_1_n1_n9 < imi_2_n1_n7 < imi_1_n1_n7 < imi_2_n3_n7 < imi_2_n1_n9 < ade_n1 < imi_1_n3_n7 < imi_1_n3_n9/imi_2_n3_n9 (see Scheme 1). The ade_n3 tautomer is very close in energy to (even 0.7 kcal/mol lower than) ade_n7, which has been observed in IR experiments as a naturally occurring tautomer of canonical adenine (ade_n9).

The first ionization energies of the various tautomers calculated as molecular orbital energies using meta-Koopman's theorem are 3 eV (BP86 potential) and 1−1.5 eV (SAOP potential) lower than vertical ionization energies calculated as the difference in energy between the radical cation and the corresponding neutral system (Δ-SCF approach).

The shape of the molecular orbitals for the different tautomers have been compared. The molecular orbitals of the different tautomers, which correspond in character, do not always have the same energetic sequence. Therefore, one should be careful when comparing these orbitals by their energetic sequence, and one should analyze their character. The π orbitals of the different tautomers show close resemblance, as do some of the σ orbitals as well. However, in many cases, the orbitals in the σ systems have different energetic orders. In the case of the nitrogen lone-pair orbitals, they also mix from one species to another, and no simple one-to-one correspondence exists.

Finally, one of the tautomeric forms studied, imi_1_n1_n9 (A*), has been suspected as a possible initiator of mutations in DNA.[1−3] Our analyses confirm that A*C is a suitable candidate for the incorporation into DNA. This base pair has the right steric shape to nicely fit into the DNA double helix. Also, it is more strongly bound than the natural Watson−Crick pair AT, and just as the latter, its hydrogen bonds receive an appreciable stabilizing contribution from orbital interactions which are nearly as large as the electrostatic component. Thus, once an A base has tautomerized to A*, the formation of the mispair A*C and thus the introduction of a mutation into the genetic information is geometrically and energetically perfectly feasible.

Acknowledgment. We thank The Netherlands organization for Scientific Research (NWO-CW), the National Research School Combination − Catalysis (NRSC-C) and the Australian Reaearch Council for financial support.

Supporting Information Available: Cartesian coordinates, orbital energies, and molecular orbitals of tautomers. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) (a) Watson, J. D.; Crick, F. H. C *Nature (London)* **1953**, *171*, 964. (b) Topal, M. D.; Fresco J. R. *Nature (London)* **1976**, *263*, 285. (c) Harris, V. H.; Smith, C. L.; Cummins, W. J.; Hamilton, A. L.; Adams, H.; Dickman, M.; Hornby, D. P.; Williams, D. M. *J. Mol. Biol.* **2003**, *326*, 1389.

(2) (a) Hanus, M.; Kabelác, M.; Rejnek, J.; Ryjácek, F.; Hobza, P. *J. Phys. Chem. B* **2004**, *108*, 2087. (b) Gorb, L.; Podolyan, Y.; Leszczynski, J.; Siebrand, W.; Fernández-Ramos, A.; Smedarchina, Z. *Biopolymers* **2002**, *61*, 77. (c) Rueda, M.; Luque, F. J.; López, J. M.; Orozco, M. *J. Phys. Chem. A* **2001**, *105*, 6575, (d) Blas, J. M.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2004**, *126*, 154.

(3) (a) Danilov, V. I.; Anisimov, V. M.; Kurita, N.; Hovorun, D. *Chem. Phys. Lett.* **2005**, *412*, 285. (b) Kawahara, S.-I.; Uchimaru, T.; Sekine, M. *J. Mol. Struct.* **2000**, *530*, 109. (c) Zhanpeisov, N. U.; Sponer, J.; Leszczynski, J. *J. Phys. Chem. A* **1998**, *102*, 10374.

(4) (a) Guckian, K. M.; Krugh, T. R.; Kool, E. T. *J. Am. Chem. Soc.* **2000**, *122*, 6841. (b) Kool, E. T.; Morales, J. C.; Guckian, K. M. *Angew. Chem.* **2000**, *112*, 1046. (c) Morales, J. C.; Kool, E. T. *J. Am. Chem. Soc.* **2000**, *122*, 1001.

(5) (a) Fonseca Guerra, C.; Bickelhaupt, F. M. *Angew. Chem.* **2002**, *114*, 2194; *Angew. Chem., Int. Ed.* **2002**, *41*, 2092. (b) Fonseca Guerra, C.; Bickelhaupt, F. M. *J. Chem. Phys.* **2003**, *119*. 4262.

(6) (a) Gupta, D.; Roitzsch, M.; Lippert, B. *Chem. Eur. J.* **2005**, *11*, 6643, (b) Gupta, D.; Huelsekopf, M.; Morell Cerdà, M.; Ludwig, R.; Lippert, B. *Inorg. Chem.* **2004**, *43*, 3386. (c) Miguel, P. J. S.; Lax, P.; Willermann, M.; Lippert, B. *Inorg. Chim. Acta* **2004**, *357*, 4552. (d) Müller, J.; Sigel, R. K. O.; Lippert, B. *J. Inorg. Biochem.* **2000**, *79*, 261. (e) Sponer, J.; Sponer, J. E.; Gorb, L.; Leszczynski, J.; Lippert, B. *J. Phys. Chem. A* **1999**, *103*, 11406. (f) Zamora, F.; Kunsman, M.; Sabat, M.; Lippert, B. *Inorg. Chem.* **1997**, *36*, 1583.

(7) (a) Plützer, Chr.; Kleinermanns, K. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4877. (b) Plützer, Chr.; Nir, E.; de Vries, M. S.; Kleinermanns, K. *Phys. Chem. Chem. Phys.* **2001**, *3*, 5469. (c) Marian, C.; Nolting, D.; Weinkauf, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3306. (d) Nir, E.; Plützer, Chr.; Kleinermanns, K.; de Vries, M. *Eur. Phys. J. D* **2002**, *20*, 317. (e) Chin, W.; Dimicoli, I.; Piuzzi, F.; Tardivel, B.; Elhanine, M. *Eur. Phys. J. D* **2002**, *20*, 347. (f) Mons, M.; Dimicoli, I.; Piuzzi, F.; Tardivel, B.; Elhanine, M. *J. Phys. Chem. A* **2002**, *106*, 5088. (g) Piuzzi, F.; Mons, M.; Dimicoli, I.; Tardivel, B.; Zhao, Q. *Chem. Phys.* **2001**, *270*, 205.

(8) (a) Laxer, A.; Major, D. T.; Gottlieb, H. E.; Fisher, B. *J. Org. Chem.* **2001**, *66*, 5463. (b) Secárová, P.; Marek, R.; Malináková, K.; Kolehmainen, E.; Hocková, D.; Hocek, M.; Sklenár, V. *Tetrahedron Lett.* **2004**, *45*, 6259.

(9) (a) Fonseca Guerra C.; Bickelhaupt, F. M. *Angew. Chem.* **1999**, *111*, 3120; *Angew. Chem., Int. Ed.* **1999**, *38*, 2942. (b) Fonseca Guerra, C.; Bickelhaupt, F. M.; Snijders J. G.; Baerends, E. J. *J. Am. Chem. Soc.* **2000**, *122*, 4117.

(10) (a) Chong, D. P.; Gritsenko, O. V.; Baerends, E. J. *J. Chem. Phys.* **2002**, *116*, 1760. (b) Gritsenko, O. V.; Braïda, B.; Baerends, E. J. *J. Chem. Phys.* **2003**, *119*, 1937. (c) Lemierre, V.; Chrostowska, A.; Dargelos, A.; Chermette, H. *J. Phys. Chem. A* **2005**, *109*, 8348.

(11) Bickelhaupt, F. M.; Baerends, E. J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley-VCH: New York, 2000; Vol. 15, pp 1−86.

(12) Fonseca Guerra, C.; Bickelhaupt, F. M.; Snijders, J. G.; Baerends, E. J. *Chem. Eur. J.* **1999**, *5*, 3581.

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

(14) (a) von Niessen, W.; Schirmer, J.; Cederbaum, L. S. *Comput. Phys. Rep.* **1984**, *1*, 57. (b) Cederbaum, L. S.; Domcke, W. *Adv. Chem. Phys.* **1977**, *36*, 205.

(15) (a) TZVP basis set: Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. *Can. J. Chem.* **1992**, *70*, 560. (b) 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.; Bakken, V.; 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.

(16) (a) Morokuma, K. *Chem. Phys.* **1971**, *55*, 1236. (b) Kitaura, K.; Morokuma, K. *Int. J. Quantum. Chem.* **1976**, *10*, 325.

(17) (a) Ziegler, T.; Rauk, A. *Inorg. Chem.* **1979**, *18*, 1755. (b) Ziegler, T.; Rauk, A. *Inorg. Chem.* **1979**, *18*, 1558. (c) Ziegler, T.; Rauk, A. *Theor. Chim. Acta* **1977**, *46*, 1.

(18) (a) Stone, A. J. *The Theory of Intermolecular Forces*; Clarendon Press: Oxford, 1996. (b) Stone, A. J. *Chem. Phys. Lett.* **1993**, *211*, 101.

(19) Poater, J.; Solà, M.; Bickelhaupt, F. M. *Chem. Eur. J.* In press.

(20) Preuss, M.; Schmidt, W. G.; Seino, K.; Furthmüller, J.; Bechstedt, F. *J. Comput. Chem.* **2004**, *25*, 112.

(21) (a) Lin, J.; Yu, C.; Peng, S.; Akiyama, I.; Li, K.; Lee, L. K.; LeBreton, P. R. *J. Am. Chem. Soc.* **1980**, *102*, 4627. (b) Orlov, V. M.; Smirnov, A. N.; Varshavsky, Y. M. *Tetrahedron Lett.* **1976**, *17*, 4377.

(22) (a) Crespo-Hernandez, C. E.; Arce, R.; Ishikawa, Y.; Gorb, L.; Leszczynski, J.; Close D. M. *J. Phys. Chem. A* **2004**, *108*, 6373. (b) Dolgounitcheva, O.; Zakrzewski, V. G.; Ortiz, J. V. *Int. J. Quantum Chem.* **2000**, *80*, 831.

(23) Wang, F.; Downtown, M. T.; Kidwani, N. *J. Theor. Comput. Chem.* **2005**, *4*, 247.