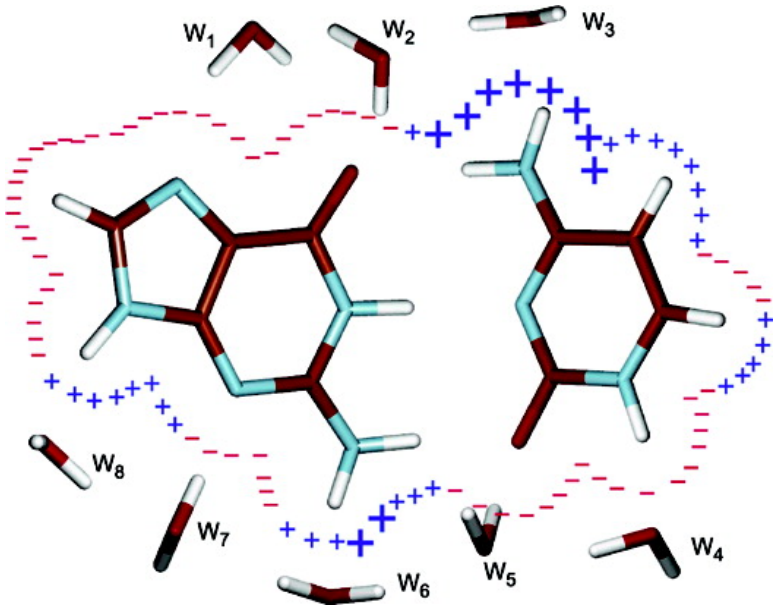
Article

Double-Proton Transfer in Adenine–Thymine and Guanine–Cytosine Base Pairs. A Post-Hartree–Fock ab Initio Study

Leonid Gorb, Yevgeniy Podolyan, Pawel Dziekonski, W. Andrzej Sokalski, and Jerzy Leszczynski

J. Am. Chem. Soc., **2004**, 126 (32), 10119-10129• DOI: 10.1021/ja049155n • Publication Date (Web): 24 July 2004



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 10 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

Journal of the American Chemical Society is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

J|A|C|S
ARTICLES
Published on Web 07/24/2004

Double-Proton Transfer in Adenine−Thymine and Guanine−Cytosine Base Pairs. A Post-Hartree−Fock ab Initio Study

Leonid Gorb,[†] Yevgeniy Podolyan,[†] Pawel Dziekonski,[‡] W. Andrzej Sokalski,[‡] and Jerzy Leszczynski*,[†]

Contribution from the Computational Center for Molecular Structure and Interactions, Department of Chemistry, Jackson State University, 1325 J.R. Lynch St., P.O. Box 17910, Jackson, Mississippi 39217-0510, and Institute of Physical and Theoretical Chemistry, Wroclaw University of Technology, Wyb. Wyspianskiego 27, 50−370 Wroclaw, Poland

Received February 16, 2004; E-mail: jerzy@ccmsi.us

Abstract: The results of a comprehensive study on the double-proton transfer in Adenine−Thymine (AT) and Guanine−Cytosine (GC) base pairs at room temperature in gas phase and with the inclusion of environmental effects are obtained. The double-proton-transfer process has been investigated in the AT and GC base pairs at the B3LYP/6-31G(d) and MP2/6-31G(d) levels of theory. It has been predicted that the hydrogen-bonded bases possess nonplanar geometries due to sp^3 hybridization of nitrogen atoms and because of the soft intermolecular vibrations in the molecular complexes. An analysis of the energetic parameters of the local minima suggests that rare AT base pair conformation is not populated due to the shallowness of this minimum, which completely disappears from the Gibbs free energy surface. The stabilization of canonic or rare forms of the DNA bases by water molecules and metal cations has been predicted by calculating the optimal configuration of charges (using differential product/transition state stabilization approach) followed by calculations of the interactions between the base pair and a water/sodium cation.

1. Introduction

The chemical reactions that involve single- or multiple-proton transfers are of permanent interest to scientists and are the subjects of numerous experimental and theoretical studies.[1] The reason for such interest is the significance of proton transfer in the most vital chemical and biological processes. Therefore, the proton transfer phenomenon occupies a special place in studies of the structures and properties of DNA bases. Indeed, many investigations of proton transfer in DNA bases have been prompted after the formulation of the famous hypothesis by Watson and Crick which suggested that the fidelity of DNA replication is directly coupled with the proton-transfer ability of the DNA bases.[2,3] However, despite various experimental and theoretical investigations that tend to confirm this hypothesis, there are still no decisive evidence that single- and double-proton transfers really play a crucial role in the fidelity of DNA synthesis; moreover, this concept is still widely discussed.[4] Nevertheless, according to Watson and Crick,[5] the fidelity of DNA synthesis or, in other words, the probability of point spontaneous mutations, depends on the possibility of DNA bases to form so-called "rare" tautomeric forms. "Rare" tautomers have chemical structures that differ from the canonic forms, described by Watson and Crick's DNA model, in the position of the hydrogen atom(s). Therefore, "rare" forms are the products of intra- and/or intermolecular proton transfer. Of course, there are also other mechanisms leading to spontaneous mutations (e.g., cytosine deamination),[6] which are beyond the scope of the current study.

One may find in Scheme 1 examples of proton transfers for each DNA base that could be of biological significance. All of

[†] Computational Center for Molecular Structure and Interactions, Department of Chemistry, Jackson State University.

[‡] Institute of Physical and Theoretical Chemistry, Wroclaw University of Technology.

(1) (a) Bell, R. P. *The Proton in Chemistry*; Cornell University Press: Ithaca, 1973. (b) Bountis, T., Ed. *Proton Transfer in Hydrogen-Bonded Systems*; NATO ASI Series 291; Plenum Press: New York, 1992. (c) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin, 1991. (d) Benderskii, V. A.; Makarov, D. E.; Wight, C. A. *Chemical Dynamics at Low Temperature. Advances in Chemical Physics 88*; Wiley & Sons: New York, 1994. (e) Bell, R. P. *The Tunnel Effect in Chemistry*; Chapman and Hall: London, 1080.

(2) (a) Kwiatkowski J. S.; Pullman, B. *Adv. Heterocycl. Chem.* **1975**, *18*, 199. (b) Nowak, M. J.; Szczepaniak, K.; Barski, A.; Shugar, D. *J. Mol. Struct.* **1980**, *62*, 47. (c) Elguero, J.; Marzin, C.; Katrizky, A. R.; Linda, P. *Adv. Heterocycl. Chem.* **1976**, *1*, 655. (d) Blackburn, G. M.; Gait, M. J. *Nucleic Acids in Chemistry and Biology*; Oxford University Press: New York, 1990; p 446.

(3) (a) Kwiatkowski, J. S.; Zelinski, T. J.; Rein, R. *Adv. Quantum Chem.* **1986**, *18*, 85. (b) Person, W. B.; Szczepaniak, K.; Szczepaniak, M.; Kwiatkowski, J. S.; Czerminski, R. J. *J. Mol. Struct.* **1989**, *194*, 239. (d) Leszczynski, J. *Int. J. Quantum Chem.* **1992**, *19*, 43. (e) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Biol. Struct. Dynam.* **1996**, *14* (1), 117. (g) Gould, I. R.; Burton, N. A.; Hall, R. J.; Hillier, I. H. *J. Mol. Struct. (THEOCHEM)* **1995**, *331*, 147. (f) Sponer, J.; Hobza, P. *J. Phys. Chem.* **1994**, *98*, 3161. (h) Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811. (i) Stewart, E. L.; Foley, C. K.; Allinger, N. L.; Bowen, J. P. *J. Am. Chem. Soc.* **1994**, *116*, 7282. (j) Leszczynski, J. *J. Phys. Chem. A* **1998**, *102*, 2357. (k) Anders, B.; Anders, H. *J. Phys. Chem.* **1997**, *101*, 3589.

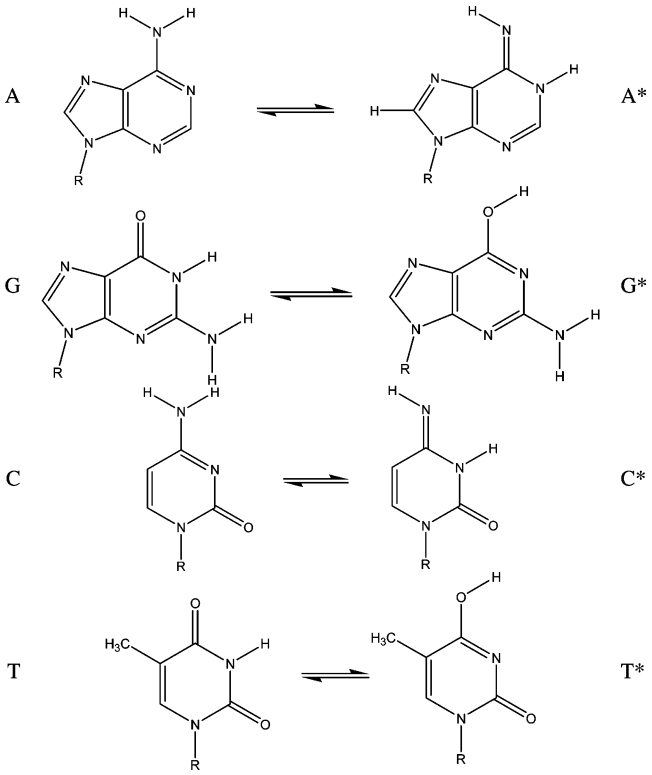
(4) Kool, E. T.; Morales, J. C.; Guckian, K. M. *Angew. Chem., Int. Ed.* **2000**, *39*, 990. Florian, J.; Goodman, M. F.; Warshel, A. *J. Am. Chem. Soc.* **2003**, *125*, 8163. Goodman, M. F. *Proc. Natl. Acad. Sci.* **1997**, *94*, 10 493.

(5) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 964.

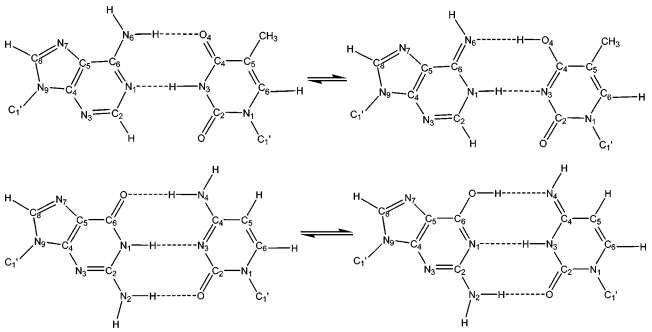
(6) Drake, J. W.; Baltz, R. H. *Annu. Rev. Biochem.* **1976**, *45*, 11.

10.1021/ja049155n CCC: $27.50

Scheme 1



Scheme 2



the processes depicted in Scheme 1 result in a single intramolecular proton transition. The biological importance of such a proton transfer may become apparent during the step of catalytic incorporation of new nucleotides into the growing DNA strand[7] when a "rare" form of a nucleic acid base forms a pair with an incorrect base.

There is another way of how "rare" forms of DNA bases could be obtained. It is represented by the equilibria drawn in Scheme 2. Formally, these two pathways result in the formation of the same tautomers. However, from a chemical point of view, the mechanisms of the tautomer formation are completely different. In the latter case, it results in a multiple-(double-)-proton transfer reaction where DNA bases in the base pair assist each other in the proton transfer. According to the pathways presented in Scheme 2, a biologically significant amount of rare tautomers could be formed before the DNA strand separation.

The following paragraphs briefly describe the current status of experimental and theoretical investigations devoted to the understanding of the mechanism of single- and double-proton-transfer reactions.

First of all, experimental and theoretical data suggest that among tautomeric equilibria represented by Scheme 1, the tautomeric transitions in isolated cytosine and guanine are expected to produce an amount of "rare" tautomers that will significantly (by several orders of magnitude) exceed the concentration of "rare" tautomers occurring in nature as evidenced by the observable frequency of spontaneous mutations. As a result, a considerable amount of "rare" tautomers (from a biological point of view) could be present when the system is in a state of equilibrium.[8] This is contrary to the cases of adenine and thymine where the corresponding equilibria are shifted toward the canonic forms more significantly. Therefore, current efforts are concentrated on the investigation of the tautomeric transitions in the guanine- and cytosine-containing species for which the rare tautomeric forms are observed experimentally.[9]

In particular, recent data suggest that the equilibrium states of isolated cytosine and guanine contain mixtures of tautomers. The experimentally measured equilibrium constants are mostly unavailable with the exception of 9-methylguanine,[8] for which the equilibrium constant of the transformation from the amino-oxo form into the amino-hydroxo form is estimated to be 1.0 $\pm$ 0.3. Much more data are available from computer simulations.[10] For instance, the relative stabilities of the low-energy tautomers are available up to the CCSD(T)/inf//RI-MP2/TZVPP level of theory for cytosine and up to the CCSD(T)/aug-cc-pVDZ//MP2/aug-cc-pVDZ level for guanine. According to these studies, the total energy differences between the guanine and cytosine tautomers presented in Scheme 1 are in the range of 0.5−1.0 kcal/mol.

It is also well-known that under common physiological conditions the various properties of DNA bases can be different from those of the isolated bases due to the influence of a number of factors. First of all, the real building block of the DNA molecule is a nucleotide and not a DNA base. Therefore, we have performed a comprehensive study in which it was shown that the tautomeric properties of nucleotides possessing an anti conformation are virtually the same as those of isolated DNA bases.[11] Since the DNA bases are partially hydrated,[12] a second important factor able to change the tautomeric properties of DNA bases is hydration. For that reason, the modeling of the influence of hydration on the tautomeric transition in DNA bases has been performed in a number of studies.[10d,13] The data available from these studies suggest that the water molecules

(7) (a) Topal, M. D.; Fresco, J. R. *Nature* **1976**, *263*, 285. (b) 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.

(8) Szczepaniak, K.; Szczepaniak, M.; Szajda, W.; Person, W. B.; Leszczynski, J. *Can. J. Chem.* **1991**, *69*, 1705.

(9) (a) Nir, E.; de Vries, M. S. *J. Chem. Phys.* **2001**, *115*, 4604. (b) Mons, M.; Dimicoli, I.; Piuzzi, F.; Tardivel, B.; Elhanine, M. *J. Phys. Chem. A* **2002**, *106*, 5088.

(10) (a) Kobayashi, R. *J. Phys. Chem. A* **1998**, *102*, 10 813. (b) Fogarasi, G. *J. Phys. Chem. A* **2002**, *106*, 1381−1390. (c) Fogarasi, G.; Szalay, P. G. *Chem. Phys. Lett.* **2002**, *356*, 383. (d) Gorb, L.; Leszczynski, J. *J. Am. Chem. Soc.* **1998**, *120*, 5024. (e) Podolyan, Y.; Gorb, L.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104*, 7346. (f) Gorb, L.; Podolyan, Y.; Leszczynski, J. *J. Mol. Struct.* (*Theochem*) **1999**, *487*, 479.

(11) Gorb, L.; Shishkin, O.; Leszczynski, J., submitted.

(12) (a) Schneider, B.; Cohen, D. M.; Berman, H. *Biopolymers* **1992**, *32*, 725. (b) Schneider, B.; Berman, H. M. *Biophys. J.* **1995**, *69*, 2661.

(13) (a) Sivanesan, D.; Subramanian, V.; Nair, B. U. *Int. J. Quantum Chem.* **2001**, *84*, 750. (b) Moroni, F.; Famulari, A.; Raimondi, M. *J. Phys. Chem.* **2001**, *105*, 1169. (c) Aleman, C. *Chem. Phys.* **1999**, *244*, 151. (d) Sivanesan, D.; Sumathi, I.; Welsh, W. J. *Chem. Phys. Lett.* **2003**, *367*, 351. (e) Moroni, F.; Famulari, A.; Raimondi, M. *J. Phys. Chem. A* **2001**, *105*, 1169. (f) Shishkin, O. V.; Gorb, L.; Leszczynski, J. *J. Phys. Chem. B* **2000**, *104*, 5357.

shift the equilibria depicted in Schemes 1 and 2 toward the canonic forms. However, even for hydrated forms the values of the equilibrium constants significantly exceed the values of the frequencies of spontaneous transitions.[14]

Finally, it should be pointed out that the equilibria presented in Schemes 1 and 2 are not necessarily established in the living organisms.[15] The recently predicted values of the barrier for intramolecular proton transfers are very high[10d,10f,14] (c.a., 30 kcal/mol). It is quite possible that some or even all of them will not reach the equilibrium state by the time necessary for the synthesis of new DNA. Therefore, the kinetics of the proton transfer should be also considered together with the thermodynamics. This approach has been introduced recently[14] for both isolated and monohydrated species.

In summary, one may conclude that despite the remaining questions (e.g., the role of the surrounding ions and molecules, estimation of the contribution of each mechanism to the frequency of spontaneous mutations, the role of proton tunneling effect,[16] etc.), the mechanism of single-proton transfer in DNA bases has been investigated relatively accurately. This is contrary to the mechanism of the pathway presented in Scheme 2. The analysis of available published data[17] suggests that a detailed mechanism of double-proton transfer remains unclear even for such relatively simple systems as the formic acid dimer,[17b,c] the formic acid−formamidine dimer,[17f] the formamidine dimer,[17c] the formamide dimer,[17e] the formamidine-formamide dimer,[17a,g,h] etc. An analysis of the current results and problems[17a] in understanding the proton transfer mechanisms taking place in these systems suggests the following: The topology of the potential surface of the double-proton transfer in relatively similar systems may be different and depends on the chemical structure, reaction conditions (vacuum or polar solvent) and the level of ab initio calculations. For example, the formamide−formamidine dimer represents the system that most closely mimics the hydrogen-bonding pattern in the adenine-thymine base pair. For this particular complex, the gas-phase calculations of the mechanism of double-proton transfer performed at the HF level suggest a stepwise mechanism via the formation of a zwitterionic intermediate in the reaction.[17g] However, much more reliable calculations that include electronic correlation (at the B3LYP/aug-cc-pVDZ and CCSD(T)/aug-cc-pVDZ//QCISD/6-31+G(d,p) levels) do not support this point. It has been concluded that the pathway of proton transfer in this system should be classified as concerted (i.e., the pathway without intermediate) and asynchronous (protons move with a time gap).[17a]

The situation becomes even more uncertain when one turns to the analysis of the proton transfer mechanism in the DNA base pairs. There are a only few studies[17h,18] that address the profile of the double-proton-transfer potential surface. However, in these papers the geometries of the local minima and saddle points have been obtained at the Hartree−Fock (HF) level of theory. Taking into account the aforementioned results for the model systems, one must conclude that a new investigation at the electron-correlated ab initio levels is required. Therefore, in the present study we used the second-order Møller−Plesset perturbation and density functional theory methods to locate the local minima and corresponding transition states for the double-proton transfer reactions in adenine−thymine (AT) and guanine−cytosine (GC) base pairs. On the basis of the results obtained from the analysis of the geometrical parameters and harmonic vibrational frequencies of the base pairs, we will address their structural characteristics, the thermodynamic and kinetic factors that govern the double-proton transfer, and finally some biological consequences related to the fidelity of the DNA synthesis.

The molecular environment of DNA complementary bases (e.g., various specific counterion configurations) may induce a considerable influence on the double-proton-transfer reaction and the stability of rare tautomeric form. Since it would be difficult to examine the great variety of possible molecular surroundings with the conventional supermolecular approach, it would be useful to solve the inverse problem that concerns defining the characteristics of an optimal molecular environment exerting optimal catalytic activity toward double-proton transfer or rare form stability. Corresponding catalytic fields could be derived within the differential transition state (product) stabilization approach[19] which has been successfully applied to model systems involving proton-transfer reactions.[20,21]

2. Methods

The ab initio LCAO-MO method was used for the study of proton transfer in AT and GC DNA base pairs. The calculations were carried out using the Gaussian 98 program package.[22] Two-dimensional adiabatic potential energy surfaces of proton transfer have been obtained at the B3LYP/6-31G(d) level of theory. The structures of the corresponding critical points on these surfaces have been reoptimized without symmetry restrictions (C_1 symmetry was assumed) by the gradient procedure at the B3LYP/6-31G(d) and MP2/6-31G(d) levels of theory. To verify that the structures have been properly attributed to the local minima or transition states, the matrixes of the energy second derivatives (Hessians) at the corresponding level of theory were checked to have zero and one imaginary eigenvalue, respectively. Single-point calculations have been performed at the MP2 level with the cc-pVDZ and cc-pVTZ basis sets.

(14) Podolyan, Y.; Gorb, L.; Leszczynski, J. *Int. J. Mol. Sci.* **2003**, *4*, 410. (b) Gorb, L.; Podolyan, Y.; Leszczynski, J.; Siebrand, W.; Fernandez-Ramos, A.; Smedarchina, Z. *Biopolymers (Nuc. Acid Sci.)* **2002**, *61*, 77.
(15) Löwdin, P. O. Quantum Genetics and Aperiodic Solids. Some Aspects of the Biological Problems of Heredity, Mutations, Aging, and Tumors in View of Quantum Theory of the DNA Molecule. Preprint No 32. Quantum Theory Project for Research in Atomic, Molecular and Solid State, Chemistry and Physics. University of Florida, Gainesville, Florida, 1962.
(16) Miura, S.; Tuckerman, M. E.; Klein, M. L. *J. Chem. Phys.* **1998**, *109*, 5290.
(17) (a) Podolyan, Y.; Gorb, L.; Leszczynski, J. *J. Phys. Chem. A* **2002**, *106*, 12 103. (b) Kim, Y. *J. Am. Chem. Soc.* **1996**, *118*, 1522. (c) Lim, J.-H.; Lee, E. K.; Kim, Y. *J. Phys. Chem. A* **1997**, *101*, 2233. (d) Loerting, T.; Liedl, K. R. *J. Am. Chem. Soc.* **1998**, *120*, 12 595. (e) Kim, Y.; Lim, S.; Kim, H.-J.; Kim, Y. *J. Phys. Chem. A* **1999**, *103*, 617. (f) Kim, Y.; Lim, S.; Kim, Y. *J. Phys. Chem. A* **1999**, *103*, 6632. (g) Zhanpeisov, N. U.; Leszczynski, J. *J. Phys. Chem. B* **1999**, *103*, 8317. (h) Bertran, J.; Oliva, A.; Rodriguez-Santiago, L.; Sodupe, M. *J. Am. Chem. Soc.* **1998**, *120*, 8159.
(18) (a) Florian, J.; Hrouda, V.; Hobza, P. *J. Am. Chem. Soc*. **1994**, *116*, 1457. (b) Florian, J.; Leszczynski, J. *J. Am. Chem. Soc.* **1996**, *118*, 3011. (c) Kryachko, E. S.; Sabin, J. R. *Int. J. Quantum Chem.* **2002**, *91*, 695.
(19) Sokalski, W. A. *J. Mol. Catal*. **1985**, *30*, 395.
(20) Dziekonski, P.; Sokalski, W. A.; Podolyan, Y.; Leszczynski, J. *Chem. Phys. Lett.* **2003**, *367*, 367.
(21) Dziekonski, P.; Sokalski, W. A.; Leszczynski, J. *Chem. Phys.* **2001**, *272*, 37.
(22) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.7; Gaussian: Pittsburgh, PA, 1998.

To correct the computed energies for the deficiencies due to the incompleteness of the basis sets, basis set extrapolation has been performed. Since the cc-pVTZ basis set often approaches the edge of computational feasibility for moderately sized systems, we used the extrapolation procedure developed by Truhlar.[23] In accordance with this technique, the HF and correlation energies are extrapolated separately. The total energy is calculated by combining the equation

$$E_{\infty}^{\mathrm{tot}} = E_{\infty}^{\mathrm{HF}} + E_{\infty}^{\mathrm{cor}} \tag{1}$$

with the following two equations for the HF and correlation energies

$$E_X^{\mathrm{HF}} = E_{\infty}^{\mathrm{HF}} + A^{\mathrm{HF}} X^{-\alpha} \tag{2}$$

$$E_X^{\mathrm{cor}} = E_{\infty}^{\mathrm{cor}} + A^{\mathrm{cor}} X^{-\beta} \tag{3}$$

where $X = 2$ for cc-pVDZ and $X = 3$ for cc-pVTZ, and A, α and β are fitting parameters.

Provided that the exponents are constants of the electronic structure method (i.e., different for the HF and MP2 methods but independent of the molecular system), the resulting expression for the infinite basis set limit energy estimate is given by

$$E_{\infty}^{\mathrm{tot}} = \frac{3^{\alpha}}{3^{\alpha} - 2^{\alpha}} E_3^{\mathrm{HF}} - \frac{2^{\alpha}}{3^{\alpha} - 2^{\alpha}} E_2^{\mathrm{HF}} + \frac{3^{\beta}}{3^{\beta} - 2^{\beta}} E_3^{\mathrm{cor}} - \frac{2^{\beta}}{3^{\beta} - 2^{\beta}} E_2^{\mathrm{cor}} \tag{4}$$

The values of the Gibbs free energy of formation have been calculated using the standard expression $\Delta G = \Delta H - T\Delta S$ at room temperature (298.15 K) where ΔH includes the thermal correction to enthalpy (H_T) calculated at the same level of theory as the geometry optimization procedure. The ΔS values have been calculated at the same levels of theory as the optimizations using the rigid rotor—harmonic oscillator—ideal gas approximation. The dimerization energy of the monomers (ΔE_{dim}) has been obtained as a sum of the calculated energy of the interaction (E_{int}) corrected for the basis set superposition error[24] and the deformation energy (E_{def}): $\Delta E_{\mathrm{dim}} = E_{\mathrm{int}} + E_{\mathrm{def}}$. The Gibbs free energy of dimerization was calculated by the subtraction of the $T\Delta S$ term from the dimerization enthalpy ΔH_{dim}. The latter was calculated by adding the thermal correction to the enthalpy term to the value of ΔE_{dim}.

To estimate the rate constants of proton-transfer reactions, the approximate instanton approach[25] has been employed as implemented in the DOIT 1.2 program.[26] The tunneling rate constant for the G*C* → GC proton transfer has been calculated using the expression

$$k_{\mathrm{r}}(T) = (\Omega_0^i / 2\pi) e^{-S_{\mathrm{I}}(T)} \tag{5}$$

where Ω_0^i is the effective tunneling frequency in the equilibrium configuration of G*C*, and $S_{\mathrm{I}}(T)$ is the multidimensional instanton action. The rate constants for the proton-transfer processes presented in Scheme 2 were calculated from $k_{\mathrm{f}}(T) = K_{\mathrm{eq}}(T)\, k_{\mathrm{r}}(T)$ where $K_{\mathrm{eq}}(T)$ is the equilibrium constant calculated using following relationship

$$K_{\mathrm{eq}}(T) = e^{-\Delta G^0/RT} \tag{6}$$

The population of vibrational levels has been calculated for 298.15 K using Boltzmann's distribution formula.

To analyze the influence of the DNA molecular environment (water molecules and metal cations) on the double-proton-transfer reaction, we applied the differential product/transition state stabilization (DPS/DTSS) approach.[19] In this method, the activation barrier lowering Δ or product relative stabilization energy is estimated from the difference in the interaction energies of any part of molecular environment C (water or cation) with the corresponding local minima (substrate S or product P) or substrate (S) and transition state (TS)[21]

$$\Delta = \Delta E(\mathrm{X,C}) - \Delta E(\mathrm{S,C}) \tag{7}$$

where X = TS or P. Negative Δ values indicate activity promoting the proton-transfer process, whereas positive values represent inhibitory activity of a particular part C of the molecular environment. Analysis of the physical nature of the catalytic activity of the hydration shell on proton-transfer reactions in model systems indicates the dominant role of the electrostatic term.[20,21] This allows the static or dynamic characteristics of molecular environment with optimal activity to be derived just from a knowledge of the superimposed transition state and substrate structures. In this case, the molecular environment contains a unitary positive point charge; Δ may be estimated as the difference of the corresponding molecular electrostatic potentials

$$\Delta_{\mathrm{s}} = V(\mathrm{X}) - V(\mathrm{S}) \tag{8}$$

where X = TS or P. The magnitude of the activation barrier lowering or relative product stabilization Δ by the unitary positive or negative charge located on the molecular van der Waals surface is indicated by the size of the plus or minus sign (catalytic fields).

Different structural parameters for the DNA base pairs such as opening, propeller twist, and buckle were calculated for the optimized geometries using the 3DNA software package.[27] Graphic representations of the molecular geometries have been made using the Molden program.[28] The potential energy surfaces have been constructed using the Surfer 5.01 program.

3. Results and Discussion

3.1. Topology of the Proton-Transfer Potential Surfaces. The topologies of the two-dimensional adiabatic potential energy surfaces for the proton transfer in adenine-thymine and guanine-cytosine DNA base pairs are shown in Figure 1. The surfaces have been constructed by varying the interatomic distances (O6(G)−H4(C), N3(C)−H1(G), and H2(G)−O2(C) in GC and H6(A)−O4(T) and H3(T)−N1(A) in AT) and by optimizing all other geometrical parameters. One may see that the potential energy surface describing the motion of the protons in the N1(G)−N3(C)−N4(C)−O6(G) region of the GC pair has three very definite critical points that correspond to the canonic structure of GC, the double-proton-transfer structure of GC (which has two hydrogen-bonded "rare" bases (G*C*)), and the transition state ($\neq$) between the two (GC → G*C*)‡. As follows from the topology of the proton transfer energy surface presented in Figure 1b, there is only one critical point which corresponds to the structure of the canonic GC base pair for the double-proton transfer between the N2(G)−O2(C) and N4(C)−O6(G) atoms. Thus, both of the considered quantum-chemical approximations suggest only the double-proton transfer in the GC base pair that results in the formation of the most stable (when isolated) rare tautomeric forms of guanine and cytosine.[10a,b,d]

(23) Truhlar, D. G. *Chem. Phys. Lett.* **1998**, *294*, 45.
(24) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(25) (a) Benderskii, V. A.; Makarov, D. E.; Wight, C. H. *Adv. Chem. Phys.* **1994**, *88*, 1. (b) Smedarchina, Z.; Fernandez-Ramos, A.; Siebrand, W. *J. Comput. Chem.* **2001**, *22*, 787. (c) Siebrand, W.; Smedarchina, Z.; Zgierski, M. Z.; Fernandez-Ramos, A. *Int. Rev. Phys. Chem.* **1999**.
(26) Smedarchina, Z.; Fernandez-Ramos, A.; Zgierski, M. Z.; Siebrand, W. DOIT 1.2, A computer program to calculate hydrogen tunneling rate constants and splittings, National Research Council of Canada: 2000.
(27) http://www.3dna.net/
(28) Schaftenaar, G.; Noordik, J. H. *J. Comput.-Aided Mol. Design* **2000**, *14*, 123.

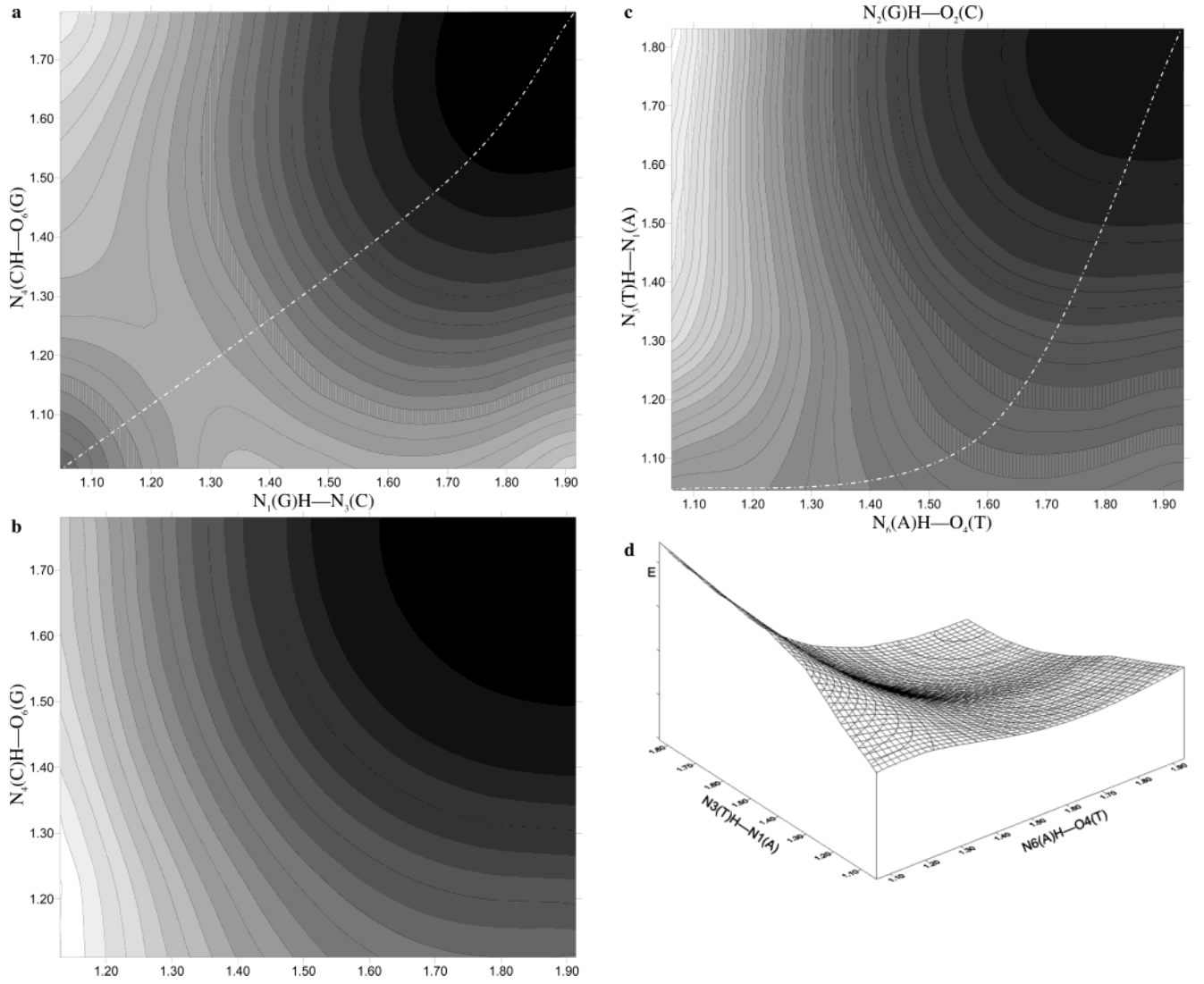


Figure 1. Potential energy surfaces of the double-proton transfer in GC (a and b) and AT (c and d) base pairs with indicated pathways (darker contours on the two-dimensional surfaces represent lower energies).

The situation is more complicated in the case of the AT base pair. The potential energy surface created with the interval between the contour levels used for GC has shown only one minimum which corresponds to the structure of the canonic AT base pair (see Figure 1c). However, when the interval is decreased one may distinguish a very shallow minimum (which is also visible on the three-dimensional potential energy surface in Figure 1d) representing the A*T* structure and the corresponding transition state located very close to it.

As was already mentioned, the structures that correspond to local minima and transition states on the potential energy surfaces presented in Figure 1 have been reoptimized at the B3LYP/6-31G(d) and MP2/6-31G(d) levels of theory without symmetry restrictions and were verified to be true local minima or transition states. The basic intermolecular distances for the optimized local minima and transition states are presented in Figure 2. Since the obtained geometrical parameters exhibit the typical and well-known tendencies peculiar to intermolecular complexes (see, for example, their analysis in ref 17a), we will discuss them only briefly.

The local minima are represented by hydrogen-bonded complexes characterized by interatomic and intermolecular distances that are usual for hydrogen-bonded systems (see Figure 2). It is important to note, however, that the intermolecular distances in both the A*T* and the G*C* hydrogen-bonded base pairs are significantly shorter than those in the canonic base pairs (except for the N2(G)H···O(C) hydrogen bond in GC base pair, which is longer). This indicates that the hydrogen bonding becomes stronger (as seen from the values of dimerization energy in Tables 3 and 4) after the double-proton transfer occurs in the AT and GC base pairs.

We believe that the levels of theory presented in this study are sufficient to describe correctly the number of saddle points on the potential energy surface. In particular, the obtained results suggest that the reaction paths that include a single-proton transfer, which results in the formation of zwitterionic stable intermediates,[18b] should not be considered at all for isolated (i.e., nonhydrated) species. This is, actually, quite natural since a single-proton transfer between the DNA bases is equivalent, in fact, to the protonation of one of the bases by another one. However, it is very well-known that the only stable gas-phase zwitterionic intermediate of such type is the $(CH_3)_3NH^+\cdots I^-$

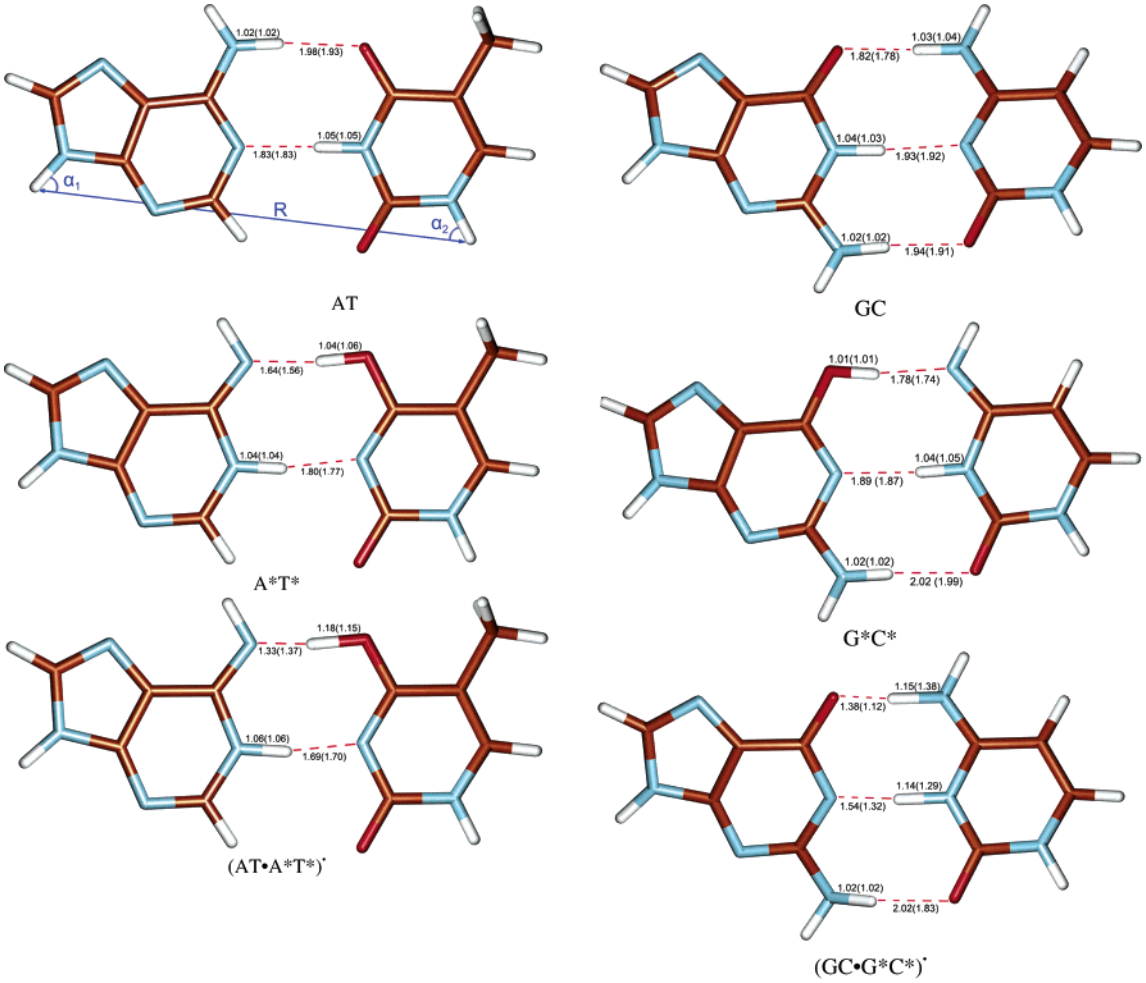


Figure 2. Basic geometrical parameters of the base pairs, their rare forms and transition states optimized at MP2 (B3LYP in parentheses) level of theory.

Table 1. Structural Parameters of the Base Pairs (R, in Å; α_1, α_2, Opening, Propeller Twist, and Buckle Angles in Degrees)[a]

structure	method	*R*	α_1	α_2	opening	propeller twist	buckle
AT	B3LYP/6-31G(d)	10.10	54.5	55.5	−0.99	0.03	0.01
	MP2/6-31G(d)	10.05	54.9	56.1	−0.02	3.31	−5.11
(AT → A*T*)‡	B3LYP/6-31G(d)	10.07	51.1	53.6	−7.39	−0.02	0.08
	MP2/6-31G(d)	10.03	50.8	54.0	−7.28	0.00	0.00
A*T*	B3LYP/6-31G(d)	10.10	51.4	54.4	−6.35	−0.01	0.06
	MP2/6-31G(d)	10.06	51.5	55.5	−5.18	−0.01	0.04
GC	B3LYP/6-31G(d)	10.21	53.4	55.3	−2.34	1.76	−3.05
	MP2/6-31G(d)	10.20	53.4	55.7	−2.01	6.51	6.31
(GC → G*C*)‡	B3LYP/6-31G(d)	10.01	52.0	53.1	−4.85	−13.50	−2.94
	MP2/6-31G(d)	10.09	51.1	51.8	−5.77	−21.87	−1.24
G*C*	B3LYP/6-31G(d)	10.27	52.3	53.0	−4.22	5.47	5.19
	MP2/6-31G(d)	10.26	52.4	53.4	−3.67	10.03	−8.02

[a] Definitions of *R*, α_1, α_2 are presented in Figure 2. For the definitions of opening, propeller twist, and buckle angles see ref 40.

Table 2. Low-lying Intermolecular Vibrational Frequencies of the Base Pairs[a]

method	γHB_s	τ_s	γHB_a	δHB_s	νHB_a	νHB_s
			AT			
B3LYP/6-31G(d)	20	32	66	102	62	110
MP2/6-31G(d)	21	33	66	96	63	111
			(A → T)‡			
B3LYP/6-31G(d)	28	42	85	112	85	210
MP2/6-31G(d)	28	41	81	106	89	195
			A*T*			
B3LYP/6-31G(d)	26	41	83	108	78	119
MP2/6-31G(d)	26	41	78	101	78	126
			GC			
B3LYP/6-31G(d)	21	34	67	96	131	138
MP2/6-31G(d)	26	36	68	89	128	135
			(G → C)‡			
B3LYP/6-31G(d)	33	37	83	118	154	244
MP2/6-31G(d)	31	51	78	117	−	166/217
			G*C*			
B3LYP/6-31G(d)	28	36	80	120	113	140
MP2/6-31G(d)	29	37	75	110	117	132/142

[a] In the vibrational mode symbols the letters ν, δ, γ, and τ denote stretching, in-plane bending, out-of-plane bending, and torsion vibrations, respectively. Subscripts *s* and *a* denote symmetric and antisymmetric, respectively.

ion pair.[29] This species is stabilized by the proton acceptor ability of $(CH_3)_3N$ and by the proton donor ability of HI which are significantly larger than those of DNA bases.

Interesting information could be obtained from the analysis of the geometrical parameters of the transition states. Recently, we have performed a comprehensive study of the double-proton transfer in a prototypic system[17a] that mimics very closely the proton transfer in the AT base pair. Analyzing the geometry of the transition state we have concluded that the proton transfer occurs concertedly and asynchronously (successively). An analysis of the data presented in Figures 1 and 2 allows one to draw the same conclusion as to the mechanism of the double-

(29) Legon, A. C.; Rego, C. A. *J. Chem. Phys.* **1993**, *99*, 1463.

Table 3. Values of the Relative Energies (not corrected for zero-point vibrational energy) of the Base Pair Dimerization,[a] Transition State, Rare Tautomeric Forms, and Dimerization of the Rare Tautomeric Forms (in kcal/mol)[b]

	ΔE_{dim}		$\Delta E^{\ddagger}$		$\Delta E_{C\rightarrow R}$		ΔE^{r}_{dim}	
method	AT	GC	AT	GC	AT	GC	AT	GC
B3LYP/6-31G(d)	−12.1	−25.2	15.4	17.3	15.2	10.8	−23.4	−18.5
MP2/6-31G(d)	−12.0	−23.9	17.3	17.8	16.3	9.7	−22.3	−17.4
MP2/cc-pVDZ// MP2/6-31G(d)	−11.4	−22.0	13.6	14.0	13.1	6.7	−21.8	−16.3
MP2/cc-pVTZ// MP2/6-31G(d)	−13.5	−25.5	12.4	13.8	12.4	7.6	−24.3	−18.9
MP2/infinite// MP2/6-31G(d)	−15.1	−27.9	11.6	13.3	12.0	7.8	−26.1	−20.7

[a] Dimerization energies have been corrected for basis set superposition error. [b] For definition of the values presented in the table see Figure 3

Table 4. Relative Gibbs Free Energies of the Base Pair Dimerization, Transition State, Rare Tautomeric Forms, and Dimerization[a]of the Rare Tautomeric Forms (in kcal/mol)[b]

	ΔG_{dim}		$\Delta G^{\ddagger}$		$\Delta G_{C\rightarrow R}$		ΔG^{r}_{dim}	
method	AT	GC	AT	GC	AT	GC	AT	GC
B3LYP/6-31G(d)	0.3	−11.8	13.8	13.4	15.2	11.1	−11.3	−5.4
MP2/6-31G(d)	0.1	−10.5	15.1	14.6	16.4	9.7	−9.8	−3.9
MP2/cc-pVDZ// MP2/6-31G(d)[c]	0.7	−8.6	11.4	10.8	13.3	6.7	−9.3	−2.8
MP2/cc-pVTZ// MP2/6-31G(d)[c]	−1.4	−12.1	10.2	10.6	12.5	7.6	−11.8	−5.4
MP2/infinite// MP2/6-31G(d)[b]	−3.0	−14.5	9.4	10.2	12.1	7.8	−13.6	−7.2

[a] Dimerization energies have been corrected for basis set superposition error. [b] For definition of the values presented in the table see Figure 4. [c] Thermal Correction to Enthalpy and the Entropy values taken from MP2/6-31G(d) calculations.

proton transfer in the AT base pair. Indeed, the process is concerted and asynchronous. An analysis of the mechanism of the double-proton transfer in the GC base pair is more complicated since the MP2 and DFT theories predict it differently. The geometries of the transition state optimized at these levels indicate (and the results of the calculated reaction path, as implemented in Gaussian with keyword IRC, confirm) that the proton transfer is concerted and asynchronous according to MP2 approximations and is concerted and rather synchronous according to DFT approximations. To further clarify, we have calculated the relative energies for the transition states corresponding to synchronous and asynchronous pathways at the MP2//B3LYP and B3LYP//MP2 levels. In other words, we have performed the calculations of the MP2 energy for the transition state optimized at the B3LYP level and vice versa. The resulting energies at the MP2 and B3LYP levels were 1.4 and 0.8 kcal/mol, respectively, higher than those for the structures optimized at the same level. Therefore, we can conclude that such a relatively small energy difference does not in fact allow us to determine decisively whether the pathway is synchronous or asynchronous. In this regard, we would like to point out that such terms as synchronous and asynchronous paths of proton transfer can be applied accurately only to the proton-transfer systems that do not undergo quantum (tunneling effect) and thermal fluctuations. In reality, both of these factors have a significant influence on the proton-transfer path resulting in the deviation of the proton-transfer trajectories from the minimum energy path.[16] Therefore, similarly to the prototypic molecules,[17] the issue regarding the geometry of the saddle points on the potential energy surface of the double-proton transfer in DNA bases remains open.

The next issue that we would like to address is the deviation of the geometry of the AT and GC base pairs from planarity. As is well-known, the DNA base pairs are planar only on average even if they are incorporated in a DNA sugar−phosphate backbone.[30] Therefore, the geometries of DNA pairs, which form DNA strands, are nonplanar, propeller-twisted, and buckled.[30,31] The observed nonplanarity of the pairs is usually attributed to external factors (e.g., base stacking). However, the results of recent ab initio Hartree−Fock calculations published in a series of papers by J. Sponer and P. Hobza[31a,b,c] suggest that in a number of mutual orientations of DNA bases, the latter are intrinsically nonplanar, mainly due to sp^3 hybridization of the nitrogen atoms in amino groups. Nevertheless, the canonic AT and GC base pairs were considered planar in those studies, probably due to the fact that the geometry optimizations have been performed at the HF level of theory. Let us analyze from this point of view the parameters, presented in Table 1, which describe the size and mutual orientation of DNA base pairs. As can be seen from the obtained results, both the local minima and the transition state of the GC pair are buckled and propeller-twisted. The value of the opening angle is also of significant magnitude. These deviations from nonplanarity are stronger in the G*C* pair and the transition state than in GC.

The only nonplanar equilibrium geometry of the AT base pair has been obtained at the MP2/6-31G(d) level. The geometries of the AT base pair from B3LYP/6-31G(d) simulations and also the A*T* base pair and the transition state for the proton transfer are predicted to be close to planar at both the MP2/6-31G(d) and B3LYP/6-31G(d) levels of theory.

Taking into account the aforementioned results, the following statement should be made. Even though the strength of the hydrogen-bonded interactions between the DNA bases is significant, the structures of the complexes are certainly not expected to be rigid and fixed to their optimized geometries, especially when one considers the properties of these species at temperatures considerably higher than 0 K. The reason is the presence of several very soft intermolecular vibrational modes that are characterized by so-called low-lying frequencies.[32] The values of these frequencies are presented in Table 2. The presence of these modes indicates that, for example, at room temperature only some fraction of the bases will possess a geometry that is close to equilibrium at any moment of time. The rest of the bases will populate the excited vibrational levels and will most likely be even more nonplanar than is described by the equilibrium geometries. To illustrate the idea, we have performed an estimation of the amount of buckled AT complexes that are initiated by the lowest vibration of 20 cm^{-1} (see Table 2) at room temperature (298 K). We found that at this temperature only 32% of AT base pairs would assume an initial planar conformation; the other 68% will populate excited vibrational states, which are definitely nonplanar. The largest

(30) Saenger, W. *Principles of Nucleic Acid Stucture*; Springer-Verlag: New York, Berlin, Heidelberg, London, Paris, Tokyo, 1988.
(31) (a) Sponer, J.; Florian, J.; Hobza, P.; Leszczynski, J. *J. Biomol. Struct., & Dynam.* **1996**, *13*, 827. (b) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 1965. (c) Sponer, J.; Hobza, P. *Int. J. Quantum Chem.* **1996**, *57*, 959. (d) Hovorun, D. M.; Gorb, L.; Leszczynski, J. *Int. J. Quantum Chem.* **1999**, *75*, 245. (e) Shishkin, O. V.; Gorb, L.; Hobza, P.; Leszczynski, J. *Int. J. Quantum Chem.* **2000**, *80*, 1116.
(32) Florian, J.; Leszczynski, J.; Jonson, B. G. *J. Mol. Struct.* **1995**, *349*, 421.

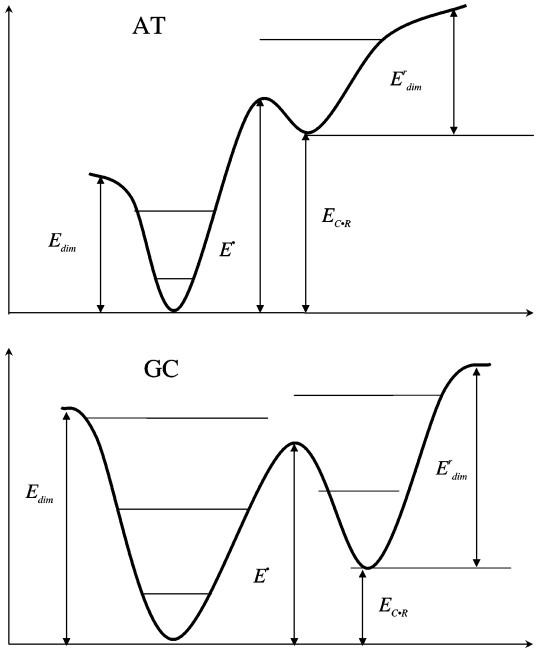


Figure 3. Qualitative representation of potential energy profiles of the double-proton transfer in AT and GC with indicated vibrational levels for the fragments involved in hydrogen bonding. (Relative energies are given in Tables 3 and 4.)

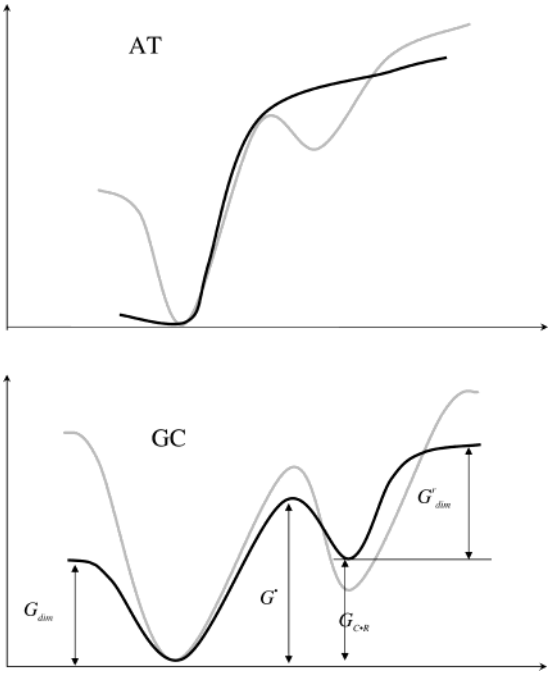


Figure 4. Qualitative representation of Gibbs free energy (black) vs potential energy (grey) profile of the double-proton transfer in AT and GC base pairs. (Relative energies are given in Tables 3 and 4.)

frequency (68 cm^{-1}) of the intermolecular vibrations that distorts the planarity of considered DNA base pairs corresponds to the twisting of the GC pair. Thus, we have also estimated the amount of complexes that possess a nonplanar geometry due to this vibration. We have found that about 27% of the population of vibrationally exited states would adopt this nonplanar geometry (for details on the use of this technique see ref 31e). Since the properties of the DNA bases are of utmost importance at room temperature, it can be concluded that both of the DNA base pairs will undergo various kinds of motions that, in general, will result in nonplanar geometries. In this regard, we would like to mention the results presented in ref 33 where, based on the results of classical molecular dynamic simulations, the internal nonplanarity of hydrated Watson−Crick AT base pairs has been predicted.

3.2. Thermodynamics and Kinetics of Double-Proton Transfer. To discuss the thermodynamic and kinetic aspects of the double-proton transfer, we decided to use two sets of data. The first one is the relative values of the total energies that do not include zero-point vibrational energy corrections. Those values are presented in Table 3. Using those values one can easily introduce a simple one-dimensional model of proton transfer that takes into account only the vibrational levels that contribute most significantly to the proton-transfer path. The definition of each parameter is clear from Figure 3. The second set of data consists of relative Gibbs free energies (Table 4); Figure 4 maps the corresponding definitions to the proton-transfer reaction profile.

Let us first analyze the data presented in Table 3. First of all, it should be mentioned that of no surprise is the fact that the data calculated at the B3LYP/6-31G(d) level are very similar to those calculated at the MP2/6-31G(d) level since these methods usually produce similar geometries and relative energies. According to these results, a one-dimensional diagram of the total energy change along the coordinate of the double-proton transfer could be represented by the profile drawn in Figure 3. We are, of course, aware of the fact that a quantitative description of proton-transfer reactions requires the consideration of at least a four-dimensional potential surface.[34] However, some simple qualitative predictions can be made even by analyzing a one-dimensional profile. Since the most important question concerning the thermodynamics of this reaction is the population of the minima on the potential energy surface, we have also presented schematically the levels of stretching vibrations for the fragments involved in hydrogen bonding, which, according to the MP2/6-31G(d) and B3LYP/6-31G(d) approximations, are located in the region beyond 3000 cm^{-1}. One may see that one or more vibrational levels are placed inside the local minima associated with the AT, GC, and G*C* structures of the DNA base pairs. This suggests that these minima should be populated. However, in the case of the A*T* structure there are no vibrational levels inside the local minimum.

To extend our analysis, we have calculated the values of the relative Gibbs free energies at 298 K where a full zero-point vibrational energy was included along with the entropy and the thermal correction to enthalpy. The most remarkable result is obtained for the AT → A*T* proton transfer. As follows from the data presented in Table 4, the values $\Delta G^{\ddagger}$ are smaller than the values $\Delta G_{C\to R}$ at all considered levels of theory. This results in the disappearance of the local minimum corresponding to the A*T* base pair on the surface of the Gibbs free energy. In addition, the A*T* structure has another interesting peculiarity. As follows from the data presented in Tables 3 and 4 and as predicted in section 3.1, the A* and T* components, which are high-energy forms, are bound to each other approximately twice

(33) Stofer, E.; Chipot, C.; Lavery, R. *J. Am. Chem. Soc*. **1999**, *121*, 9503.

(34) Spirko, V.; Cejchan, A.; Lutchyn, R.; Leszczynski, J. *Chem. Phys. Lett.* **2002**, *355*, 319. Fernandes-Ramos, A.; Smedarchina, Z.; Rodriguez-Otero, J. *J. Chem. Phys*. **2001**, *114*, 1567.

as strong as the A and T components. A similar result was previously obtained in ref 18a.

Let us now draw attention to the values of the Gibbs free energy related to the interaction of the canonic adenine and thymine. Recent computational MD/QM analysis[35] reveals that neither the Watson−Crick nor the Hoogsteen structures of the AT base pair correspond to the global minimum on the Gibbs free energy surface. Therefore, in general, relatively small values of ΔG_{dim} obtained in this study are not surprising. However, only negative values of ΔG_{dim} would suggest that the canonic AT structure is a local minimum on the Gibbs free energy surface. One may see that those negative values have been reached at the MP2 level using only the extended cc-pVTZ and infinite basis sets. Precisely these data should be considered the most accurate since, according to the benchmark calculations of the interaction energy in small intermolecular complexes, the MP2 results are reasonably close to the CCSD(T) values obtained with the same basis sets.[36]

In contrast to the AT base pair, the potential energy surface and the Gibbs free energy surface for the double-proton transfer in the GC base pair has two clear minima. Therefore, the equilibrium between the canonic GC and G*C* forms of the guanine-cytosine base pair should be established. The MP2/infinite//MP2/6-31G(d) level of the calculations predicts the value of the equilibrium constant at 298 K to be approximately 2.0×10^{-6}.

Finally, we have estimated the values of the rate constants that describe the kinetics of the proton transfer in forward and reverse directions. At the MP2/infinite//MP2/6-31G(d) level they amount to 1.2×10^{7} sec^{-1} and 1.2×10^{13} sec^{-1}, respectively. It means that the equilibrium concentration of the G*C* form of the base pair will be achieved instantly as soon as guanine and cytosine form a canonic structure of the base pair.

3.3. Stabilizing and Destabilizing Factors for the Double-Proton Transfer in DNA Base Pairs. Several factors should be taken into account if one wishes to address the mechanism of the double-proton exchange in a DNA strand. Among them are environmental effects resulting from the stacking interaction, nucleotide backbone structure, hydration, and the influence of cations. The influence of the last two factors (hydration and metal cations) can be predicted using the differential product/transition state approach.[19] To perform such an analysis we have presented in Figures 5 and 6 the optimal configuration of charges that should stabilize the formation of rare forms through the double-proton-transfer processes AT → A*T* and GC → G*C*. In other words color-coded fields presented in Figures 5 and 6 and defined by eqs 7 and 8 reflect directly the kind of charge placed near van der Waals contour that will induce AT → A*T* or GC → G*C* rearrangement.

Let us discuss the influence of hydration first. The most convenient way to interpret the data visualized in Figures 5 and 6 is to compare them with the results of the polyhydration influence on the relative energies of AT, A*T*, GC, G*C*. A similar investigation has been performed in our previous study at the MP2/6-31G(d)//HF/6-31G(d) level for the hydration of the AU, A*U*, CiC, and C*iC* base pairs (iC denotes isocytosine, which has the same structural pattern as the six-membered heterocyclic part of guanine).[37] The results of the relative energies obtained previously[35] can be very easily

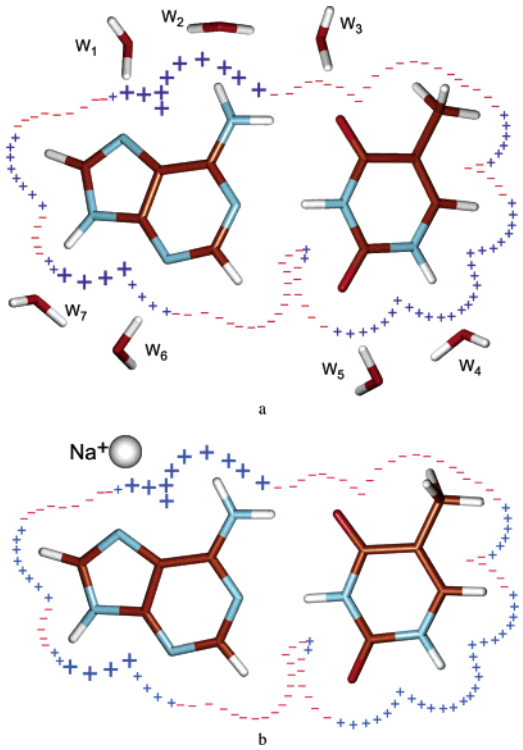


Figure 5. Optimal configuration of charges that should stabilize the A*T* base pair (blue area denotes positive charges; red area denotes negative charges) mapped on (a) AT base pair hydrated by seven water molecules and, (b) AT base pair coordinated by Na^+ cation.

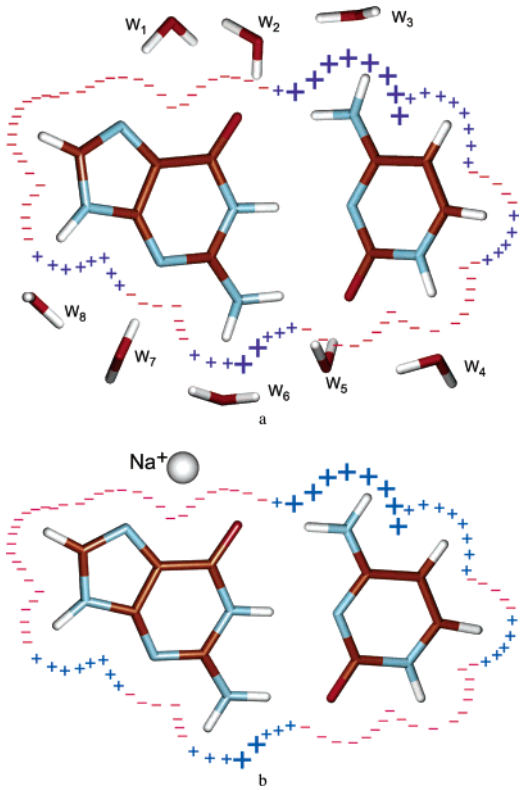


Figure 6. Optimal configuration of charges that should stabilize the G*C* base pair (blue area denotes positive charges; red area denotes negative charges) mapped on (a) GC base pair hydrated by eight water molecules and, (b) GC base pair coordinated by Na^+ cation.

(35) Kratochvil, M.; Sponer, J.; Hobza, P. *J. Am. Chem. Soc.* **2000**, *122*, 3495.
(36) Sponer, J.; Hobza, P. *Chem. Phys. Lett.* **1997**, *267*, 263. Sponer, J.; Hobza, P. *J. Phys. Chem. A* **2000**, *104*, 4592.

explained using the maps presented in Figures 5 and 6 and, therefore, applied directly to the currently studied base pairs. However, to clarify the discussion, it should be mentioned that only the data concerning the positions of water molecules obtained previously have been used to construct the initial aqueous hydration shells surrounding the AT, A*T*, GC, and G*C* DNA base pairs. Afterward, all structures of the polyhydrates have been reoptimized at the B3LYP/6-31G(d) level of theory. Since the structures of the aqueous hydration shells do not change significantly from the canonic to the rare tautomeric forms, we have presented only the hydrated structures of the canonic forms in Figures 5 and 6.

A comparison of the optimal charge positions with the positions of water molecules for the optimized hydration shells suggests that the water molecules are not able to stabilize the "rare" tautomers in either the case of the AT or the GC base pairs. This conclusion is based on the following results. The orientation of only two water molecules (W5 and W6) in the hydration shell of the AT base pair corresponds to the requirements of the optimal charge configuration. Therefore, only two out of seven water molecules are able to stabilize the "rare" configurations. Indeed, it has been found that the A*T* base pair hydrated with seven water molecules no longer represents a local minimum on the potential energy surface. Instead, it spontaneously transforms into a single-proton-transfer structure possessing a hydrogen atom transferred from the N3 atom of thymine to the N1 atom of adenine. This structure is 7.5 kcal/mol less stable than the canonic AT pair. These predictions are in complete correspondence with the results published in ref 37 and those obtained in the current investigation. Similar results have been obtained for the hydration of the GC base pair. Only one (W4) out of eight water molecules stabilizes the G*C* structure over the GC structure. As a result, both minima remain on the potential energy surface, but the canonic structure of the GC base pair becomes 17.5 kcal/mol more stable than the G*C* structure. This value is 6.7 kcal/mol higher (see Table 3) than that for the nonhydrated GC species at the same level of theory.

The maps of the optimal charge configuration are also able to provide information concerning the electrostatic influence of the metal cations. Since the most likely position of the metal cations is known for both DNA base pairs (N7 atoms of purine's component),[38] we have modeled the influence of the cations being placed only in this particular site. The map presented in Figure 5 suggests that one should expect some stabilization of the A*T* structure. For the GC base pair, one should observe a stabilization of the canonic structure rather than the rare tautomer. To verify this prediction, complexes of AT and GC base pairs with a Na^+ cation have been studied in this work. It was found that an interaction with a Na^+ cation results in greater stabilization of the rare A*T* form than the canonic AT structure (relative energy of the A*T* becomes 12.6 kcal/mol) and the canonic GC structure rather than the G*C* structure (relative energy of G*C* becomes 15.1 kcal/mol). These numbers, however, should be taken only as an indication of the trends in the stabilization of the canonic vs the rare structures of DNA base pairs by a metal cation since the geometrical parameters of the base pairs were frozen and only the position of the cation has been optimized (relaxation of the base pair geometry would lead to a different optimal charge configuration). The obtained data are completely in agreement with the predictions based on the optimal charge configurations.

4. Conclusions and Possible Biological Consequences

We have performed an extensive, high-level quantum-chemical study of the geometric and energetic characteristics that accompany the double-proton transfer in adenine−thymine and guanine−cytosine DNA base pairs. The following most important conclusions can be drawn.

Two sources of intrinsic nonplanarity of isolated DNA bases have been revealed. The first one originates from a partial sp^3 hybridization of the bases' amino groups. This source of nonplanarity is already recognized in the literature.[31] However, the current investigation suggests that the conclusion of internal nonplanarity of DNA base pairs should be extended to the canonic structure of the GC base pairs and most likely to the AT base pairs. The second source of nonplanarity has a dynamic nature. Due to the high flexibility of the base pairs originating from a number of intermolecular vibrations with low-lying frequencies, some excited vibrational levels are populated at room temperature. Therefore, at every moment of time a significant fraction of isolated DNA base pairs will possess a nonplanar geometry.

We predict the absence of local minimum corresponding to the isolated A*T* DNA base pair. In other words, the interaction of canonic forms of adenine and thymine will not promote the formation of "rare" adenine and thymine tautomers. Therefore, the interaction of adenine with thymine does not contribute at all to the formation of A* and T* structures in the DNA base pairs.

In contrast, the interaction of the canonic structures of guanine and cytosine can result in the formation of "rare" forms. The magnitude of the equilibrium constant that describes this equilibrium estimated at MP2/cc-pVTZ//MP2/6-31G(d) and MP2/infinite//MP2/6-31G(d) (see Table 3) is approximately 10^{-6}. As it was already mentioned, this equilibrium will be established instantly. It is worth noting that this value of the equilibrium constant is 3 orders of magnitude smaller than the one estimated earlier for the gas-phase tautomerization of single isolated guanine and cytosine bases. Therefore, the interaction of guanine with cytosine is the factor decreasing the gas-phase equilibrium concentration of "rare" tautomeric forms. Nevertheless, since the value of the equilibrium constant still exceeds the values of the frequency of spontaneous mutations (which range from 10^{-8} to 10^{-10}),[7a] the amount for the G* and C* species is expected to be greater than the observable natural rate of spontaneous mutations. This is due to the fact that external stabilization factors have not been taken into account in our calculations.

The data presented suggest that among the considered external stabilization factors the influence of water molecules is crucial for the stabilization of the canonic structure of the GC base pair. Since our model is still very simple compared to the real DNA molecule surrounding, we decided not to make the estimation of the equilibrium constant in this case. We are also

(37) Zhanpeisov, N. U.; Leszczynski, J. *J. Phys. Chem. B* **1998**, *102*, 9109. Zhanpeisov, N. U.; Leszczynski, J. *J. Phys. Chem. A* **1998**, *102*, 6167.
(38) Sponer, J.; Sabat, M.; Burda, J. V.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. B* **1999**, *103*, 2528. Sponer, J.; Burda, J. V.; Sabat, M.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. A* **1998**, *102*, 5951. Gresh, N.; Sponer, J. E.; Spackova, N.; Leszczynski, J.; Sponer, J. *J. Phys. Chem. B* **2003**, *107*, 8669.

aware that the amount of water molecules considered in this study is probably greater than that of the water molecules hydrating the DNA bases under physiological conditions.[30] Nevertheless, the stabilization energy of 6.7 kcal/mol obtained in this study should be considered as an indication of a strong additional factor which stabilizes the canonic G and C species when guanine and cytosine form a base pair. This conclusion is in good agreement with known experimental observations suggesting that the rate of spontaneous mutations occurring in the nonreplicated genome is significantly slower than the one taking place during DNA replication.[39]

Acknowledgment. This work was facilitated by ONR Grant No. N00014-98-1-0592, NIH Grant No. G1 2RR13459-21, by NSF-CREST Grant No. HRD-0318519, and partial support from Wroclaw University of Technology.

JA049155N

(39) Auerbach, C. *Mutation Research. Problems, Results and Perspectives*; John Wiley and Sons: New York, 1976.
(40) Dickerson, R. E. et al. *J. Mol. Biol.* **1989**, *205*, 787.