The effects of tautomerization and protonation on the adenine–cytosine mismatches: a density functional theory study

Hamid Reza Masoodi*, Sotoodeh Bagheri and Mahsa Abareghi

Faculty of Science, Department of Chemistry, Vali-e-Asr University of Rafsanjan, P.O. Box 77176, Rafsanjan, Iran

Communicated by Ramaswamy H. Sarma

(Received 17 April 2015; accepted 11 July 2015)

In the present work, we demonstrate the results of a theoretical study concerned with the question how tautomerization and protonation of adenine affect the various properties of adenine–cytosine mismatches. The calculations, in gas phase and in water, are performed at B3LYP/6-311++G(d,p) level. In gas phase, it is observed that any tautomeric form of investigated mismatches is more stabilized when adenine is protonated. As for the neutral mismatches, the mismatches containing amino form of cytosine and imino form of protonated adenine are more stable. The role of aromaticity on the stability of tautomeric forms of mismatches is investigated by NICS(1)ZZ index. The stability of mispairs decreases by going from gas phase to water. It can be explained using dipole moment parameter. The influence of hydrogen bonds on the stability of mismatches is examined by atoms in molecules and natural bond orbital analyses. In addition to geometrical parameters and binding energies, the study of the topological properties of electron charge density aids in better understanding of these mispairs.

Keywords: adenine; cytosine; tautomerization; protonation; mismatch

Introduction

Prototropic tautomerism of nucleotide bases is believed to induce the formation of rare tautomers of DNA bases (Colominas, Luque, & Orozco, 1996; Schreiber & González, 2007; Sponer & Hobza, 2003). In some cases, the complexes of rare tautomers are more stable than ground state one. For example, Samijlenko et al. indicated that the three rare adenine tautomers generate more stable complexes with CH₃COO⁻ and Na⁺ ions in comparison with the ground state case (Samijlenko, Krechkivska, Kosach, & Hovorun, 2004). The rare tautomeric forms can also cause mispairing during replication and, if it escapes repair, eventually may lead to a point mutation (Lippert & Gupta, 2007; Topal & Fresco, 1976). Furthermore, the mispairing can be incorporated into the DNA strand without disturbing the geometrical and energetic demands of the DNA double helix on the base-pair formation (Guckian, Krugh, & Kool, 2000; Harris et al., 2003). Das and Lyngdoh have proposed that the configuration of a solitary base pair is a major factor to predict whether it would occur at the wobble position or not (Das & Lyngdoh, 2014). Moreover, Watson and Crick explained that tautomerization alters the hydrogen bonding partners and therefore could enable mismatches to assume the structure of canonical base pairs (Watson & Crick, 1953). For instance, Wang et al. have observed adenine–cytosine mismatch using high-resolution X-ray crystallographic analysis of a DNA polymerase that catalyzes replication in crystal (Wang, Hellinga, & Beese, 2011). In adenine–cytosine mispair, if adenine (A) tautomerizes to its imino form (from hereon designated A*), the latter will pair with cytosine (C) instead of thymine (T). This behavior is also observed for imino form of cytosine (denoted as C*). On the other words, the C* incorrectly pairs with A. This mismatch was theoretically characterized by Danilov, Anisimov, Kurita, and Hovorun (2005) and Fonseca Guerra, Bickelhaupt, Saha, and Wang (2006). Danilov et al. indicated that the formation energy of A·C* is more favorable than A*·C. They also studied the nature of their hydrogen bonding based on the energy decomposition analysis of Morokuma–Kitaura and the reduced-variational-space methods to elucidate the possibility of the tautomerization of DNA base pairs (Danilov et al., 2005). Also, Fonseca Guerra et al. confirmed that A*·C mispair is a suitable candidate for the incorporation into DNA (Fonseca Guerra et al., 2006). Hovorun and Brovarets’ have indicated that exactly the A·C* base mispair is an active player of the point mutational events and is effectively dissociated by the replication machinery into the A and C* monomers in contrast to the A*·C base mispair, playing the mediated role of a provider of the A·C* base mispair in DNA that is synthesized (Brovarets’ & Hovorun, 2015c). In order to understand

*Corresponding author. Email: h.r.masoodi@vru.ac.ir

© 2015 Taylor & Francis
the origin of the spontaneous transitions, Hovorun and Brovarets’ have also proposed the new physicochemical mechanism on the basis of tautomerization, induced by the interaction of the DNA polymerase recognition center with the canonical nucleotide bases, of the pyrimidine bases in wobble base pairs A-C and G-T which transition into the pairs A-C* and G*-T accordingly (Brovarets’ & Hovorun, 2009). In other work, the tautomeric transition between wobble A-C(w) mismatch and Watson–Crick-like A-C*(WC) base mispair has theoretically studied by proceeding non-dissociatively via the sequential proton transfer between the bases through the planar, highly stable, and zwiterionic TS\(^{A-\rightarrow C.\rightarrow A.C(w)\rightarrow A.C*(WC)}\) transition state. Here, the authors have suggested that biologically significant A-C(w) → A-C*(WC) tautomerization is kinetically controlled pathway for the formation of the enzymatically competent Watson–Crick-like A-C*(WC) DNA base mispair in the essentially hydrophobic recognition pocket of the high-fidelity DNA polymerase responsible for the occurrence of spontaneous point AC/CA incorporation errors during DNA biosynthesis (Brovarets’ & Hovorun, 2015a). Recently, the interesting practical applications of non-Watson–Crick base pairs have also been investigated. Theoretical calculations predict A-C, A-G, and C-C mispairs as ideal to be incorporated into DNA sequences for developing a molecular switch. The experimental proof further substantiates the conceivability of an A-C mispair-based pH switch. In addition, it has been observed that pH-dependent molecular switches can be constructed harnessing the stability of a protonated A-C mispair (Jissy & Datta, 2014).

Protonation of nucleic acid bases plays a significant role in many biochemical (i.e., enzymatic reactions, stabilization of triplex structures) and mutagenic processes (Sinden, 1994; Wilcox, Ahluwalia, & Bevilacqua, 2011). The protonation/deprotonation of base in any canonical nucleoside significantly perturbs DNA-like conformations. Nevertheless, the ionization mechanism cannot explain entirely the nature of the spontaneous transitions (Brovarets’, Zhurakivsky, & Hovorun, 2010). Halder et al. have benchmarked the conventional formalisms for modeling the process of protonation and concluded that considering water as a proton donor might provide a physicochemically relevant picture of the relative order of protonation propensity of different sites of the nucleobases. Also, while the availability of stabilization possibilities determines the feasibility of occurrence of protonated bases, their occurrence context and specific functional roles are important factors determining their occurrence propensities (Halder, Halder, Bhattacharyya, & Mitra, 2014). Specifically, the protonation of adenine, adenosine, and adenine-containing nucleotides have been studied extensively in both gas and condensed phases by theoretical and experimental means (Russo, Toscano, Grand, & Jolibois, 1998; van Zundert et al., 2011). For example, Marian et al. produced protonated adenine by electrospray ionization (Marian, Nolting, & Weinkauf, 2005). In the other work, protonated adenine was produced after the ionization of cold neutral adenine dimers (Cheong et al., 2011). Also, Rajabi et al. investigated the structure of the protonated adenine dimer by infrared multiple photon dissociation spectroscopy and electronic structure calculations (Rajabi, Theel, Gillis, Beran, & Fridgen, 2009).

Base mispairing may also be brought about by a charged nucleobase. Such processes could in fact circumvent the necessity of tautomerization (Goodman, 1995). The present work deals with the mispairs involving adenine nucleobase, in its neutral and protonated forms, and cytosine. Here, we focus on the mispairs which can exist in a cognate base-pair conformation in the gas phase and in water. Protonated adenine comes in various tautomeric forms. This complexity arises from the fact that the two stable neutral adenine tautomers can be protonated at different sites. On the basis of experimental data and theoretical calculations, the four lowest ones in energy as well as their imino forms are shown in Scheme 1. All the other tautomers of the protonated adenine have considerably higher energies and are not studied here. Also, 1H,9H-adeninium (1H-9H-A(+)\textsuperscript{+}) is not considered in this manuscript whereas it leads to wobble base paring.

**Computational details**

All calculations have been implemented in the Gaussian 09 suite of programs (Frisch et al., 2009) at the spin-restricted level. The density functional theory (DFT) method where electron correlation is taken into account by means of non-local exchange and correlation functional is emerging as a cost-effective alternative to the time consuming (Hertwig & Koch, 1995; Leuilliot, Ghomi, Scalmani, & Berthier, 1999; Shukla & Leszczynski, 2000). Here, the geometries were optimized at the B3LYP/6-311++G(d,p) level of theory. The B3LYP/6-311++G(d,p) has been successfully applied on similar systems recently studied and have been verified to give accurate normal mode frequencies, characteristics of intra- and intermolecular hydrogen bonds and geometries (Brovarets’, Zhurakivsky, & Hovorun, 2013a; Brovarets’, Zhurakivsky, & Hovorun, 2013b; Matta, 2010; Ponomareva, Yurenko, Zhurakivsky, Mourik, & Hovorun, 2014; Wiberg, 2004; Yurenko, Zhurakivsky, Samijlenko, & Hovorun, 2011). Solute–solvent interactions can have dramatic effects on molecular energies, structures, and properties (Erikson, 2001; Leszczynski, 1995). If it were necessary to consider each solvent molecule as a separate molecule, the computational cost of modeling a solvent-mediated chemical reaction would grow prohibitively high. Modeling the solvent as a polarizable continuum, rather than individual molecules,
makes \textit{ab initio} computation feasible. In many cases, solvation effects can be computed very effectively in the framework of continuum solvation models (Dolney et al., 2000; Tomasi & Persico, 1994). In these models, the bulk of the solvent is represented as a structureless polarizable medium characterized mainly by its dielectric constant. Even when specific interactions require the introduction of some solvent molecules, strongly bound to the solute, the continuum picture is still very useful and often necessary (Cossi, Rega, Scalmani, & Barone, 2003). To examine the effects of water solvent on the stability of adenine–cytosine mismatches, the polarizable continuum model (PCM) (Barone & Cossi, 1998) was performed at the B3LYP/6-311++G(d,p) level. The PCM is one of the most reliable continuum solvation procedures and is commonly used in computational chemistry to model solvation effects (Kannappan, Suganthi, & Sathyanarayanamoorthi, 2014; Tomasi, Mennucci, & Cammi, 2005). Such a success is mainly because of the continuous improvements, both in terms of computational efficiency and generality, made by all the people involved in the PCM project. The result of these efforts is that nowadays, PCM, with all its different variants, is the default choice in many computational codes to couple a quantum–mechanical description of a molecular system with a continuum description of the environment (Mennucci, 2012). The symmetrical constraint (Cs) was considered during optimization. It has been previously shown that the effect of constraining the nucleobase geometry to Cs symmetry on the bond energy of natural Watson–Crick pairs (which are close to Cs symmetric) is in most cases very small or zero (Fonseca Guerra, Bickelhaupt, Snijders, & Baerends, 2000; Paragi et al., 2008). The nature of the stationary points was confirmed.
by frequency calculations at the same level of theory. All the structures are found to be at local minima on the potential energy surface. To obtain credible complexation energies in gas phase and in water, the single-point energies of the species were further refined at the MP2/6-311++G(d,p)//B3LYP/6-311++G(d,p) level of theory.

The aromaticity of adenine and cytosine tautomers was also investigated by nucleus-independent chemical shift (NICS) calculations. It is defined as the negative value of the absolute shielding computed at a ring center or at some other interesting points of a system. Rings with large negative NICS values are considered aromatic. As shown by Lazzeretti and Aihara (Aihara, 2002; Lazzeretti, 2000; Lazzeretti, 2004), NICS values at the geometrical center of the ring (NICS(0)) contain important spurious contributions from the in-plane tensor components that are not related to aromaticity. NICS(1) (1Å above/below the plane of the ring) essentially reflects π-effects and it is a better indicator of the ring current than the value at the center, because at this point, the effects of the local σ-bonding contributions are diminished (Corminboeuf, Heine, Seifert, Schleyer, & Weber, 2004; von Ragüe Schleyer et al., 2001). The out-of-plane component of the NICS(1), NICS(1)ZZ, correctly reflects the π-electron effects and probably is a better descriptor of aromaticity (Ebrahim, Habibi, Masoodi, & Gholipour, 2009). For determining NICS (1)ZZ values, the NMR calculations were performed at the B3LYP/6-311++G(d,p) level using GIAO formalism (Wolinski, Hinton, & Pulay, 1990).

The topological electron charge density was also analyzed by the atoms in molecules (AIM) method (Bader, 1990), using AIM2000 program (Biegler König & Schönbahm, 2002) on the wave functions obtained at the B3LYP/6-311++G(d,p) level.

Moreover, the population analysis was performed by the natural bond orbital (NBO) method (Reed, Curtiss, & Weinhold, 1988) at the B3LYP/6-311++G(d,p) level on the optimized structures using NBO program (Glendenning, Reed, Carpenter, & Weinhold, 1998) under Gaussian 09 program package.

Results and discussion

A comparison between amino and imino forms of investigated species

We first review the neutral and protonated adenine and also neutral cytosine tautomers. Starting from the two most stable tautomers of neutral adenine, 9H-adenine (9H-A) and 7H-adenine (7H-A), different tautomers of protonated adenine can be constructed. The most stable tautomers are 1H-9H-adeninium (1H-9H-A\(^+\)), 3H-7H-adeninium (3H-7H-A\(^+\)), 3H-9H-adeninium (3H-9H-A\(^+\)), and 7H-9H-adeninium (7H-9H-A\(^+\)). These tautomers are gathered in Scheme 1. Other tautomeric structures have higher energies and are not examined in the following discussion. As previously mentioned, a cognate base-pair conformation is not obtained in the presence of 1H-9H-A\(^+\) (see Scheme 2(a)). Hence, it is not studied in this manuscript. The dominant form of cytosine in biological systems is 1H-keto-amino tautomer (C) and it is considered here.

A typical adenine–cytosine mismatch is given in Scheme 2(b). It is observed that adenine–cytosine mismatch can exist in two forms: A⁻…C and A⁻…C*.

In order to compare the stability of amino and imino forms of adenine and cytosine in gas phase and in water, the cohesive energy or the binding energy (BE) per atom for any tautomer was calculated according to the following formula (Rohrer, 2001; Roohi & Bagheri, 2013):

\[
BE = \frac{aE_C + bE_N + dE_H + eE_O + fE_{H^+}}{a + b + d + e + f}
\]

where \(a, b, d, e, \) and \(f\) are the number of C, N, H, O, and H⁺ species, respectively. \(E_C, E_N, E_H, E_O, \) and \(E_{H^+}\) are the ground state total energies of C, N, H, O, and H⁺, respectively, and \(E\) is the total energy of the optimized tautomer. As shown from Table 1, the energetic order of neutral and protonated forms of adenine in gas phase and in water is as following: 3H-7H-A\(^+\) > 3H-9H-A\(^+\) > 7H-9H-A\(^+\) > 7H-9H-A\(^+\) > 3H-7H-A\(^+\) > 3H-9H-A\(^+\) > 9H-A > 7H-A > 9H-A > 7H-A > 9H-A > 7H-A. It is also found that the BE value of C is higher than that of C*. These findings indicate that the protonated forms of adenine are more stable than neutral ones. Also, the amino form is more possible than imino case for any species. Here, the aromaticity of six-membered ring of any species is examined at both amino and imino forms. With reference to the values of NICS(1)ZZ, it is concluded that the amino form is higher in aromaticity than the imino case (see Table 1). Thus, the more stability of amino form may be attributed to aromaticity factor. It is observed that the BE value increases by going from gas phase to water. As can be seen from Table 1, the dipole moment of adenine and cytosine tautomers in water is greater than that in gas phase. Whereas water is a polar solvent, the more stability of tautomers in water may be ascribed to hydration effects.

Theoretical study of adenine–cytosine mismatches

Here, the complexation energy has been calculated using equation

\[
\Delta E = E_{\text{complex}} - \sum E_{\text{mon}}
\]

where \(E_{\text{complex}}\) and \(E_{\text{mon}}\) are optimized energies of adenine–cytosine mismatch and each individual component monomer, respectively. The results in Table 2 indicate that the \(\Delta E\) values are negative. In other words, the complexation energies show an upward trend as is
observed for the \( A^* \cdots C, A \cdots C^*, A^{(+)\cdots}C, \) and \( A^{(+)\cdots}C^* \) mismatches. In gas phase, any tautomeric form of adenine–cytosine mispair is more stabilized when adenine is protonated. It is also found that the \( \Delta E \) value of \( A^* \cdots C / A^{(+)\cdots}C \) complexes is more negative than that of \( A \cdots C^*/A^{(+)\cdots}C^* \) ones, respectively, indicating that \( A^* \cdots C \) and \( A^{(+)\cdots}C \) mismatches are more stable.

The aromaticity of six-membered ring of adenine and cytosine changes during complexation (see Table 3). A meaningful relationship can be found between the stability of adenine–cytosine mismatches and aromaticity changes (\( \Delta \text{NICS}(1)_{ZZ} \)). From aromaticity point of view, it seems that the imino forms of adenine and cytosine are more suitable for mismatch formation. Whereas the \( \Delta \text{NICS}(1)_{ZZ} \) in adenine (at both A and A* forms) is greater than that in cytosine, it is expected that the \( A^{*}\cdots C/A^{(+)\cdots}C \) to be more possible than \( A\cdots C^*/A^{(+)\cdots}C^* \).

Table 1. The values of BE (kJ mol\(^{-1}\)), aromaticity (ppm), and dipole moment (Debye) in considered monomers*.

|       | BE               | NICS(1)\(_{ZZ}\) | \( \mu \) |
|-------|------------------|------------------|---------|
| 9H-A  | 463.89, 466.19   | −20.96, −20.20   | 2.45, 3.44 |
| 9H-A* | 460.57, 463.30   | −4.25, −6.41     | 3.88, 5.65 |
| 7H-A  | 461.58, 465.48   | −23.90, −21.73   | 6.97, 10.34 |
| 7H-A* | 459.28, 463.05   | −6.77, −7.44     | 3.12, 4.37 |
| 3H-7H-A\(^{(+)\cdots}\) | 495.75, 509.72 | −16.59, −16.14   | 2.73, 3.65 |
| 3H-7H-A\(^{(+)\cdots}\) | 490.94, 505.96 | −1.40, −3.30     | 2.10, 3.03 |
| 3H-9H-A\(^{(+)\cdots}\) | 495.37, 509.33 | −14.57, −15.26   | 4.07, 5.99 |
| 3H-9H-A\(^{(+)\cdots}\) | 489.21, 505.19 | 0.64, −2.56      | 8.32, 11.98 |
| 7H-9H-A\(^{(+)\cdots}\) | 493.54, 508.69 | −21.25, −20.11   | 7.79, 10.66 |
| 7H-9H-A\(^{(+)\cdots}\) | 491.02, 506.26 | −3.10, −5.25     | 5.59, 7.52 |
| C     | 448.51, 453.03   | −4.74, −6.31     | 6.75, 9.65 |
| C*    | 448.01, 451.14   | 0.71, −0.35      | 4.94, 6.63 |

Note: *The normal and italic data correspond to calculations in gas phase and in water, respectively.
mispairs, respectively. This result is confirmed by the energetic order of investigated complexes. Compared to gas phase, the absolute values of $\Delta E$ ($|\Delta E|$) decrease in water. Here, the role of dipole moment ($\mu$) is investigated. It is observed that the investigated mismatches have lower dipole moments than the total dipole moments of the monomers. The stronger interaction of polar solvent (water) with monomers leads to a decrease in stability of adenine–cytosine mismatches.

In addition, we have calculated the interaction ($\Delta E_{\text{int}}$) and deformation ($\Delta E_{\text{def}}$) energies of the considered systems (see Table 2). The $\Delta E_{\text{int}}$ values were evaluated as the difference between the complexation energy and the sum of the energies of the separated monomers, with the same geometries as they have in the complex (i.e., frozen geometries). The difference between the $\Delta E$ and the $\Delta E_{\text{def}}$ is the deformation energy of the monomers. While the value of $\Delta E_{\text{int}}$ is negative and makes a positive contribution to the $\Delta E$, this behavior is reversed for $\Delta E_{\text{def}}$. For any investigated system, the $\Delta E_{\text{int}}$ and $\Delta E_{\text{def}}$ values in $A^\ast\ldots C/A(+)\ldots C$ are, respectively, greater than those in $A\ldots C/A(+)\ldots C$. It is observed that the contribution of $\Delta E_{\text{int}}$ is dominant in all complexes. Also, the difference between $\Delta E_{\text{int}}$ and $\Delta E_{\text{def}}$ is amplified in the presence of protonated adenine. Considering the difference between $\Delta E_{\text{int}}$ and $\Delta E_{\text{def}}$ it is expected that the $A^\ast\ldots C$ and $A(+)\ldots C$ mismatches to be more stable in neutral and positively charged adenine–cytosine mismatches. This result is in accord with the order of $\Delta E$ values.

The adenine–cytosine mismatch is stabilized by three hydrogen bonds (H-bonds). As shown in Scheme 2(b), they are denoted by HB1, HB2, and HB3. In the following, the H-bonds are examined using geometrical, topological, and energetic parameters. The bond length of $A\ldots H$ (A is proton acceptor) is often treated as a rough measure of the strength of H-bond (Checińska & Grabowski, 2006). For studied mismatches, the bond lengths of H-bonds are gathered in Table S1 (see supplementary materials). It can be seen that the length of HB1 in $A^\ast\ldots C$ is shorter than that in $A\ldots C^\ast$. An opposite order is observed for mismatches containing protonated adenine. These observations can be explained on the basis of atomic charges of N(1) and H(1) atoms. The natural charges (the nuclear charge minus the summed natural populations of natural atomic orbitals on the atom) on the H and N atoms ($q_H$, $q_N$) obtained using NBO calculations at the B3LYP/6-311++G(d,p) level of theory are given in Table S2 (see supplementary materials). The N(1) and H(1) atoms involved in HB1 have negative and positive charges, respectively. The increase in the absolute value of $q_{N(1)}$ ($|q_{N(1)}|$) and $q_{H(1)}$ leads to stronger HB1. The protonation of adenine leads to increasing $q_{H(1)}$ and decreasing $|q_{N(1)}|$ in its amino and imino forms, respectively. Thus, the HB1 is, respectively, strengthened/weakened in $A^\ast\ldots C/A(+)\ldots C$ forms. Compared to $A\ldots C^\ast$, it is also found that the length of HB2 in $A^\ast\ldots C$ is shortened. This difference is intensified in the presence of protonated adenine. In $A^\ast\ldots C$ mismatch, the H(2) and N(2) atoms involved in HB2 correspond to adenine and cytosine, respectively. This contribution is reversed for $A\ldots C^\ast$ mispair. The increase

Table 2. The values of complexation, interaction and deformation energies (kJ mol$^{-1}$), and the dipole moment changes (Debye) on complexation$^a$.

|   | $\Delta E$  | $\Delta E_{\text{int}}$ | $\Delta E_{\text{def}}$ | $\Delta \mu$ |
|---|------------|------------------------|------------------------|-------------|
| 9H-A...C$^+$ | $-54.75$, $-34.79$ | $-66.80$, $-46.31$ | $-60.54$, $-38.42$ | $5.78$, $3.63$ | $-4.30$, $-5.82$ |
| 9H-A...C$^-\ast$ | $-84.46$, $-41.59$ | $-94.81$, $-55.79$ | $-94.45$, $-46.51$ | $9.99$, $4.92$ | $-6.84$, $-10.29$ |
| 7H-A...C$^+$ | $-54.31$, $-34.76$ | $-64.95$, $-46.68$ | $-61.02$, $-38.79$ | $6.71$, $4.03$ | $-2.85$, $-3.61$ |
| 7H-A...C$^-\ast$ | $-77.74$, $-40.76$ | $-88.32$, $-54.98$ | $-86.51$, $-45.71$ | $8.77$, $4.95$ | $-0.21$, $-0.28$ |
| 3H-7H-A$^-\ast$, C$^+$ | $-90.92$, $-37.01$ | $-103.34$, $-50.68$ | $-102.38$, $-42.00$ | $11.46$, $4.99$ | $-2.08$, $-3.12$ |
| 3H-7H-A$^-\ast$, C$^-\ast$ | $-156.40$, $-57.04$ | $-168.01$, $-72.33$ | $-180.33$, $-67.49$ | $23.93$, $10.46$ | $-5.50$, $-7.91$ |
| 3H-9H-A$^-\ast$, C$^+$ | $-87.80$, $-36.79$ | $-100.12$, $-49.45$ | $-97.24$, $-41.19$ | $9.44$, $4.40$ | $-2.15$, $-3.77$ |
| 3H-9H-A$^-\ast$, C$^-\ast$ | $-163.88$, $-58.07$ | $-176.46$, $-73.39$ | $-189.65$, $-69.18$ | $25.77$, $11.12$ | $-11.09$, $-15.19$ |
| 7H-9H-A$^-\ast$, C$^-\ast$ | $-83.59$, $-36.67$ | $-94.25$, $-48.87$ | $-93.82$, $-40.69$ | $10.23$, $4.02$ | $-0.07$, $-0.92$ |

Table 3. The aromaticity changes of adenine and cytosine bases (ppm) on complexation.$^a$

|   | $\text{ANICS(1)zz}$ (adenine) | $\text{ANICS(1)zz}$ (cytosine) |
|---|-------------------------------|-------------------------------|
| 9H-A...C$^+$ | $2.66$, $2.10$ | $0.14$, $0.37$ |
| 9H-A...C$^-\ast$ | $-2.8$, $-1.03$ | $1.41$, $1.58$ |
| 7H-A...C$^+$ | $2.87$, $2.47$ | $0.33$, $0.37$ |
| 7H-A...C$^-\ast$ | $-2.83$, $-1.02$ | $1.38$, $1.65$ |
| 3H-7H-A$^-\ast$, C$^+$ | $3.36$, $2.42$ | $-0.89$, $0.36$ |
| 3H-7H-A$^-\ast$, C$^-\ast$ | $-3.23$, $-0.86$ | $0.34$, $1.44$ |
| 3H-7H-A$^-\ast$, C$^+$ | $3.38$, $2.38$ | $-0.80$, $0.09$ |
| 3H-9H-A$^-\ast$, C$^-\ast$ | $-2.20$, $-0.63$ | $0.40$, $1.37$ |
| 7H-9H-A$^-\ast$, C$^+$ | $3.53$, $2.35$ | $-0.73$, $0.21$ |
| 7H-9H-A$^-\ast$, C$^-\ast$ | $-2.97$, $-1.03$ | $0.71$, $1.43$ |

Note: $^a$The normal and italic data correspond to calculations in gas phase and in water, respectively.

$^b$The $\Delta E$ values were obtained at MP2/6-311++G(d,p)/B3LYP/6-311++G(d,p) level.
in $q_{H(2)}$ and $|\rho_{O(2)}|$ leads obviously to strengthen HB2. In gas phase, the $q_{H(2)}$ in C* and $|\rho_{O(2)}|$ in C are, respectively greater than those in A* and A. In comparison with $q_{H(2)}$ values in C* and A*, the difference in $|\rho_{O(2)}|$ values of A and C is further revealed. Hence, it is expected that the HB2 in A* ...C to be stronger than that in A ...C*. In water, the values of $q_{H(2)}$ in A* and $|\rho_{O(2)}|$ in C are both higher than those in C* and A, respectively. The protonation of adenine is accompanied by increasing $q_{H(2)}$ in A(+) ...C and decreasing $|\rho_{O(2)}|$ in A(+) ...C. Thus, HB2 in A(+) ...C is obviously stronger than that in A ...C*. The HB3 includes H and O atoms of adenine and cytosine, respectively. It is observed that the length of HB3 in A* ...C is shorter than that in A ...C*. This behavior can be ascribed to natural atomic charges on O (3) and H(3) atoms. As shown in Table S2, the $q_{H(2)}$ and $|\rho_{O(2)}|$ values in A* and C are, respectively, higher than those in A and C*. The protonation of adenine intensifies this difference.

Topological criteria were also proposed to detect the existence of H-bond (Grabowski, Sokalski, & Leszczynski, 2005; Popelier, 2000). The bond critical points (BCPs) of the N ...H and O ...H interactions were found and the features of them were analyzed since it is well known that characteristics of BCPs, such as the electron densities ($\rho_{HB}$), their laplacians ($\nabla^2 \rho_{HB}$), and the energetic properties ($H_{HB}$) of BCPs, allow us to categorize interactions, and these topological parameters are also treated as measures of H-bond strength (Domagala & Grabowski, 2005; Gálvez, Gómez, & Pacios, 2003). The values of $\rho_{HB}$, $\nabla^2 \rho_{HB}$, and $H_{HB}$ at the BCPs were evaluated by the means of AIM approach at the B3LYP/6-311++G(d,p) level of theory (see supplementary materials, Table S3). A typical molecular graph is shown in Figure 1. In general, the $\rho_{HB}$ value can be a useful parameter for describing the strength of H-bonds (Grabowski, 2006). The order of $\rho_{HB1}$ values is identical in gas phase and in water. In neutral complexes, the value of $\rho_{HB1}$ in A* ...C is higher than that in A ...C*. An opposite behavior is observed in the presence of protonated adenine. Rozas et al. have introduced a classification of H-bonds according to their strength. Weak H-bonds show both $\nabla^2 \rho_{HB}$ and $H_{HB}$ values positive; for medium H-bonds, $\nabla^2 \rho_{HB} > 0$ and $H_{HB} < 0$, and also for strong H-bonds, the $\nabla^2 \rho_{HB}$ as well as $H_{HB}$ are negative (Rozas, Alkorta, & Elguero, 2000). In gas phase, $\nabla^2 \rho_{HB1}$ and $H_{HB1}$ values in A ...C* form are positive, indicating that HB1 considered as weak H-bonds while A* ...C form is characterized by the positive $\nabla^2 \rho_{HB1}$ and negative $H_{HB1}$ showing that HB1 may be classified as medium H-bonds. This order is reversed for positively charged complexes. In water, it is also observed that the HB1 in all neutral complexes can be considered as medium H-bonds. The $\rho_{HB2}$ values in gas phase and in water indicate that HB2 in A* ...C form is stronger than that in A ...C*. This difference is intensified in the presence of protonated adenine. According to Rozas classification, HB2 in all neutral complexes is of medium type. The presence of protonated adenine increases $\rho_{HB2}$ in A(+) ...C form and decreases it in A(+) ...C* case so that HB2 in A(+) ...C and A(+) ...C* forms can be classified as medium and weak H-bonds, respectively. In any investigated system, the $\rho_{HB2}$ decreases by going from gas phase to water. Considering $\rho_{HB3}$ value, it is concluded that HB3 in A* ...C form is stronger than that in A ...C* case. This difference is intensified by protonation of adenine. On the basis of $\nabla^2 \rho_{HB3}$ and $H_{HB3}$ values, HB3 can be characterized as weak H-bonds in all complexes. Compared to gas phase, the value of $\rho_{HB3}$ in A* ...C form increases in water while a reverse behavior is observed for A ...C*. Also, $\rho_{HB3}$ values in both forms of A(+) ...C and A(+) ...C* decrease by going from gas phase to water. The AIM results are confirmed by H-bond lengths.

The conventional H-bond energies were evaluated by the empirical Iogansen’s formula (Iogansen, 1999):

$$\text{E}_{\text{HB}} = \frac{1}{2} \left( q_{H} \cdot q_{O} \right)$$

Figure 1. The molecular graphs for 9H-adenine–cytosine mismatch obtained using AIM analysis. Small red spheres, small yellow spheres, and lines represent BCPs, ring critical points, and bond paths, respectively.
where \( E_{\text{log HB}} \) is a hydrogen bond energy in kcal \( \text{mol}^{-1} \), \( \Delta \nu \) is the magnitude of the red shift (relative to the free molecule) in \( \text{cm}^{-1} \) of the stretching mode of H-bonded AH groups in an AH…B complex (A and B = N, O). The partial deuteration was applied to minimize the effect of mechanical resonances (Brovarets’ & Hovorun, 2010).

The logansen’s formula has frequently been applied for studying various base pairs (Brovarets’ & Hovorun, 2013b; Brovarets’, Yurenko, & Hovorun, 2015; Brovarets’, Zhurakivsky, & Hovorun, 2015). For example, Brovarets’ et al. have studied Hyp−Hyp base pair at B3LYP/6-311++G(d,p) level of theory, where Hyp and Hyp’ are keto and enol tautomers of the hypoxanthine, respectively. They estimated the energy of NH…N H-bond in Hyp−Hyp base pair using Iogansen \( E_{\text{log}} \). On the basis of their results, the energy of NH…N H-bond in gas phase is obtained equal to 29.16 kJ \( \text{mol}^{-1} \) (Brovarets’, Zhurakivsky & Hovorun, 2013c). In other work, the logansen’s formula has been used to examine NH…N H-bond in DNA base mispair containing amino and imino tautomers of cytosine. The results obtained at B3LYP/6-311++G(d,p) level in gas phase showed that the energies of upper and lower NH…N H-bonds correspond to 27.87 and 27.07 kJ \( \text{mol}^{-1} \) (Brovarets’ & Hovorun, 2013a). In the present work, the results in Table 4 indicate that the ranges of \( E_{\text{HB1}} \) and \( E_{\text{HB2}} \), in gas phase, are from 19.87 to 35.93 and 11.92 to 40.46 kJ \( \text{mol}^{-1} \), respectively. In water, these ranges correspond to 21.25–28.90 and 15.06–34.00 kJ \( \text{mol}^{-1} \).

Nevertheless, the strength of CH…O H-bond cannot be examined using Iogansen’s formula. Thus, the energies of the intermolecular H-bonds were also evaluated by the empirical Espinosa–Molins–Lecomte (EML) formula (Espinosa, Molins, & Lecomte, 1998) based on the electron density distribution at the BCPs of the H-bonds:

\[
E_{\text{EML HB}} = 0.5 V(r)
\]

where \( V(r) \) is the value of a local potential energy at the BCPs. The values of \( E_{\text{EML HB}} \) QUOTE \( E_{\text{HB}} \), and \( V(r) \) are reported in Tables 4 and S3. Nikolaenko et al. indicated that the EML formula usually overestimates the energy of the classical H-bonds (Nikolaenko, Bulavin, & Hovorun, 2012). Whereas the evaluation of energy of CH…O H-bond is only possible using EML formula, however, the correlations between \( E_{\text{EML HB}} \) values and energetic, geometrical and topological properties are considered in the following sections.

As observed from Table 4, the energy values of CH…O H-bond lie within the range 2.30–15.48/1.74–9.93 kJ \( \text{mol}^{-1} \) in gas phase/in water. Brovarets’ et al. have investigated the physico-chemical nature and energetic of the non-conventional CH…O H-bonds in the biologically important natural nucleobase pairs using EML formula. Their results lie within 1.88–16.27/2.59–17.55 kJ \( \text{mol}^{-1} \) at DFT/MP2 levels of theory (Brovarets’, Yurenko & Hovorun, 2014).

Table 4. The values of H-bond energy (kJ \( \text{mol}^{-1} \)) estimated by EML and logansen’s formulas\(^{a,b}\).

| \( E_{\text{HB1}} \) | \( E_{\text{HB2}} \) | \( E_{\text{HB3}} \) | \( \Sigma E_{\text{HB}} \) |
|----------------|----------------|----------------|----------------|
| \( 9\text{H}−\text{A}\ldots\text{C}^{+} \) | −26.08, −28.59 | −39.29, −31.69 | −3.80, −3.10 | −69.17, −63.38 |
| 23.36, 23.44 | 24.13, 21.25 | − | 47.49, 44.69 |
| \( 9\text{H}−\text{A}^{+}\ldots\text{C} \) | −36.96, −29.17 | −41.29, −40.25 | −3.88, −4.28 | −82.13, −73.71 |
| 28.35, 24.81 | 22.61, 23.20 | − | 50.96, 48.01 |
| \( 7\text{H}−\text{A}^{+}\ldots\text{C}^{+} \) | −27.41, −29.90 | −37.89, −31.08 | −3.71, −3.12 | −69.01, −64.09 |
| 21.91, 23.25 | 23.55, 20.97 | − | 45.46, 44.21 |
| \( 7\text{H}^{+}\ldots\text{C} \) | −33.45, −28.18 | −39.67, −39.83 | −3.90, −4.37 | −77.03, −72.37 |
| 26.96, 24.28 | 22.37, 23.22 | − | 49.33, 47.51 |
| \( 3\text{H}−\text{H}^{+}\ldots\text{A}^{+}\ldots\text{C}^{+} \) | −58.07, −41.47 | −30.03, −24.77 | −2.67, −2.07 | −90.77, −86.30 |
| 35.93, 28.90 | 12.45, 15.22 | − | 48.38, 44.13 |
| \( 3\text{H}−\text{H}^{+}\ldots\text{A}^{+}\ldots\text{C} \) | −16.11, −20.32 | −91.01, −72.80 | −15.48, −9.93 | −122.60, −103.05 |
| 20.25, 21.25 | 38.24, 33.39 | − | 58.49, 54.64 |
| \( 3\text{H}−\text{H}^{+}\ldots\text{A}^{+}\ldots\text{C}^{+} \) | −52.38, −41.17 | −29.35, −23.30 | −2.49, −1.74 | −84.22, −66.21 |
| 33.55, 28.24 | 11.92, 15.06 | − | 45.47, 43.30 |
| \( 3\text{H}−\text{H}^{+}\ldots\text{A}^{+}\ldots\text{C} \) | −18.73, −21.28 | −101.12, −74.94 | −14.80, −9.52 | −134.65, −105.74 |
| 21.55, 21.65 | 40.46, 34.00 | − | 62.01, 55.64 |
| \( 7\text{H}−\text{H}^{+}\ldots\text{A}^{+}\ldots\text{C} \) | −55.91, −46.34 | −27.65, −27.10 | −2.30, −2.55 | −85.86, −65.98 |
| 34.80, 27.04 | 14.64, 17.94 | − | 49.44, 44.99 |
| \( 7\text{H}−\text{H}^{+}\ldots\text{A}^{+}\ldots\text{C} \) | −16.02, −22.53 | −68.63, −51.59 | −10.62, −6.47 | −95.27, −80.59 |
| 19.87, 22.02 | 32.43, 27.07 | − | 52.31, 49.09 |

Notes: \(^{a}\)The normal and italic data correspond to calculations in gas phase and in water, respectively.
\(^{b}\)The results in first and second rows were obtained using EML and logansen’s formulas, respectively.
The effects of tautomerization and protonation on the adenine–cytosine mismatches

Table 4 indicate that for any neutral complex, the trend in $E_{\text{EML,HB}}$ values is $E_{\text{EML,HB2}} > E_{\text{EML,HB1}} > E_{\text{EML,HB3}}$. Moreover, the values of $E_{\text{EML,HB1}}$, $E_{\text{EML,HB2}}$, and $E_{\text{EML,HB3}}$ in A*…C form are greater than those in A*…C* case. The protonation of adenine has a potential impact on the strength of H-bonds. In positively charged complexes, the order of $E_{\text{EML,HB}}$ values is different in A*…C and A*…C* forms. The $E_{\text{EML,HB}}$ value in A*…C* increases in the following order $E_{\text{EML,HB1}} > E_{\text{EML,HB2}} > E_{\text{EML,HB3}}$, while this trend in A*…C is $E_{\text{EML,HB3}} > E_{\text{EML,HB1}} > E_{\text{EML,HB2}}$. As shown in Scheme 2(b), the N and H atoms of A* contribute to HB1 and HB2, respectively. This behavior is reversed for A. With regard to the natural atomic charges, the protonation of adenine increases/decreases the tendency of H/N atoms to form H-bond. Whereas HB3 consists of the H atom of adenine, the protonation of adenine may be expected to have a considerably favorable influence on the strength of HB3.

Surprisingly, in comparison with neutral complexes, the $E_{\text{EML,HB3}}$ value decreases in A*…C* form and increases in A*…C case. These findings may be attributed to cooperativity effects (Brovarets’ & Hovorun, 2014; Brovarets’ & Hovorun, 2015b; Brovarets’, Zhuravkivsky & Hovorun, 2014). In systems with multiple H-bonds, the strength of one H-bond is affected by an adjacent H-bond. Here, the reinforcement/weakening of HB2 in A*…C/A*…C* forms, respectively, lead to increasing/decreasing $E_{\text{EML,HB1}}$ value. This cooperativity effect is confirmed by geometrical and topological parameters of HB3. In the following, the total BE of H-bonds ($\sum E_{\text{EML,HB}}$) is calculated for any mismatch. In accord with $|\Delta E|$, the absolute value of $\sum E_{\text{EML,HB}}$ ($\sum E_{\text{EML,HB}}$) in A*…C/A*…C* is greater than that in A*…C*/A*…C*. In comparison with gas phase, $\sum E_{\text{EML,HB}}$ decreases in water. As shown in Figure 2(a), a direct relationship can be found between $\sum E_{\text{EML,HB}}$ and $|\Delta E|$ values.

To gain more insight into the influences of tautomerization and protonation on H-bonds, the NBO analysis has been performed on the investigated complexes. The NBO calculations show that the most important donor–acceptor interaction in HB1 and HB2 is LpN → $\sigma^*_{N-H}$ while it is LpO → $\sigma^*_{C-H}$ in HB3 (see Scheme 2(b)). The energy values of these interactions ($E^{(2)}_{\text{HB}}$) are gathered in Table S4 (see supplementary materials). For neutral complexes in gas phase, the value of $E^{(2)}_{\text{HB1}}$ in A*…C form is greater than that in A*…C* case. An opposite behavior is observed for positively charged mismatches. With the exception of 7H-A*…C and 7H-A*…C* forms, the order of $E^{(2)}_{\text{HB1}}$ values in water is similar to gas phase. In water, the value of $E^{(2)}_{\text{HB1}}$ in 7H-A*…C* is greater than that in 7H-A*…C. The difference of $E^{(2)}_{\text{HB1}}$ values in tautomeric forms of any mismatch is reduced in water. It may be attributed to solvent effects on donor–acceptor interactions. In comparison with any mismatch in gas phase, the LpN(1) → $\sigma^*_{N-H(1)}$ interaction in A*…C*/A*…C is reinforced/weakened in water so that $E^{(2)}_{\text{HB1}}$ value in 7H-A*…C* would be even greater than that in 7H-A*…C. The value of $E^{(2)}_{\text{HB2}}$ in A*…C is also higher than that in A*…C*. This difference is intensified by the protonation of adenine. Compared to gas phase, the LpN(2) → $\sigma^*_{N-H(2)}$ interactions in A*…C/A*…C* mismatches become stronger/weaker in water, while this interaction in both forms of A*…C and A*…C* is generally weakened by going from gas phase to water. Similar to $E^{(2)}_{\text{HB2}}$, the value of $E^{(2)}_{\text{HB3}}$ in A*…C is greater than that in A*…C*. This difference is enhanced in the presence of protonated adenine. As going from gas phase to water, LpO(3) → $\sigma^*_{C-H(3)}$ interaction is weakened/strengthened in A*…C*/A*…C* forms while it is weakened in both forms of A*…C* and A*…C*. For any mismatch, the total $E^{(2)}_{\text{HB}}$ ($\sum E^{(2)}_{\text{HB}}$) is calculated. In agreement with $|\Delta E|$ and $\sum E_{\text{EML,HB}}$, the value of $\sum E^{(2)}_{\text{HB}}$ in A*…C/A*…C forms is higher than that in A*…C*/A*…C*. In comparison with gas phase, it is observed that $\sum E^{(2)}_{\text{HB}}$ decreases in water. A direct relationship can be seen between $\sum E^{(2)}_{\text{HB}}$ and $|\Delta E|$ values in Figure 2(b).
Conclusions
In this manuscript, the effects of tautomerization and protonation of adenine on the various properties of adenine–cytosine mismatches have been investigated. Here, we examined the mispairs which can exist in a cognate base-pair conformation in the gas phase and in water. At first, the neutral and protonated adenine and also neutral cytosine tautomers were studied. With reference to the values of NICS(1)\_zz, it was concluded that the amino forms of adenine and cytosine is higher in aromaticity than the imino case. Thus, the more stability of amino form may be attributed to aromaticity factor. It was observed that the stability of species increases by going from gas phase to water. The dipole moment of adenine and cytosine tautomers in water is greater than that in gas phase. Whereas water is a polar solvent, the more stability of tautomers in water may be ascribed to hydration effects.

The results show that the any tautomeric form of adenine–cytosine mispair in gas phase is more stabilized when adenine is protonated. The ΔE value of A*…C/A(+)…C complexes is, respectively, more negative than that of A…C/A(+)…C* ones, indicating that A*…C and A(+)…C mismatches are more stable. From aromaticity point of view, it seems that the imino forms of adenine and cytosine are more suitable for mismatch formation. Whereas the changes of NICS(1)\_zz in adenine (at both A and A* forms) are greater than that in cytosine, it is expected that the A*…C and A(+)…C to be more possible than their corresponding mispairs. This result is confirmed by the energetic order of investigated complexes. Compared to gas phase, the |ΔE| values decrease in water. It is observed that the investigated mismatches have lower dipole moments than the total dipole moments of the monomers. The stronger interaction of polar solvent (water) with monomers leads to decreasing stability of adenine–cytosine mismatches. For any investigated mispair, the interaction and deformation energies were calculated. These values in A*…C/A(+)…C are, respectively, greater than those in A…C*/A(+)…C*. It is observed that the contribution of interaction energy is dominant in all complexes.

The adenine–cytosine mispair is stabilized by three H-bonds. The nature and strength of H-bonds were characterized by AIM analysis. The energies of the intermolecular H-bonds were evaluated by the empirical logansen’s and EML formulas. Also, the most important donor–acceptor interactions were studied using NBO analysis. The excellent correlations were found between the complexation energy and the results of AIM and NBO analyses.

Supplementary material
The supplementary material for this paper is available online at http://dx.doi.org/10.1080/07391102.2015.1072734.

Disclosure statement
No potential conflict of interest was reported by the authors.

References
Aihara, J. (2002). Nucleus-independent chemical shifts and local aromaticities in large polycyclic aromatic hydrocarbons. Chemical Physics Letters, 365, 34–39.
Bader, R. F. W. (1990). Atoms in molecules: A quantum theory. Oxford: Oxford University Press.
Barone, V., & Cossi, M. (1998). Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. Journal of Physical Chemistry A, 102, 1995–2001.
Biegler König, F., & Schönbohm, J. (2002). Update of the AIM2000-program for atoms in molecules. Journal of Computational Chemistry, 23, 1489–1494.
Brovarets’, O. O., & Hovorun, D. M. (2009). Physicochemical mechanism of the wobble DNA base pairs Gua·Thy and Ade·Cyt transition into the mismatched base pairs Gua‘Thy and Ade‘Cyt’ formed by the mutagenic tautomers. Ukrainica Bioorganica Acta, 2, 12–18.
Brovarets’, O. O., & Hovorun D. M. (2010). Stability of mutagenic tautomers of uracil and its halogen derivatives: The results of quantum-mechanical investigation. Biopolymers and Cell, 26, 295–298.
Brovarets’, O. O., & Hovorun, D. M. (2013a). Atomistic nature of the DPT tautomerisation of the biologically important C·C DNA base mispair containing amino and imino tautomers of cytosine: A QM and QTAIM approach. Physical Chemistry Chemical Physics, 15, 20091–20104.
Brovarets’, O. O., & Hovorun, D. M. (2013b). Atomistic understanding of the C·T mismatched DNA base pair tautomerization via the DPT: QM and QTAIM computational approaches. Journal of Computational Chemistry, 34, 2577–2590.
Brovarets’, O. O., & Hovorun, D. M. (2014). Why the tautomerization of the G·C Watson–Crick base pair via the DPT does not cause point mutations during DNA replication? QM and QTAIM comprehensive analysis. Journal of Biomolecular Structure & Dynamics, 32, 1474–1499.
Brovarets’, O. O., & Hovorun, D. M. (2015a). Tautomeric transition between wobble A·C DNA base mispair and Watson–Crick-like A·C mismatch: Microstructural mechanism and biological significance. Physical Chemistry Chemical Physics, 17, 15103–15110.
Brovarets’, O. O., & Hovorun, D. M. (2015b). The nature of the transition mismatches with Watson–Crick architecture: The G*·T or G·T* DNA base mispair or both? A QM/QTAIM perspective for the biological problem. Journal of Biomolecular Structure & Dynamics, 33, 925–945.
Brovarets’, O. O., & Hovorun, D. M. (2015c). The physicochemical essence of the purine-pyrimidine transition mismatches with Watson–Crick geometry in DNA: A·C’ versus A’·C. A QM and QTAIM atomistic understanding. Journal of Biomolecular Structure & Dynamics, 33, 28–55.
Brovarets’, O. O., Yurenko, Y. P., & Hovorun, D. M. (2014). Intermolecular CH⋅⋅⋅O/N H-bonds in the biologically important pairs of natural nucleobases: A thorough quantum-chemical study. *Journal of Biomolecular Structure & Dynamics*, 32, 993–1022.

Brovarets’, O. O., Yurenko, Y. P., & Hovorun, D. M. (2015). The significant role of the intermolecular CH⋅⋅⋅O/N hydrogen bonds in governing the biologically important pairs of the DNA and RNA modified bases: A comprehensive theoretical investigation. *Journal of Biomolecular Structure & Dynamics*, 33, 1624–1652.

Brovarets’, O. O., Zhurakivsky, R. O., & Hovorun, D. M. (2010). Is there adequate ionization mechanism of the spontaneous transitions? Quantum-chemical investigation. *Biopolymers and Cell*, 26, 398–405.

Brovarets’, O. O., Zhurakivsky, R. O., & Hovorun, D. M. (2013a). DPT tautomerization of the long A·A Watson–Crick base pair formed by the amino and imino tautomers of adenine: Combined QM and QTAIM investigation. *Journal of Molecular Modeling*, 19, 4223–4237.

Brovarets’, O. O., Zhurakivsky, R. O., & Hovorun, D. M. (2013b). The physico-chemical “anatomy” of the tautomeration through the DPT of the biologically important pairs of hypoxanthine with DNA bases: QM and QTAIM perspectives. *Journal of Molecular Modeling*, 19, 4119–4137.

Brovarets’, O. O., Zhurakivsky, R. O., & Hovorun, D. M. (2013c). The physico-chemical mechanism of the tautomerisation via the DPT of the long Hyp·⋅··Hyp Watson–Crick base pair containing rare tautomer: A QM and QTAIM detailed look. *Chemical Physics Letters*, 578, 126–132.

Brovarets’, O. O., Zhurakivsky, R. O., & Hovorun, D. M. (2014). Does the tautomeric status of the adenine bases change upon the dissociation of the A*·Asyn Topal-Fresco DNA mismatch? A combined QM and QTAIM atomic insight. *Physical Chemistry Chemical Physics*, 16, 3715–3725.

Brovarets’, O. O., Zhurakivsky, R. O., & Hovorun, D. M. (2015). DPT tautomerisation of the wobble guanine–thymine DNA base mismatch is not mutagenic: QM and QTAIM arguments. *Journal of Biomolecular Structure & Dynamics*, 33, 674–689.

Chečińska, L., & Grabowski, S. J. (2006). F-H–F–F-C hydrogen bonds – The influence of hybridization of carbon atom connected with F-acceptor on their properties. *Chemical Physics*, 327, 202–208.

Cheong, N. R., Nam, S. H., Park, H. S., Ryu, S., Song, J. K., Park, S. M., ... Jouvet, C. (2011). Photofragmentation in selected tautomers of protonated adenine. *Physical Chemistry Chemical Physics*, 13, 291–295.

Colominas, C., Luque, F. J., & Orozco, M. (1996). Tautomerism and protonation of guanine and cytosine. Implications in the formation of hydrogen-bonded complexes. *Journal of the American Chemical Society*, 118, 6811–6821.

Cominboeuf, C., Heine, T., Seifert, G., Schleyer, P. V. R., & Weber, J. (2004). Induced magnetic fields in aromatic [n]annulenes – Interpretation of NICS tensor components. *Physical Chemistry Chemical Physics*, 6, 273–276.

Cossi, M., Rega, N., Scalmani, G., & Barone, V. (2003). Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *Journal of Computational Chemistry*, 24, 669–681.

Danilov, V. I., Anisimov, V. M., Kurita, N., & Hovorun, D. (2005). MP2 and DFT studies of the DNA base rare pairs: The molecular mechanism of the spontaneous substitution mutations conditioned by tautomerism of bases. *Chemical Physics Letters*, 412, 285–293.

Das, G., & Lyngdoh, R. H. D. (2014). Configuration of wobble base pairs having pyrimidines as anticodon wobble bases: Significance for codon degeneracy. *Journal of Biomolecular Structure & Dynamics*, 32, 1500–1520.

Dolney, D. M., Hawkins, G. D., Winget, P., Liótard, D. A., Cramer, C. J., & Truhlar, D. G. (2000). Universal solvation model based on conductor-like screening model. *Journal of Computational Chemistry*, 21, 340–366.

Domagała, M., & Grabowski, S. J. (2005). CH⋯N and CH⋯S–hydrogen bonds-influence of hybridization on their strength. *The Journal of Physical Chemistry A*, 109, 5683–5688.

Ebrahimi, A., Habibi, M., Masoodi, H. R., & Gholipour, A. R. (2009). Relationship between calculated NMR data and intermolecular hydrogen bond properties in X-pyridine⋯HF. *Chemical Physics*, 355, 67–72.

Erikson, L. A. (Ed.). (2001). *Theoretical biochemistry: Processes and properties of biological systems*. Amsterdam: Elsevier Science.

Espinosa, E., Molins, E., & Lecomte, C. (1998). Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. *Chemical Physics Letters*, 283, 170–173.

Fonseca Guerra, C., Bickelhaupt, F. M., Saha, S., & Wang, F. (2006). Adenine tautomers: Relative stabilities, ionization energies, and mismatch with cytosine. *The Journal of Physical Chemistry A*, 110, 4012–4020.

Fonseca Guerra, C., Bickelhaupt, F. M., Snijders, J. G., & Baerends, E. J. (2000). Hydrogen bonding in DNA base pairs: Reconciliation of theory and experiment. *The Journal of the American Chemical Society*, 122, 4117–4128.

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., ... Fox, D. J. (2009). *Gaussian 09 (Revision A.02)*. Wallingford, CT: Gaussian.

Gálvez, O., Gómez, P. C., & Pacios, L. F. (2003). Variation of the intermolecular distance of properties dependent on the electron density in cyclic dimers with two hydrogen bonds. *The Journal of Chemical Physics*, 118, 4878–4895.

Glendening, E. D., Reed, A. E., Carpenter, J. E., & Weinhold, F. (1998). NBO Version 3.1.

Goodman, M. F. (1995). Mutations caught in the act. *Nature*, 378, 237–238.

Grabowski, S. J. (2006). Hydrogen bonding – New insights. Dordrecht: Springer Science & Business Media.

Grabowski, S. J., Sokalski, W. A., & Leszczynski, J. (2005). How short can the H–H intermolecular contact be? New findings that reveal the covalent nature of extremely strong interactions. *The Journal of Physical Chemistry A*, 109, 4331–4341.

Guckian, K. M., Krugh, T. R., & Kool, E. T. (2000). Solution structure of a nonpolar, non-hydrogen-bonded base pair surrogate in DNA. *Journal of the American Chemical Society*, 122, 6841–6847.

Halder, A., Halder, S., Bhattacharyya, D., & Mitra, A. (2014). Feasibility of occurrence of different types of protonated base pairs in RNA: A quantum chemical study. *Physical Chemistry Chemical Physics*, 16, 18383–18396.
von Ragué Schleyer, P., Manoharan, M., Wang, Z. X., Kiran, B., Jiao, H. J., Puchta, R., & van Eikema Hommes, N. J. R. V. E. (2001). Dissected nucleus-independent chemical shift analysis of π-aromaticity and antiaromaticity. Organic Letters, 3, 2465–2468.

Wang, W., Hellinga, H. W., & Beese, L. S. (2011). Structural evidence for the rare tautomer hypothesis of spontaneous mutagenesis. Proceedings of the National Academy of Sciences, 108, 17644–17648.

Watson, J. D., & Crick, F. H. C. (1953). Genetical implications of the structure of deoxyribonucleic acid. Nature, 171, 964–967.

Wiberg, K. B. (2004). Basis set effects on calculated geometries: 6-311++G** vs. aug-cc-pVDZ. Journal of Computational Chemistry, 25, 1342–1346.

Wilcox, J. L., Ahluwalia, A. K., & Bevilacqua, P. C. (2011). Charged nucleobases and their potential for RNA catalysis. Accounts of Chemical Research, 44, 1270–1279.

Wolinski, K., Hinton, J. F., & Pulay, P. (1990). Efficient implementation of the gauge-independent atomic orbital method for NMR chemical shift calculations. Journal of the American Chemical Society, 112, 8251–8260.

Yurenko, Y. P., Zhurakivsky, R. O., Samijlenko, S. P., & Hovorun, D. M. (2011). Intramolecular CH…O hydrogen bonds in the A1 and B1 DNA-like conformers of canonical nucleosides and their Watson–Crick pairs. Quantum chemical and aim analysis. Journal of Biomolecular Structure and Dynamics, 29, 51–65.