Is the DPT Tautomerization of the Long A·G Watson–Crick DNA Base Mispair a Source of the Adenine and Guanine Mutagenic Tautomers? A QM and QTAIM Response to the Biologically Important Question

Ol’ha O. Brovarets,[a,b,c] Roman O. Zhurakivsky,[a] and Dmytro M. Hovorun*[a,b,c]

Introduction

The rare tautomeric hypothesis of spontaneous point mutagenesis postulated by Watson and Crick[1] is one of the most consistent from the physicochemical point of view and biologically productive approach to the nature of spontaneous point mutations in DNA. It has successfully stood the test of time (a lot of experimental[2–7] and theoretical evidences[8–21] have been received in its support) and is now considered a classical provision that underlie this hypothesis and which, figuratively speaking, can be considered as its cornerstones.

1. Mutagenic tautomers of the DNA bases possess large enough lifetime in order not to kinetically limit the processes of their improper pairing with the DNA bases in the canonical tautomeric form in the course of the replication of the latters. Currently, reliable theoretical data are received[22–28] which clearly show that the lifetime of the mutagenic tautomers of both canonical and modified DNA bases that are mutagenic compounds exceeds not only the time used by the replication machinery for the enzymatic incorporation of the nucleotide into the structure of the synthesized DNA double helix (~10⁻³ s[22]), and even the time of the DNA replication in a living cell (~10⁻¹ s[22,29]). It was elucidated that the high stability of the mutagenic tautomers of the nucleotide bases is provided by the lack of the intramolecular hydrogen bonds (H-bonds) in them that are the channels of the effective proton migration.[22–25]

2. Generation of the mutagenic tautomers of the DNA bases[30–35] in a living cell occurs due to the interaction of the DNA bases between themselves as well as due to the other factors of endogenous origin, primarily water...
molecules. A thorough analysis of the literature shows that these issues are now intensively studied and this problem is still far from its final solution. In particular, we showed that the tautomerization as of the A,T[38] and G–C[39] Watson–Crick DNA base pairs, so of the short CT Watson–Crick DNA base pair[40] via the double proton transfer (DPT) along the intermolecular H-bonds (so-called Löwdin's model[41,42]) does not generate mutagenic tautomers of the DNA bases, as the effect of their quantum protection against mutagenic tautomerization takes place in these structures. The same effect hinders the mutagenic tautomerization of the DNA bases by the DNA-binding proteins,[43] that are contained in the replisome.[44,45] A considerable number of researchers consider the endogenous water as the source of the mutagenic tautomerization of the DNA bases in the cell, starting from one water molecule[26,27,36] and ending with microhydration.[10,14,18,27,37] However, the lack in the literature of the estimations of the probability of water molecules penetration into the significantly hydrophobic base-pair binding pocket of a replicative DNA polymerase[46,47] which is a part of the replisome,[44,45] as well as the direct experimental data on their most probable location in this pocket is appeared as an Achilles’ heel of this approach.

This article, which is a logical extension of the previous works on a kindred topic,[8–21,38–40,48–50] is devoted to the finding of the new physicochemical mechanisms of the DNA bases mutagenic tautomerization. An entry point of this article is to find out whether the DPT tautomerization of the long A–G Watson–Crick DNA base mispair, which is considered in the literature as the purine–pyrimidine transversion inducing point mutations in DNA,[51–65] produces A* and G* mutagenic tautomers at its dissociation in the base-pair recognition pocket of the DNA polymerase. This possibility was not considered in the literature before.

In addition, the A–G base mispair is an interesting biological object, as it is quite often found in the spatial structure of the RNA[66–71] and, moreover, it is regarded in the literature as a promising nanostructure for the demands of biomolecular electronics.[72–75] However, in both cases the A–G base mispair is considered as a static structure, that does not tautomerise.

Using a combination of approaches, namely quantum-mechanical (QM) calculations, quantum theory of atoms in molecules (QTAIM) and methodology of the sweeps of the structural, electron-topological, energetical, polar, and natural bond orbital (NBO) properties of the specific intermolecular contacts and base pairs, we investigated the A–G ↔ A*–G* tautomerization via the DPT for the first time.

Our findings have shown that the A–G ↔ A*–G* tautomerization via the DPT according to the Löwdin’s mechanism,[41,42] represents itself a concerted (i.e., this reaction involves no stable intermediates) and asynchronous (i.e., both protons involved in the H-bonds move with a time gap) process. The A*–G* base mispair, formed as a result of the A–G → A*–G* tautomerization via the DPT, was found to be dynamically unstable,[38–40] because the zero-point energy of the corresponding vibrational mode, which frequency becomes imaginary at the TS A–G → A*–G*, greatly exceeds the value of the reverse barrier of the A–G → A*–G* tautomerization. It should be noted that the value of the Gibbs free energy of activation for the reverse reaction of the A–G → A*–G* tautomerization is negative at all levels of QM theory. Consequently, the lifetime of the A*–G* base mispair is extremely short (4.83 × 10–14 s at the MP2/cc-pVQZ/B3LYP/6–311+/+G(d,p) level of QM theory). Therefore any of the six low-frequency intermolecular vibrations[76] can develop during this period of time. This means that during DNA replication the dissociation of the A*–G* base mispair into the A* and G* monomers does not change their tautomeric status, that is, it occurs according to the scheme A*–G* → A–G → A + G. All in all, Löwdin’s mechanism[41,42] does not work in this case.

### Computational methods

All calculations have been carried out with the Gaussian’09 suite of programs.[77] Geometries and harmonic vibrational frequencies of the A–G and A*–G* DNA base mispairs and the TS A–G → A*–G* of their mutagenic tautomerization via the DPT were obtained using density functional theory (DFT)[78] with the B3LYP hybrid functional,[79] which includes Becke’s three-parameter exchange functional (B3)[80] combined with Lee, Yang, and Parr’s (LYP) correlation functional[81] in connection with Pople’s 6–311G(d,p) basis set in vacuum. A scaling factor of 0.9668[82–84] has been used in this work at the B3LYP QM level of theory to correct the harmonic frequencies of all the studied structures. We performed single point energy calculations at the correlated MP2 level of theory[85] with the 6–311++G(d,p) and 6–311+G(3df,2pd) Pople’s[86–88] and cc-pVTZ and cc-pVQZ Dunning’s cc-type[89,90] basis sets for the B3LYP/6–311++G(d,p) geometries to consider electronic correlation effects as accurately as possible. MP2/6–311+G(2df,pd)//B3LYP/6–311++G(d,p), MP2/6–311+G(3df,2pd)// B3LYP/6–311++G(d,p), MP2/cc-pVTZ//B3LYP/6–311++G(d,p), and MP2/cc-pVQZ//B3LYP/6–311++G(d,p) levels of theory have been successfully applied on similar systems recently studied and have been verified to give accurate normal mode frequencies, barrier heights, characteristics of intra- and intermolecular H-bonds and geometries.[38–40,91] Moreover, an excellent agreement between computational and experimental NMR, UV, and IR spectroscopic data[92,93] evidences that the levels of theory applied for the single-point energy calculations (MP2/6–311++G(2df,pd), MP2/6–311++G(3df,2pd), MP2/ cc-pVTZ, and MP2/cc-pVQZ), as well as the method used for the geometry optimization (B3LYP/6–311++G(d,p)) are reliable. The DFT method has been recommended in the literature for describing tautomerization phenomena of the H-bonded nucleobase pairs, because it has shown a good balance between computational cost and accuracy and therefore can be considered as the shortest way to MP2 results.[38,39,94–96] Furthermore, the DFT method has also been proved to be hugely popular for the study of the vibrations of the constituents of the nucleic acids.[85,86]
The correspondence of the stationary points to local minimum or TS\_\text{AG} \rightarrow A^*G^* on the potential energy landscape has been checked by the absence or the presence, respectively, of one and only one imaginary frequency corresponding to the normal mode that identifies the reaction coordinate. TS\_\text{AG} \rightarrow A^*G^* was located by means of Synchronous Transit-guided Quasi-Newton method.\textsuperscript{[97,98]}

Because the stationary points and TS\_\text{AG} \rightarrow A^*G^* were located, the reaction pathway was established by following the IRC in the forward and reverse directions from the TS using the Hess-based predictor-corrector integration algorithm\textsuperscript{[99–101]} with tight convergence criteria. These calculations eventually ensure that the proper reaction pathway, connecting the expected reactants and products on each side of the TS, has been found. We have investigated the evolution of the energetical, geometrical, polar, electron-topological and NBO characteristics of the specific intermolecular contacts and base pairs along the reaction pathway establishing them at the each point of the IRC.

The electronic interaction energies $E_{\text{int}}$ have been computed at the MP2/6-311++G(2df,pd) level of theory for the geometries optimized at the DFT B3LYP/6-311++G(d,p) level of theory as the difference between the total energy of the base pair and the energies of the isolated monomers. In each case, the interaction energy was corrected for the basis set superposition error\textsuperscript{[102,103]} through the counterpoise procedure.\textsuperscript{[104,105]}

The Gibbs free energy $G$ values for all structures were obtained at room temperature ($T = 298.15$ K) in the following way:

$$G = E_{\text{el}} + E_{\text{corr}},$$

where $E_{\text{el}}$—the electronic energy, $E_{\text{corr}}$—the thermal correction.

The lifetime $\tau$ of the A$^*$G$^*$ base mispair was calculated using the formula\textsuperscript{[106]}:

$$\tau = \frac{\ln 10^3}{k_t + k_r}.$$  

To estimate the values of the forward $k_f$ and reverse $k_r$ rate constants for the A-G $\leftrightarrow$ A$^*$G$^*$ tautomerization reaction:

$$k_t = \Gamma \cdot \frac{k_B T}{h} \cdot e^{-\frac{\Delta G_{\text{t}}}{RT}},$$

we applied the standard TS theory\textsuperscript{[106]} in which quantum tunneling effects are accounted by the Wigner's tunneling correction,\textsuperscript{[107]} that is adequate for the DPT reactions\textsuperscript{[38–40,48–50]}:

$$\Gamma = 1 + \frac{1}{24} \left( \frac{h \nu}{k_B T} \right)^2$$

where $k_B$—Boltzmann's constant, $T$—absolute temperature, $h$—Planck's constant, $\Delta G_{\text{t}}$—the Gibbs free energy of activation for the DPT reaction in the forward and reverse directions ($T = 298.15$ K), $R$—universal gas constant, $\nu$—the magnitude of the imaginary frequency at the TS\_AG $\rightarrow$ A$^*$G$^*$.

Bader's QTAIM was applied to analyze electron density.\textsuperscript{[108]}

The topology of the electron density distribution has been examined in detail using program package AIMPAT\textsuperscript{[109]} with all the default options. Wave functions were obtained at the level of theory used for geometry optimization. The presence of a bond critical point (BCP), namely the so-called (3,−1) BCP and a bond path between donor and acceptor, as well as the positive value of the Laplacian at this BCP ($\Delta \rho \geq 0$), were considered as criteria for the DH-bond and H-bond formation.\textsuperscript{[110–114]} Moreover, another five Koch and Popelier's criteria dealing with the changes of atomic properties (positive charge increase $\Delta q$, dipolar polarization decrease $\Delta M$, reduction in atomic volume $\Delta V$, energetic destabilization $\Delta E$, and mutual penetration of donor (H$_{\text{D}}$) and acceptor (H$_{\text{A}}$) hydrogen atoms $\Delta r$) were applied to test the N2H$\_2$--HC2 contact as a DH-bond.

The energies of the conventional intermolecular H-bonds $E_{\text{HB}}$ in the A-G and A$^*$G$^*$ base mispairs were evaluated by the empirical logansen's formula\textsuperscript{[116]}:

$$E_{\text{HB}} = 0.33 \cdot \sqrt{\Delta V - 40},$$

where $\Delta V$—the magnitude of the redshift (relative to the free molecule) of the stretching mode of the H-bonded groups involved in the H-bonding. The partial deuteration of the CH, NH, and NH$_2$ groups was applied to eliminate the effect of vibrational resonances.\textsuperscript{[38,114]}

The energy of the N1H$\_1$--N1 H-bond in the TS\_AG $\rightarrow$ A$^*$G$^*$ was estimated by the Nikolaienko–Bulavin–Hovorun formula\textsuperscript{[117]}:

$$E_{\text{HB}} = -2.03 + 225 \cdot \rho,$$

where $\rho$ is the electron density at the (3,−1) BCP of the H-bond.

The energies of the C2H$\_2$--N2 H-bond in the A$^*$G$^*$ base mispair and TS\_AG $\rightarrow$ A$^*$G$^*$ and of the all intermolecular H-bonds $E_{\text{HB}}$ under the investigation of the sweeps\textsuperscript{[38–40,48,49,50]} of their energies were evaluated by the empirical Espinosa–Molins–Lecomte (EML) formula\textsuperscript{[118,119]} based on the electron density distribution at the (3,−1) BCPs of the H-bonds:

$$E_{\text{HB}} = 0.5 \cdot V(r),$$

where $V(r)$—the value of a local potential energy density at the (3,−1) BCPs.

Moreover, the belonging of the N2H$\_2$--HC2 DH-bond in the A-G base mispair to the true H-bonds was estimated by means of Grunenberg's compliance constants formalism.\textsuperscript{[120–122]}

To study the charge transfer property in the interacting orbitals of the N2H$\_2$--HC2 DH-bond in the A-G base mispair, we have resorted to the NBO analysis,\textsuperscript{[123]} which interprets the electronic wave function in terms of a set of occupied Lewis and a set of unoccupied non-Lewis localized orbitals. The second-order Fock matrix analysis was carried out to evaluate interaction between donor (i) and acceptor (j) bonds.
The result of such interaction is a migration of the electron density from the idealized Lewis structure into an empty non-Lewis orbital $r$. For each donor ($i$) and acceptor ($j$) bonds, the stabilization energy is:

$$E^{(2)} = \Delta E_{ij} = q_i F(i, j)^2 / \varepsilon_j - \varepsilon_i$$

where $q_i$ is the donor orbital occupancy, $\varepsilon_j$ and $\varepsilon_i$ are diagonal elements, and $F(i, j)$ is the off diagonal element of the NBO Fock matrix.

The atomic numbering scheme for the nucleobases is conventional.

### Results and Discussion

#### Structural, electron-topological, and energetical properties of the intermolecular H-bonds in the A-G and A*-G* DNA base mispairs and $T_{A,G} \rightarrow \text{A*-G*}$ of their mutual transformation

Tables (1–4) and Figures 1–10 show the obtained results. Their discussion starts from the structural, electron-topological, and energetical characteristics of the intermolecular H-bonds in the A-G and A*-G* DNA base mispairs and $T_{A,G} \rightarrow \text{A*-G*}$ of their tautomerization via the DPT.

The biologically important A-G DNA base mispair ($\angle C6N1(A)N1C2(G) = -160.3^\circ$), containing canonical tautomers of the A and G DNA bases, and A*-G* DNA base mispair...
It is evident that the D base mispair (N1H bridge, N1H/C1/C1/C1) obtained at the MP2/cc-pVQZ//B3LYP/6–311\[84,125–131\]. The A double-stranded DNA, is the weakest interaction.

The nonplanar TS\(A\rightarrow A^*\) (\(\Delta G_{TS} = 9.63\) kcal mol\(^{-1}\)) and \(\Delta E_{TS} = 11.46\) kcal mol\(^{-1}\) obtained at the MP2/cc-pVQZ//B3LYP/6–311++G(d,p) level of QM theory) is stabilized by the two antiparallel O6H–N6 (5.68 kcal mol\(^{-1}\)) and N1H–N2 (6.51 kcal mol\(^{-1}\)) H-bonds and one N2H–C2 DH-bond (0.68 kcal mol\(^{-1}\)), whereas the tautomerized A\(^*\) base mispair (\(\Delta G = 10.07\) kcal mol\(^{-1}\)) and \(\Delta E = 9.58\) kcal mol\(^{-1}\) obtained at the MP2/cc-pVQZ//B3LYP/6–311++G(d,p) level of QM theory) is stabilized by the two antiparallel canonical O6H–N2 (10.88 kcal mol\(^{-1}\)) and N1H–N1 (7.01 kcal mol\(^{-1}\)) H-bonds and one weak C2H–N2 H-bond (0.42 kcal mol\(^{-1}\)) (Tables 1 and 2).

Table 3. Change of atomic properties of the donor H\(_d\) and acceptor H\(_a\) hydrogen atoms involved in the intermolecular N2H\(_d\)–H\(_a\)C2 DH-bond in the A–G DNA base mispair, that tautomerizes via the DPT, obtained at the B3LYP/6–311++G(d,p) level of theory in vacuum.

<table>
<thead>
<tr>
<th>(\Delta G^{(a)})</th>
<th>(\Delta E^{(a)})</th>
<th>(\Delta \Delta G^{(b)})</th>
<th>(\Delta \Delta E^{(c)})</th>
<th>(\Delta \Delta G^{(d)})</th>
<th>(\Delta \Delta E^{(e)})</th>
</tr>
</thead>
<tbody>
<tr>
<td>kcal mol(^{-1})</td>
<td>kcal mol(^{-1})</td>
<td>kcal mol(^{-1})</td>
<td>kcal mol(^{-1})</td>
<td>kcal mol(^{-1})</td>
<td>kcal mol(^{-1})</td>
</tr>
<tr>
<td>MP2/6–311++G(2df,pd)</td>
<td>9.78</td>
<td>9.30</td>
<td>9.09</td>
<td>10.92</td>
<td>-0.69</td>
</tr>
<tr>
<td>MP2/6–311++G(3df,2pd)</td>
<td>10.02</td>
<td>9.53</td>
<td>9.51</td>
<td>11.34</td>
<td>-0.51</td>
</tr>
<tr>
<td>MP2/cc-pVTZ</td>
<td>9.71</td>
<td>9.22</td>
<td>9.36</td>
<td>11.19</td>
<td>-0.35</td>
</tr>
<tr>
<td>MP2/cc-pCVD</td>
<td>10.07</td>
<td>9.58</td>
<td>9.63</td>
<td>11.46</td>
<td>-0.44</td>
</tr>
</tbody>
</table>

(a) The atomic charge, a.u. (b) The dipolar polarization, a.u. (c) The atomic volume, a.u. (d) The energy of the atom, a.u. (e) The radius of the donor H-d

The spectroscopic data collected in Table 1 confirm geometrical results. The shift in the frequency of the stretching mode of the AH donor group (the difference between the frequency for the AH group in the monomer and in the base pair) is positive (shift to the red) for all conventional H-bonds, whereas—negative for the N2H–C2 DH-bond and the C2H–N2 H-bond.

Physicochemical characteristics of the N2H–C2 DH-bond in the A–G base mispair and of the C2H–N2 H-bond in the A\(^*\)–G\(^*\) base mispair and TS\(A\rightarrow A^*\). The values of the Grunenberg’s compliance constant equal 35.795 Å/mdyn for the N2H–C2 DH-bond in the A–G DNA base mispair, 47.419 Å/mdyn for the C2H–N2 H-bond in the

Table 4. Energetic and kinetic characteristics of the A–G \(\rightarrow A^*\) tautomerization via the DPT in vacuo obtained at the different levels of QM theory for the geometry calculated at the B3LYP/6–311++G(d,p) level of QM theory.

<table>
<thead>
<tr>
<th>Level of QM theory</th>
<th>(\Delta G^{(a)})</th>
<th>(\Delta E^{(b)})</th>
<th>(\Delta \Delta G^{(c)})</th>
<th>(\Delta \Delta E^{(d)})</th>
<th>(\Delta \Delta G^{(e)})</th>
</tr>
</thead>
<tbody>
<tr>
<td>kcal mol(^{-1})</td>
<td>kcal mol(^{-1})</td>
<td>kcal mol(^{-1})</td>
<td>kcal mol(^{-1})</td>
<td>kcal mol(^{-1})</td>
<td>kcal mol(^{-1})</td>
</tr>
<tr>
<td>MP2/6–311++G(2df,pd)</td>
<td>9.78</td>
<td>9.30</td>
<td>9.09</td>
<td>10.92</td>
<td>-0.69</td>
</tr>
<tr>
<td>MP2/6–311++G(3df,2pd)</td>
<td>10.02</td>
<td>9.53</td>
<td>9.51</td>
<td>11.34</td>
<td>-0.51</td>
</tr>
<tr>
<td>MP2/cc-pVTZ</td>
<td>9.71</td>
<td>9.22</td>
<td>9.36</td>
<td>11.19</td>
<td>-0.35</td>
</tr>
<tr>
<td>MP2/cc-pCVD</td>
<td>10.07</td>
<td>9.58</td>
<td>9.63</td>
<td>11.46</td>
<td>-0.44</td>
</tr>
</tbody>
</table>

(a) The relative Gibbs free energy of the A\(^*\)–G\(^*\) base pair (\(\Delta G_{GA} = 0.00\) kcal mol\(^{-1}\)). [b] The relative electronic energy of the A\(^*\)–G\(^*\) base pair (\(\Delta E_{GA} = 0.00\) kcal mol\(^{-1}\)). [c] The Gibbs free energy of activation for the forward reaction of the A–G\(\rightarrow A^*\)–G\(^*\) DPT tautomerization (\(T = 298.15\) K). kcal mol\(^{-1}\). [d] The activation electronic energy for the reverse reaction of the A–G\(\rightarrow A^*\)–G\(^*\) DPT tautomerization (\(T = 298.15\) K). kcal mol\(^{-1}\). [e] The activation electronic energy for the reverse reaction of the A–G\(\rightarrow A^*\)–G\(^*\) DPT tautomerization. [f] The frequency of the vibrational mode in the A–G base pair, which becomes imaginary in the TS\(A\rightarrow A^*\) tautomerization obtained at the level of geometry optimization, cm\(^{-1}\). [g] The zero-point vibrational energy associated with this normal mode.

(a) The lifetime of the A\(^*\)–G\(^*\) base pair, s. (b) The time necessary to reach 99.9% of the equilibrium concentration of the A–G reactant and the A\(^*\)–G\(^*\) product of the A–G\(\rightarrow A^*\)–G\(^*\) tautomerization reaction via the DPT, s.
A*G* DNA base mispair and 37.515 Å/mdyn for the C2H—N2 H-bond in the TS
A—G — A*G* tautomerization, indicating that they are attractive closed-shell interactions [110–114,132].

We have compared the NBO charges on the hydrogen atoms of the N2H and C2H groups in order to distinguish between the H-bond donating (donor) and accepting (acceptor) groups [49]. It was found that the hydrogen atom of the N2H group in the G base bears a greater positive charge (0.388 e) and thus can be considered as the H-bonding donor (Hd), than the hydrogen atom of the C2H group in the A base (0.168 e) in the A—G base mispair, serving as the H-bonding acceptor (Ha).

The analysis of the electron density distribution, atomic properties and their changes of the donor Hd and acceptor Hs hydrogen atoms involved in the N2Hd—HsC2 DH-bond in the AG base mispair (Table 3) allows us to establish that it completely satisfies all eight “two-molecule” Koch and Popelier’s criteria for the identification of the H-bonds [114,115]. So, we established the presence of a (3,−1) BCP, a bond path between the donor Hd and acceptor Hs hydrogen atoms, a positive value of the Laplacian of the electron density and also that the charge of the donor Hd hydrogen atom increases, its dipolar polarization and atomic volume decrease, the energy of the donor Hd hydrogen atom increases and the mutual...
penetration is the positive value (i.e., the atomic radius of the bonded atom is shorter than of the nonbonded atom) for both the donor $H_d$ and acceptor $H_a$ hydrogen atoms upon the $N_2H/C_1/C_1/C_1H/C_2$ bond formation.

The $N_2H/C_1/C_1/C_1H/C_2$ and $C_2H/C_1/C_1/C_1N_2$ specific interactions in the $A/C_1G$, $A^{*}/C_1G^{*}$ stationary points, respectively, do not meet the geometric requirements for the H-bonding, as the distances between the donor and acceptor groups and atoms exceed the sum of corresponding Bondi’s van der Waals radii \cite{133}: $d_{N_2H/C_1/C_1/C_1H/C_2} (2.469 \text{ Å}) > 2r_{HvdW} (2.40 \text{ Å})$ and $d_{C_2H/C_1/C_1/C_1N_2} (3.229$ and $3.085 \text{ Å}) > r_{HvdW} + r_{NvdW} (2.75 \text{ Å})$. However, a van der Waals cutoff is not the physical limit of the long-range electrostatic H-bond interaction and can act beyond this distance\cite{111,114}. Moreover, the hydrogen bond radius for the CH group has been recently revised in the literature and established to be $1.10 \pm 0.20 \text{ Å}$ and it appears that a CH group could have a radius larger than $1.2 \text{ Å}$ when involved in the H-bonding\cite{134}. Nevertheless, the value of the $N_2H/C_1/C_1/C_1H/C_2$ bond angle in the $A/C_1G$ base mispair, that is equal to $124.6^\circ$, and of the $C_2H/C_1/C_1/C_1N_2$ bond angle, that is equal to $120.8^\circ$ in the $A^{*}/C_1G^{*}$ base mispair and $176.6^\circ$ in the $T_{S_{A/C_1G^{*}}}A^{*}/C_1G^{*}$ suggests that these interactions contribute toward the structural stabilization (Tables 1 and 2).

Electron-topological analysis shows that there is specific interaction between the donor $H_d$ and acceptor $H_a$ hydrogen atoms in the $N_2H/C_1/C_1/C_1H/C_2$ bond (Tables 1 and 2). NBO analysis predicts transfer of charge from $\sigma(N_2-H)$ bonding orbital to $\sigma^*(C_2-H)$ antibonding orbital. The second-order perturbation energy $E^{\(2\)}$, characterizing the strength of this interaction, is equal to $0.14 \text{ kcal/mol}^{-1}$ in the $A/C_1G$ base mispair.

For the first time, we have established the spectroscopic manifestations of the $N_2H/C_1/C_1/C_1H/C_2$ bond in the $A/C_1G$ base mispair and of the $C_2H/C_1/C_1/C_1N_2$ bond in the $A^{*}/C_1G^{*}$ base mispair. Thus, in particular, we have shown that the $\nu(N_2H)$ frequency of the stretching vibration of the $N_2H$ donor group of the $N_2H/C_1/C_1/C_1H/C_2$ bond shifts to the red by $0.5 \text{ cm}^{-1}$ and its IR intensity increases in $\sim 7$ times, whereas the $\nu(C_2H)$ frequency of the stretching vibration of the $C_2H$ donor group of $C_2H/C_1/C_1/C_1N_2$ bond shifts to the red by $0.5 \text{ cm}^{-1}$ and its IR intensity increases in $\sim 7$ times.

Figure 3. Profile of the dipole moment $\mu$ along the IRC of the $A/C_1G \rightarrow A^{*}/C_1G^{*}$ tautomerization via the DPT obtained at the B3LYP/6–311++G(d,p) level of theory in vacuo.

Electron-topological analysis shows that there is specific interaction between the donor $H_d$ and acceptor $H_a$ hydrogen atoms in the $N_2H/C_1/C_1/C_1H/C_2$ bond (Tables 1 and 2). NBO analysis predicts transfer of charge from $\sigma(N_2-H)$ bonding orbital to $\sigma^*(C_2-H)$ antibonding orbital. The second-order perturbation energy $E^{\(2\)}$, characterizing the strength of this interaction, is equal to $0.14 \text{ kcal/mol}^{-1}$ in the $A/C_1G$ base mispair.

For the first time, we have established the spectroscopic manifestations of the $N_2H/C_1/C_1/C_1H/C_2$ bond in the $A/C_1G$ base mispair and of the $C_2H/C_1/C_1/C_1N_2$ bond in the $A^{*}/C_1G^{*}$ base mispair. Thus, in particular, we have shown that the $\nu(N_2H)$ frequency of the stretching vibration of the $N_2H$ donor group of the $N_2H/C_1/C_1/C_1H/C_2$ bond shifts to the red by $0.5 \text{ cm}^{-1}$ and its IR intensity increases in $\sim 7$ times, whereas the $\nu(C_2H)$ frequency of the stretching vibration of the $C_2H$ donor group of $C_2H/C_1/C_1/C_1N_2$ bond shifts to the red by $0.5 \text{ cm}^{-1}$ and its IR intensity increases in $\sim 7$ times.

Figure 4. Profiles of: (a) the electron density $\rho$; (b) the Laplacian of the electron density $\Delta \rho$; (c) the ellipticity $\varepsilon$ at the $(3,-1)$ BCPs of the covalent and H-bonds, and (d) the energy of the H-bond $E_{HB}$ estimated by the EML formula\cite{118,119}, along the IRC of the $A/G \rightarrow A^{*}/G^{*}$ tautomerization via the DPT obtained at the B3LYP/6–311++G(d,p) level of theory in vacuo. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]
Figure 5. Profiles of: (a) the distance $d_{A\rightarrow B}$ between the electronegative A and B atoms; (b) the distance $d_{AH\rightarrow HB}$ between the hydrogen and electronegative A or B atoms; and (c) the angle $\angle AH\rightarrow B$ of the AH–B H-bonds along the IRC of the A–G $\rightarrow$ A*–G* tautomerization via the DPT obtained at the B3LYP/6–311++G(d,p) level of theory in vacuo.

Figure 6. Profiles of: (a) the electron density $\rho$; (b) the Laplacian of the electron density $\Delta\rho$; (c) the ellipticity $\iota$, and (d) the energy $E_{HB}$ estimated by the EML formula\(^{[118,119]}\) at the (3,–1) BCPs of the N2H–HC2 DH-bond (highlighted in green) and the C2H–N2 H-bond (highlighted in red) along the IRC of the A–G $\rightarrow$ A*–G* tautomerization via the DPT obtained at the B3LYP/6–311++G(d,p) level of theory in vacuo.
the C2H···N2 H-bond shifts to the red by 3.2 cm$^{-1}$ and its IR intensity decreases in $1.3$ times under the formation of the bond (Table 1). At the same time, the $\nu$(C2H) frequency of the out-of-plane bending vibration increases by 20.4 cm$^{-1}$, while its IR intensity decreases in $1.1$ times under the formation of the C2H···N2 H-bond in the A$^*$-G$^*$ base mispair. These spectral changes are characteristic for the weak H-bonds [38,110,114]. Taking into account the aforementioned arguments, the N2H···HC2 DH-bond and the C2H···N2 H-bond can be considered as true H-bonds [38,39,49,114],[135–137].

Thermodynamical and dynamical stability of the A-G and A$^*$-G$^*$ DNA base mispairs

In the result of our painstaking work it was found that the A-G ($\Delta E_{\text{int}} = -17.54$ kcal mol$^{-1}$ and $\Delta G_{\text{int}} = -3.57$ kcal mol$^{-1}$) and A$^*$-G$^*$ ($\Delta E_{\text{int}} = -23.62$ kcal mol$^{-1}$ and $\Delta G_{\text{int}} = -11.02$ kcal mol$^{-1}$) base mispairs are thermodynamically stable structures, as their Gibbs free energies of interaction, obtained at the MP2/6–311++G(2df,pd)//B3LYP/6–311++G(d,p) level of theory in vacuum, are less than zero, differing greatly from each other. Notably, the interaction energies increase for the tautomerized A$^*$-G$^*$ base mispair comparably with the starting A-G base mispair. Furthermore, the electronic and Gibbs free energies of interaction of the A-G and A$^*$-G$^*$ base mispairs are less than those for the G-C Watson–Crick base pair [39]. It is important to note that the intermolecular H-bonds and the DH-bond make a major contribution into the stabilization of the A-G ($[E_{\text{N6H-C6}} + E_{\text{N1H-N1}} + E_{\text{N2H-CH2}}]/[\Delta E_{\text{int}}] = 73.4\%$) and A$^*$-G$^*$ ($[E_{\text{O6H-N6}} + E_{\text{N1H-N1}} + E_{\text{C2H-N2}}]/[\Delta E_{\text{int}}] = 77.5\%$) DNA base mispairs. This energy relationship cannot be considered as the unique physicochemical characteristic exceptionally of these base mispairs, as it was also observed for the other H-bonded base pairs [38–40,43,114].

It was found that the A$^*$-G$^*$ base mispair is dynamically unstable structure [38–40,43,94] as the zero-point energy of the
corresponding vibrational mode, which frequency becomes imaginary at the TS of A → A* via the DPT (Table 3). It should be noted that the value of the Gibbs free energy of activation for the reverse reaction of the A → A* tautomerization is negative at all levels of QM theory. Moreover, this statement is objective and does not depend on the chosen quantum-chemical level of theory (Table 3).

As a result the lifetime of the A* base mispair is extremely short and is equal to 4.83 × 10^{-14} s at the MP2/cc-pVQZ//B3LYP/6–311 + + G(d,p) level of theory (Table 4). Therefore, any of the six low-frequency intermolecular vibrations (22.0, 33.0, 52.5, 76.5, 102.7, and 109.8 cm^{-1}) can develop during this time span, as their periods are noticeably greater than this time interval. This observation additionally indicates that the A* base mispair is not dynamically stable structure: this means that the A → A* via the DPT [38,39,41,42] does not work in the case of the A-G base mispair.

The time τ_{99.9\%} necessary to reach 99.9\% of the equilibrium concentration of the starting A-G and the final A* base mispair is equal to 4.45 × 10^{-13} s obtained at the MP2/cc-pVQZ//B3LYP/6–311 + + G(d,p) level of theory (Table 4).

Structural and electron-topological architecture of the established nine key points along the IRC of the A → A* tautomerization via the DPT

With a view to investigate the energetical, structural, polar, electron-topological, and NBO reorganizations of the A-G base mispair and intermolecular specific contacts stabilizing it along the IRC, we have performed the calculations of the electronic energy, the first derivative of the electronic energy with respect to the IRC, the dipole moment of the base pair, the distances and the angle of the intermolecular H-bonds, the electron density, the Laplacian of the electron density, ellipticity, and the energy at the (3, 2) BCPs of the intrapair H-bonds, the NBO charges of the hydrogen atoms involved in the intermolecular interactions, the dihedral angles, the glycosidic angles, and the distance between the glycosidic hydrogens at each step along the IRC of the A-G tautomerization. In such a way, we obtained the scanning (so-called sweeps) of these characteristics presented in Figures 2–10.

We revealed altogether the nine key points [38–40,48–50] for the A-G → A* tautomerization, based on the changes of
the electron density and geometry of the intermolecular H-bonds in the A-G base mispair along the IRC of the A-G → A*-G* tautomerization via the DPT (Figs. 1, 2, 4, and 5).

The key point 1, IRC = −12.06 Bohr. The starting structure along the IRC pathway is the A-G DNA base mispair, stabilized by the N6-H–O6 and N1-H–N1 H-bonds and the N2H-..HC2 DH-bond (Tables 1 and 2 and Figs. 1, 2, 4–7).

The key point 2, IRC = −3.72 Bohr. The structure of the base pair, for which the H–N1 chemical bond of the A base is significantly weakened and the N1-H bond essentially becomes the covalent bond, for which $\Delta p$ equals zero ($\rho_{N1-H} = 0.113$ a.u.; $\Delta \rho_{N1} = 0.000$ a.u.; $d_{N1-H} = 1.415$ Å; $d_{N1-N1} = 2.630$ Å; $\angle N1-H-N1 = 177.2^\circ$) (Figs. 1, 2, 4, and 5). The maximum value of the energy of the N1-H bond is reached at this key point (Fig. 4d). Interestingly, that one out of the two extrema of the first derivative of the electron energy with respect to the IRC—$d\epsilon/d\text{IRC}$ is reached exactly at this key point (Fig. 4d). Moreover, precisely at the key point 2 the A and G bases, acting in this case as the reagents of the DPT reaction, lose their chemical individuality.

The key point 3, IRC = −3.51 Bohr. This structure is characterized by the equivalent loosened N1-H and H-N1 covalent bonds. Dependencies of the geometrical and electron-topological characteristics at the BCPs of these equivalent chemical bonds intersect exactly at this key point, forming a $\chi^2$-like graph for the loosened N1-H–N1 covalent bond ($\rho_{N1-H} = 0.149$ a.u.; $\Delta \rho_{N1} = 0.096$ a.u.; $d_{N1-H} = 1.307$ Å; $d_{N1-N1} = 2.632$ Å; $\angle N1-H-N1 = 177.7^\circ$) (Table 2 and Figs. 1, 2, 4, and 5).

The key point 4, IRC = −3.25 Bohr. The structure of the base pair, for which the H–N1 covalent bond of the N1-H–N1 covalent bridge becomes the H–N1 H-bond (Figs. 1 and 2). A characteristic feature of this structure is a zero value of $\Delta \rho$ at the (3,—1) BCP of the H–N1 H-bond ($\rho_{H-N1} = 0.104$ a.u.; $\Delta \rho_{H-N1} = 0.000$ a.u.; $d_{H-N1} = 1.453$ Å; $d_{N1-N1} = 2.643$ Å; $\angle N1-H-N1 = 178.3^\circ$) (Table 2 and Figs. 1, 2, 4, and 5). The maximum value of the energy of the H–N1 H-bond is attained at this key point (Fig. 4d).

The key point 5, IRC = −0.11 Bohr. At this structure situated quite close to the TS$_{A\rightarrow A^*}$, the N6-H chemical bond of the A base is significantly weakened and the H–O6 H-bond actually becomes the H–O6 covalent bond, for which $\Delta p$ equals zero ($\rho_{H-O6} = 0.137$ a.u.; $\Delta \rho_{H-O6} = 0.000$ a.u.; $d_{H-O6} = 1.287$ Å; $d_{N6-O6} = 2.485$ Å; $\angle N6-H-O6 = 174.8^\circ$) (Table 2 and Figs. 1, 2, 4, and 5). The maximum value of the energy of the H–O6 H-bond is reached exactly at this key point (Fig. 4d).

The key points 6 and 7, IRC = 0.00 Bohr. The TS$_{A\rightarrow A^*}$ of the tautomerization via the DPT is stabilized by the N6-H–O6 covalent bridge, N1-H–N1 and C2-H–N2 canonical H-bonds and represents itself the structure, characterized by the equivalent loosened N6-H and H–O6 covalent bonds (Tables 1 and 2 and Figs. 1, 2, 4, and 5). Dependencies of the geometrical and electron-topological characteristics at the (3,—1) BCPs of these equivalent chemical bonds intersect exactly at this key point, forming a $\chi^2$-like graph for the loosened N6-H–O6 covalent bond ($\rho_{N6-H} = \rho_{H-O6} = 0.167$ a.u.; $\Delta \rho_{N6-H} = \Delta \rho_{H-O6}$).

Separation of the reaction pathway into the reactant, TS and product regions

These nine key points have been used in this study to divide the reaction pathway into the reactant, TS and product regions (138–141) of the A-G → A*-G* tautomerization via the DPT (Figs. 1 and 2). This underlying separation can be done quite naturally and unambiguously by taking the reaction force minimum and the reaction force maximum as the boundaries for these regions (138–141). It was established, basing on the analysis of the sweeps of the physicochemical characteristics of the intermolecular interactions along the IRC, that the reactant and product regions, where nucleotide bases do not lose their chemical individuality, are located between the key points 1–2 (−12.06 ± 3.72 Bohr) and 8–9 (0.26 ± 2.88 Bohr), respectively. The TS region, where eventually the DPT reaction occurs and the A and G bases lose their chemical individuality, is quite narrow and located between the key points 2 and 8 (−3.72 ± 0.26 Bohr), where the Laplacian of the electron density $\Delta \rho$ at the (3,—1) BCPs of the N1-H and N6-H H-bonds, respectively, vanishes. Interestingly, as it was noted earlier, the extremum of the first derivative of the dependency of the electronic energy on the IRC—$d\epsilon/d\text{IRC}$ is reached exactly at the key points 2 and 8 (Figs. 1 and 2b).

We established that the electronic energy necessary to bring the donor and acceptor atoms of the intermolecular interactions as close as possible to each other to acquire such mutual deformation and orientation, that eventually lead to the DPT reaction, that is the electronic energy difference between the key points 2 and 1, is equal to 8.08 kcal mol$^{-1}$, representing 64.1% of the TS$_{A\rightarrow A^*}$ electronic energy relatively to the A-G base mispair. The comparably small amount of the electronic energy (1.30 kcal mol$^{-1}$) releases during the relaxation of the key point 8 to the key point 9, that is obtained as the electronic energy difference between the key points 8 and 9,
representing 10.3% of the TS_{A\rightarrow A^*G^*} electronic energy relatively to the A-G base mispair. These observed data evidence that the greatest amount of energy is spent for the rebuilding and reorganization of the A and G bases within the A-G base mispair actually before the initiation of the chemical reaction that occurs at the TS region.

Variations of the polar, electron-topological, energetical, and geometrical physicochemical properties of the conventional H-bonds in the A-G base mispair along the IRC of the A-G \rightarrow A^*G^* tautomerization through the DPT

The absolute value of the dipole moment $\mu$ of the studied base pairs significantly alters within the range of values 4.97 ± 8.18 D along the IRC of the A-G \rightarrow A^*G^* tautomerization (Fig. 3). The $\Omega$-like profile of the dipole moment has a maximum (8.18 D) in the vicinity of the TS_{A\rightarrow A^*G^*} (IRC = −1.15 Bohr) (Fig. 3).

The electron-topological (the electron density $\rho$, the Laplacian of the electron density $\nabla^2\rho$, the ellipticity $\varepsilon$ and the energy $E_{\text{HB}}$, estimated by the EML formula [118,119]) at the (3,−1) BCPs and geometrical (the distance $d_{A\rightarrow B}$ between the electronegative A and B atoms, the distance $d_{\text{AHHB}}$ between the hydrogen and electronegative A or B atoms and the angle $\angle$AH−B of the AH−B H-bond) properties of the canonical intermolecular H-bonds are presented in Figures 4 and 5, respectively.

The dependencies of the electron density $\rho$, the Laplacian of the electron density $\nabla^2\rho$, the ellipticity $\varepsilon$ and the energy $E_{\text{HB}}$, estimated by the EML formula [118,119], at the (3,−1) BCPs of the H-bonds are nonmonotonic. The graphs of the electron density $\rho$ and the Laplacian of the electron density $\nabla^2\rho$ at the (3,−1) BCPs of the N1−H and H−N1 bonds intersect exactly at the key point 3, while of the N6−H and H−O6 bonds—at the key points 6 and 7, that are one and the same (Figs. 4a and 4b). Moreover, the shapes of these profiles are similar to each other, indicating a strong correlation between them [111,112].

In contrast, the graph of the ellipticity $\varepsilon$ at the (3,−1) BCP of the N1−H bond crosses with the graphs for the N1−H and N6−H bonds at the key point 2 (Fig. 4c). This behavior of the ellipticity $\varepsilon$ indicates that the intermolecular bonds are very sensitive to the dynamical behavior of the base pair and their energies are modulated by the low-frequency intermolecular vibrations of the base pair that tautomerizes [38–40,48–50]. The values of the $\rho$ (0.030 ± 0.321 a.u.), $\nabla^2\rho$ (−1.870 ± 0.163 a.u.), and $\varepsilon$ (0.014 ± 0.067) parameters lie within a wide range and are in line with the results, presented in our recent works [38–40,48–50].

It is important to note that the upper N6H−O6 and the middle N1H−N1 H-bonds in the A-G base mispair exist within the range from the key point 1 to 2 and from the key point 1 to 5 inclusively, reaching their maxima at the key points 2 and 5, respectively. The upper O6H−N6 and the middle N1H−N1 H-bonds in the A^*G* base mispair exist within the range from the key point 4 to 9 and from the key point 8 to 9 inclusively, reaching their maxima at the key points 4 and 8, respectively (Fig. 4d).

Notably, the $d_{A\rightarrow B}$ and $d_{\text{AHHB}}$ distances and the angles $\angle$AH−B of the N1H−N1, N6H−O6, and O6H−N6 H-bonds nonmonotonically vary along the IRC of the A-G \rightarrow A^*G^* tautomerization via the DPT (Figs. 5). The minima of the $d_{N6−O6}$ distance is observed exactly at the key point 2, in which actually electronic and geometric rebuildings of the base pair are activated in order to ensure the course of the reaction, and of the $d_{\text{N6−O6}}$ distance—exactly at the key points 6 and 7 (Fig. 5a). Thus, the $d_{\text{N6−O6}}$ and $d_{\text{N1−N1}}$ distances intersect exactly at the key point 3, while the $d_{\text{N6H}}$ and $d_{\text{O6H}}$ distances—exactly at the key points 6 and 7, in which starts the relaxation to the product—the A^*G* base mispair (Figs. 5b). At this also both the $\angle$N6H−O6H−N6H (171.2 ± 175.1°) and $\angle$N1H−N1H (171.8 ± 178.4°) angles significantly vary along the IRC (Fig. 5c).

Moreover, by analyzing the profiles of the $\angle$AH−B angle, the $d_{A\rightarrow B}$ and $d_{\text{AHHB}}$ distances of the AH−B H-bond, the $r(H−H)$ distance between the glycosidic protons, the $\chi_1$ (\angle$N9H(A)(A)$ and $\chi_2$ (\angle$N9H(G)(G)$) glycosidic angles, the $\angle$N1C2N2H2(H) and $\angle$C6N1(A)N1C2(G) dihedral angles, we arrived at the conclusion that the A-G base mispair “breathes” throughout the tautomerization process (Figs. 5, 7, 9, and 10). The compression of the starting A-G base mispair, occurring mainly at the TS region, arises due to the contraction of the distances between the N1 and N1 nitrogen atoms (by 0.333 Å) and the N6 nitrogen and O6 oxygen atoms (by 0.400 Å) (Fig. 5a).

The dependency of the $r(H−H)$ distance between the H9 glycosidic protons varies in the range 11.948 ± 12.375 Å and acquires its minimal value exactly at the key point 2 (Fig. 9a). Interestingly, the $r(H−H)$ distance for the A-G base mispair (12.375 Å) exceeds the same value for the A^*G* base mispair (12.188 Å) (Fig. 9a). The oscillations of the values are observed for the $\chi_1$ and $\chi_2$ glycosidic angles (43.9 ± 46.4° and 42.9 ± 45.9°, respectively) (Fig. 9b).

The $\angle$N1C2N2H2(H) dihedral angle lying in the range 33.9 ± 42.6° reflects the rotation of the amino group relative to the plane of the G base, while the $\angle$C6N1(A)N1C2(G) dihedral angle lying in the range −171.4 ± 169.2° displays the twists of the A and G bases relative to each other (Fig. 10). The profiles of these dihedral angles nonmonotonically change along the IRC of the A-G \rightarrow A^*G* tautomerization via the DPT (Fig. 10).
Changes of the physicochemical characteristics of the N2H–HC2 DH-bond and the C2H–N2 H-bond along the IRC of the A–G ↔ A*–G* tautomerization via the DPT

Extremely interesting situation is observed for the third specific interaction in the A–G base mispair, as in the course of the A–G ↔ A*–G* reaction the hydrogen atom of the C2H group serving as an acceptor in the N2H–HC2 DH-bond becomes the donor of the C2H–N2 H-bond at the IRC = −10.07 Bohr. It was found that the A–G ↔ A*–G* tautomerization is assisted by the third N2H–HC2 DH-bond that converts into the C2H–N2 H-bond without discontinuity and bifurcation at the IRC = −10.07 Bohr (Figs. 1, 6 and 7). In our recent work,[49] it was obtained the similar result according the C2H–HC2 DH-bond that assists the A–A* ↔ A–A* tautomerization via the DPT, for which the donor and acceptor hydrogen atoms interchange in the course of the reaction.

The profiles of the $\rho$, $\Delta \rho$, $\epsilon$, $E_{HB}$, $d_{AB}$, $d_{AH\rightarrow HB}$, and $\angle AH\rightarrow B$ parameters of the N2H–HC2 DH-bond and the C2H–N2 H-bond are presented in Figures 6 and 7. We have previously shown that the N2H–HC2 DH-bond and the C2H–N2 H-bond are true H-bonds, as they meet all known to date criteria of H-bonding.[38,39,49,114,135–137]

Profiles of the electron density $\rho$, the Laplacian of the electron density $\Delta \rho$, the ellipticity $\epsilon$, and the energy $E_{HB}$, estimated by the EML formula,[118,119] at the (3,−1) BCP of the N2H–HC2 DH-bond transforming into the C2H–N2 H-bond along the IRC of the A–G ↔ A*–G* tautomerization are bell-shaped with a slightly asymmetric top (Fig. 6). Moreover, the shapes of the $E_{HB}$, $\rho$ and $\Delta \rho$ profiles are similar to one another, indicating a strong correlation between these values, for which the correlation coefficient $\gamma$ consists 0.992/0.991 between $E_{HB}$ and $\rho$ values, 0.969/0.908 between $E_{HB}$ and $\Delta \rho$ values, 0.962/0.900 between $\rho$ and $\Delta \rho$ values obtained at the B3LYP/6–31G(d,p)/MP2/6–31G(d,p) levels of theory, respectively, as evidenced by the recently published data for the CH–N H-bonds revealed in the biologically important base pairs.[138] Notably, the values of the $\rho$, $\Delta \rho$, $\epsilon$, and $E_{HB}$ values for the N2H–HC2 DH-bond insignificantly vary within the range: 0.0037 ± 0.0038 a.u., 0.01369 ± 0.01384 a.u., 0.372 ± 0.503 and 0.648 ± 0.656 kcal·mol$^{-1}$, respectively, reaching their minima exactly at the A–G base mispair (Fig. 6). It was drawn attention to the fact that the $\rho$, $\Delta \rho$, and $E_{HB}$ values for the C2H–N2 H-bond, varying within a wide range 0.0029 ± 0.0057 a.u., 0.0095 ± 0.0171 a.u. and 0.45 ± 0.91 kcal·mol$^{-1}$, respectively, reach their minima exactly at the A*–G* base mispair (Tables 1 and 2 and Figs. 6a, 6b and 6d). This nonmonotone behavior of the ellipticity $\epsilon$ at the (3,−1) BCP of the C2H–N2 H-bond (0.194 ± 0.617) shows the dynamical instability of the C2H–N2 H-bond and the modulation of its energy by the low-frequency intermolecular vibrations of the base mispair[38–40,48–50] (Fig. 6c), moreover, the modulation amplifies at the weakening of the contact. These data coincide with the range of values obtained in our recent work.[110–114]

The $d_{A\rightarrow B}$ and $d_{AH\rightarrow HB}$ distances nonmonotonically increase, whereas the $\angle N2H\rightarrow H$ angle of the N2H–HC2 DH-bond and the $\angle C2H\rightarrow N2$ angle of the C2H–N2 H-bond, that attains in the course of the reaction several local minima in the vicinity of the 2–4 and 5–8 key points, nonmonotonically decrease along the IRC of the A–G ↔ A*–G* tautomerization (Fig. 7). There are considerable jumps for the geometric characteristics of the third interaction in a base mispair at the N2H–HC2 DH-bond → C2H–N2 H-bond transition along the IRC of the A–G ↔ A*–G* tautomerization (Fig. 7). The minimal values of the geometric characteristics are observed for the N2H–HC2 DH-bond, indicating a close proximity of donor and acceptor groups, while the maximal values—for the C2H–N2 H-bond.

Sweeps of the NBO charges of the hydrogen atoms involved in the third interaction evidence that the charge of the donor H$_4$ hydrogen atom invoked in the N2H$_4$–H$_2$C2 DH-bond bears the greater charge than the acceptor H$_4$ hydrogen atom and remains so during the formation of the C2H–N2 H-bond along the IRC of the A–G ↔ A*–G* tautomerization via the DPT (Fig. 8). However, these graphs do not intersect with each other. It can be asserted that strictly speaking the third interaction can be considered as a partially charge-assisted H-bond in the TS region.

Asynchronous concerted mechanism of the A–G ↔ A*–G* tautomerization via the DPT

Analyzing the sweeps of the geometrical and electron-topological characteristics of the intermolecular H-bonds, namely the observed spacings of the $\gamma$-like crossings on the $d_{NH\rightarrow HOB}$, $d_{NH\rightarrow HNN}$, $\rho$, $\Delta \rho$ and $\epsilon$ profiles, we came to the conclusion that the A–G ↔ A*–G* tautomerization proceeds through the asynchronous concerted mechanism (Tables 1 and 2 and Figs. 1, 2, 4–7).

The A–G base mispair converts into the A*–G* base mispair by the sequential migration of the N1 nitrogen atom of the G base along the N1H–N1 H-bond to the N1 nitrogen atom of the A base, forming in such a way the A*–G* zwiterionic base pair. Then, the second mobile proton, localized at the N6 oxygen atoms, the proton reaches the O6 covalent bridge, N1H–O6 complex, N1H–N1 and C2H–N2 H-bonds, the proton reaches the O6 oxygen atom, that eventually leads to the formation of the tautomerized A*–G* base mispair involving mutagenic tautomers of the A and G DNA bases (Tables 1 and 2 and Fig. 1).

Concluding Remarks

In conclusion, we found out that the A–G ↔ A*–G* tautomerization via the DPT occurs via the asynchronous concerted mechanism.

It was established that the A–G base mispair is stabilized by the N6H–O6 and N1H–N1 H-bonds and N2H–HC2 DH-bond, whereas the A*–G* base mispair—by the O6H–N6 and N1H–N1 H-bonds and weak C2H–N2 H-bond and the $T_{AG}$ ↔ $A*\leftrightarrow G* by the N6–H–O6 covalent bridge, N1H–N1 and C2H–N2 H-bonds. Interestingly, the A–G ↔ A*–G* tautomerization via the DPT is accompanied by the third N2H–HC2 DH-
bond, that smoothly converts into the third weak C2H–N2 H-bond at the IRC = −10.07 Bohr.

Based on the sweeps of the energies of the intermolecular interactions, it was established that the N6H—O6 H-bond is anticooperative to the two others N1H—N1 H-bond and N2H—C2 H2 D2-bond in the A-G base mispair, mutually weakening each other, whereas the three O6H—N6, N1H—N1 and C2H—N2 H-bonds are cooperative in the A*-G* base mispair, mutually reinforcing each other.

The nine key points, two of which are accidentally degenerate, were detected and completely investigated along the IRC of the A⋅G −→ A*-G* tautomerization via the DPT, among them three key points are stationary structures—the A-G reactant (the key point 1), the TS_A⋅G −→ A*-G* (the key point 6 that coincides with the key point 7) and the A*-G* product (the key point 9). The others six key points are defined by the structural and electron-topological rearrangements of the long A-G Watson–Crick base pair along the IRC of the A⋅G −→ A*-G* tautomerization via the DPT.

The A*-G* base mispair was found to be dynamically unstable structure, because the zero-point energy of the corresponding vibrational mode, which frequency becomes imaginary at the TS_A⋅G −→ A*-G*, greatly exceeds the value of the reverse barrier of the A⋅G −→ A*-G* tautomerization. The lifetime of the A*-G* base mispair is extremely short (4.83 × 10−14 s), that is determined by the negative value of the Gibbs free energy for the A*-G* −→ A-G transition at all levels of QM theory. Moreover, any of the six low-frequency intermolecular vibrations can develop during this period of time. These observations indicate that the classical Löwdin’s mechanism of the origin of the spontaneous point mutations arising during DNA replication does not work in this case.

Acknowledgments

This work was partially supported by the Science and Technology Center in Ukraine (STCU) within the project no. 5728 for 2012–2014 years and by the State Fund for Fundamental Research (SFFR) of Ukraine within the Ukrainian-Japanese project no. F 52.2/001 for 2013–2014 years. O.O.B. was supported by the Grant of the President of Ukraine to support scientific research of young scientists for 2012 year from the Ministry of Education and Science, Youth and Sports of Ukraine. Authors thank Bogolyubov Institute for Theoretical Physics of the National Academy of Sciences of Ukraine for providing calculation resources and software. This work was performed using computational facilities of joint computer cluster of SSI “Institute for Single Crystals” of the National Academy of Sciences of Ukraine and Institute for Scintillation Materials of the National Academy of Sciences of Ukraine incorporated into Ukrainian National Grid. The authors sincerely thank Dr. Fernando R. Clemente (Gaussian, Inc.) for the technical support of the work.

Keywords: amino and imino tautomers of the adenine · keto and enol tautomers of the guanine · adenine-guanine (A-G) DNA base mispair · sweeps of the energetic · electron-topological · geometric · polar · NBO parameters along the IRC · the double proton transfer · cooperativity of the H-bonds · NH–HC dihydrogen bond · CH–N H-bond · B3LYP and MP2 levels of QM theory · QTAIM analysis

References


Received: 21 September 2013
Revised: 12 November 2013
Accepted: 30 November 2013
Published online on 00 Month 2013