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

Ol’ha O. Brovarets*a,b and Dmytro M. Hovorun*a,b*

*aDepartment of Molecular and Quantum Biophysics, Institute of Molecular Biology and Genetics, National Academy of Sciences of Ukraine, 150 Akademika Zabolotnoho Str., Kyiv 03680, Ukraine; bDepartment of Molecular Biotechnology and Bioinformatics, Institute of High Technologies, Taras Shevchenko National University of Kyiv, 2-h Akademika Hlushkova Ave., Kyiv 03022, Ukraine

Communicated by Ramaswamy H. Sarma

(Received 18 April 2014; accepted 13 May 2014)

This study provides the first accurate investigation of the tautomerization of the biologically important guanine* thymine (G*·T) DNA base mispair with Watson–Crick geometry, involving the enol mutagenic tautomer of the G and the keto tautomer of the T, into the G·T* mispair (ΔG = .99 kcal mol−1, population = 15.8% obtained at the MP2 level of quantum-mechanical theory in the continuum with ε = 4), formed by the keto tautomer of the G and the enol mutagenic tautomer of the T base, using DFT and MP2 methods in vacuum and in the weakly polar medium (ε = 4), characteristic for the hydrophobic interfaces of specific protein–nucleic acid interactions. We were first able to show that the G*·T ↔ G·T* tautomerization occurs through the asynchronous concerted double proton transfer along two antiparallel O6H···O4 and N1···HN3 H-bonds and is assisted by the third N2H···O2 H-bond, that exists along the entire reaction pathway. The obtained results indicate that the G·T* base mispair is stable from the thermodynamic point of view complex, while it is dynamically unstable structure in vacuum and dynamically stable structure in the continuum with ε = 4 with lifetime of 6.4·10−12 s, that, on the one side, makes it possible to develop all six low-frequency intermolecular vibrations, but, on the other side, it is by three orders less than the time (several ns) required for the replication machinery to forcibly dissociate a base pair into the monomers during DNA replication. One of the more significant findings to emerge from this study is that the short-lived G·T* base mispair, which electronic interaction energy between the bases (−23.76 kcal mol−1) exceeds the analogical value for the G·C Watson–Crick nucleobase pair (−20.38 kcal mol−1), “escapes from the hands” of the DNA replication machinery by fast transforming into the G*·T mismatch playing an indirect role of its supplier during the DNA replication. So, exactly the G*·T mismatch was established to play the crucial role in the spontaneous point mutagenesis.

Keywords: spontaneous point mutations; transition mismatches; DPT tautomerization; mutagenic tautomer; DNA replication; hydrogen bond cooperativity; B3LYP and MP2; QTAIM

Introduction

The accuracy of the DNA replication in the living cell is a very important biological phenomenon that determines the genetic stability of all living things on the Earth (Kornberg & Baker, 1992). It is well known that the point (gene) mutations play the dual role: evolutionary, on the one side, for maintaining genetic information over many generations and, on the other side, they can be considered as a key cause of different illnesses, in particular cancer in humans (Drake, 1991; Friedberg et al., 2006; Vonborstel, 1994). Transitions, as events that occur most frequently and, therefore, are often eliminated from the genome by the reparations systems of the cell (Lee, Popodi, Tang, & Foster, 2012; Rogozin and Pavlov, 2003), occupy a prominent place among the point mutations, the underlying cause of which is an accidental formation of the irregular base pairs with geometry close to those of the Watson–Crick base pairs in the base pair recognition pocket of the DNA polymerase (Goodman, 1997; Hwang & Taylor, 2005; Kool, 2002; Minnick, Liu, Grindley, Kunkel, & Joyce, 2002), e.g. of the G·T mismatch in the recognition pocket of the DNA polymerase λ, that is well poised for catalysis (Bebenek, Pedersen, & Kunkel, 2011). Unfortunately, their physicochemical nature remains understandable only in the most general features (Topal & Fresco, 1976). This is obviously due to the lack of the well-grounded from the physicochemical point of view and internally non-contradictable approach or ideology to the successful solution of these problems. In fairness it must be said that only the so-called Watson and Crick’s tautomeric hypothesis formulated shortly after the establishment of the spatial structure of the DNA can pretend on the role of this ideology (Watson & Crick, 1953). Nowadays, the intensive accumulation of both experimental and theoretical data continues within

*Corresponding author. Email: dhovorun@imbg.org.ua

© 2014 Taylor & Francis
the framework of this hypothesis (Brovarets’ & Hovorun, 2013c; Danilov, Anisimov, Kurita, & Hovorun, 2005; Danilov et al., 2009; Wang, Hellenga, & Beece, 2011).

Thus, in our previous work (Brovarets’ & Hovorun, 2013c) by applying DFT and MP2 quantum chemical methods and methodology of the scans of the electron-topological, energetic, geometrical, and polar characteristics along the intrinsic reaction coordinate (IRC) of the tautomerization, it was shown that the lifetime \((5.76 \times 10^{-10})\) s of the thermodynamically and dynamically stable A*-C base mispair \((\Delta G = 2.72 \text{ kcal mol}^{-1})\), population \(= 1.02 \%\) obtained at the MP2 level of quantum-mechanical (QM) theory in the continuum with \(\varepsilon = 4\) is by order less than the time (several ns) required for the replication machinery to forcibly dissociate a base pair into the monomers during DNA replication. These data suggest that the enzymatically competent A-C base mispair possessing Watson–Crick geometry, that was experimentally fixed in the recognition pocket of the DNA polymerase (Wang et al., 2011), is in the A-C* tautomeric form (here and below the mutagenic tautomers (Brovarets’ & Hovorun, 2010a, 2011a, 2013d; Furmanchuk et al., 2011; Kosenkov et al., 2009; Platonov, Samijlenko, Sudakov, Kondratyuk, & Hovorun, 2005; Samijlenko, Yurenko, Stepanyugin, & Hovorun, 2012) of the canonical DNA bases are marked with asterisks), associated by the three cooperative N6H···N4, N1···HN3, and C2H···O2 H-bonds. Notably, the C2H···O2 specific contact was established to be the real H-bond (Brovarets’ & Hovorun, 2013c), which acts as a third point of support, thereby ensuring the similarity of the structurally dynamical properties of this mismatch to the classical Watson–Crick DNA base pairs.

This paper is a logical continuation of the previous one (Brovarets’ & Hovorun, 2013c). It is aimed to give an unambiguous answer to the biologically important question, issued in its title, and is focused on the thorough elucidation of the physico-chemical mechanism of the G*-T→G*T* tautomerization through the double proton transfer (DPT) to understand the role of each of these base mispairs in the production of the spontaneous point replication errors. In fact, the incorrect G*-T and G-T* purine–pyrimidine base mispairs with Watson–Crick architecture and mechanism of their tautomerization through the DPT are the subjects of the quantum chemical study both in vacuum and in the continuum with \(\varepsilon = 4\), typical for the hydrophobic interiors of proteins (Garcia-Moreno et al., 1997) and the hydrophobic interfaces of specific protein–nucleic acid interactions (Bayley, 1951; Brovarets’, Yurenko, Dubey, & Hovorun, 2012; Dewar & Storch, 1985; Mertz & Krishtalik, 2000; Petruska, Sowers, & Goodman, 1986).

In this paper, for the first time we have succeeded in the theoretical survey of the G*-T→G*T* DPT tautomerization. It was established that this reaction proceeds in the ground electronic state through a single transition state following the concerted mechanism, in which the proton-transfer process occurs asynchronously without forming any stable intermediates. The G-T* DNA base mispair with the population equal to 15.8% was found to be the dynamically stable structure with lifetime of \(6.39 \times 10^{-12}\) s (in the continuum with \(\varepsilon = 4\)), that is by three orders less than the time required for the replication machinery to forcibly dissociate a base pair into the monomers during DNA replication. This work also reports the evolution of the energetic, electron-topological, geometric, polar, and natural bond orbital (NBO) parameters (Brovarets’ & Hovorun, 2013e, 2014b; Brovarets’, Zhurakivsky, & Hovorun, 2013b, 2013c; Brovarets’, Kolomiets’, & Hovorun, 2012), describing the course of the tautomerization via the DPT along the IRC. It was established based on the energetic scans, that the O6H···O4/H···O6, N3H···N1/N1H···N3 and N2H···O2 H-bonds in the G*-T/G/T* DNA base mispairs, respectively, are significantly cooperative.

**Computational methods**

All calculations have been carried out with the Gaussian 09 programs suite (Frisch et al., 2010). Geometries and harmonic vibrational frequencies of the G*-T/G/T* DNA base mispairs and the transition state (TS\(_{G^*\cdot T\rightarrow G\cdot T^*}\)) of their tautomerization via the DPT were obtained using Density Functional Theory (DFT) with B3LYP hybrid functional (Tirado-Rives & Jorgensen, 2008), which includes Becke’s three-parameter exchange functional (B3) (Parr & Yang, 1989) combined with Lee, Yang and Parr’s (LYP) correlation functional (Lee, Yang, & Parr, 1988) and second-order Møller–Plesset perturbation theory (MP2) (Frisch, Head-Gordon, & Pople, 1990) in connection with Pople’s 6-311++G(d,p) basis set in vacuum and in the continuum with a low dielectric constant of \(\varepsilon = 4\), which is characteristic for the hydrophobic interiors of proteins (Garcia-Moreno et al., 1997) and hydrophobic interfaces of specific protein–nucleic acid interactions (Bayley, 1951; Brovarets’, Yurenko, et al., 2012; Dewar & Storch, 1985; Mertz & Krishtalik, 2000; Petruska et al., 1986), by using the Conductor-Like Polarizable Continuum Model (Barone & Cossi, 1998; Cossi, Rega, Scalmani, & Barone, 2003). In this article the “DFT(\(\varepsilon = 1\))/DFT(\(\varepsilon = 4\))/MP2(\(\varepsilon = 1\))/MP2(\(\varepsilon = 4\))” abbreviation denotes the representation of the results obtained at: DFT level of theory in vacuum/DFT level of theory in the continuum with \(\varepsilon = 4\)/MP2 level of theory in vacuum/MP2 level of theory in the continuum with \(\varepsilon = 4\), respectively. Scale factors of .9580 and .9531 (Brovarets’ & Hovorun, 2013e, 2014b) (for the existing approaches to the definition of this parameter see papers (Palafox, 2000; Palafox, Iza, & Gil, 2002; Palafox & Rastogi, 2002)) have been used in the present work at the B3LYP and MP2 levels of theory, respectively, to correct
the harmonic frequencies of all the studied base pairs. DFT and MP2 levels of QM theory have established itself well exactly in the studies of proton transfer reactions (Arabi & Matta, 2011; Brovarets’ & Hovorun, 2013a, 2013b, 2013e, 2014b; Brovarets’ et al., 2013b, 2013c; Brovarets’, Zhurakivsky, & Hovorun, 2013d). Furthermore, it was found that the DFT and MP2 methods allow to obtain adequate barrier heights, characteristics of intra- and intermolecular H-bonds, and geometries (Brovarets’ & Hovorun, 2010b, 2010c, 2010d, 2010e, 2011b; Brovarets’, Zhurakivsky, & Hovorun, 2010; Samijlenko et al., 2012; Matta, 2010; Lozynski, Rusinska-Roszak, & Mack, 1998; Palafax, 2014).

Geometry optimizations were followed by the single-point energy calculations in vacuum and in the continuum with a low dielectric constant (ε = 4) at the DFT and MP2 levels of theory using a wide variety of basis sets, in particular, Pople’s (Frisch, Pople, & Binkley, 1984; Hariharan & Pople, 1973; Krishnan, Binkley, Seeger, & Pople, 1980) and Dunning’s basis sets (Dunning, 1989; Kendall, Dunning, & Harrison, 1992), augmented with polarization and/or diffuse functions. They are: 6-311++G(X), where X = G(d,p), G(2df,pd), G(3df,2pd), and cc-pVTZ (altogether 4 basis sets).

The correspondence of the stationary points to minimum on the potential energy landscape or TS, located by means of Synchronous Transit-guided Quasi-Newton method (Peng, Ayala, Schlegel, & Frisch, 1996; Peng & Schlegel, 1993), 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.

Once the stationary points were located, the reaction pathway was established by performing the IRC calculations in the forward and reverse directions from the TS$^{G^*\rightarrow T\rightarrow G\rightarrow T^*}$ to reach the reactant (G*·T base mispair) and product (G·T* base mispair) using the Hessian-based predictor-corrector integration algorithm (Hratchian & Schlegel, 2004, 2005a, 2005b) with tight convergence criteria, that eventually ensure that the proper reaction pathway has been found.

By using the IRC data we have established the profiles of the physico-chemical properties of the DNA base mispairs and intermolecular contacts stabilizing them, in particular 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,−1) BCPs of the intrapair H-bonds, the NBO charges of the hydrogen atoms involved in the intermolecular H-bonds, the glycosidic angles, the distance between the glycosidic hydrogens and dihedral angle along the reaction pathway of the G*·T$\rightarrow$G·T* DPT tautomerization establishing them at the each point along the IRC in vacuum and in the continuum with ε = 4.

The electronic interaction energies in vacuum $E_{\text{int}}$ have been computed at the MP2/6-311++G(2df,pd) level of theory for the geometries optimized at the DFT and MP2 levels of theory as the difference between the total energy of the base mispair and the energies of the isolated monomers. In each case the interaction energy was corrected for the basis set superposition error (BSSE) (Boys & Bernardi, 1970; Gutowski, Van Lenthe, Verbeek, Van Duijneveldt, & Chalasinski, 1986) through the counterpoise procedure (Sordo, Chin, & Sordo, 1988; Sordo, 2001). The interaction energies in solution $E_{\text{sol}}$ have been calculated as the sum of gas phase interaction energies and a solvation correction term (Brovarets’ & Hovorun, 2014b):

$$E_{\text{sol}} = E_{\text{int}} + (E_{\text{cosmo}} - E_{\text{gas phase}}),$$  \hspace{1cm} (1)

where $E_{\text{cosmo}}$ and $E_{\text{gas phase}}$ are the BSSE uncorrected interaction energies in solution and gas phase, respectively, evaluated for the geometries optimized in solution.

The Gibbs free energy $G$ values for all structures was obtained in the following way:

$$G = E_{\text{el}} + E_{\text{corr}},$$  \hspace{1cm} (2)

where $E_{\text{el}}$ – the electronic energy, $E_{\text{corr}}$ – thermal correction. We applied the standard TS theory (Atkins, 1998) to estimate the activation barriers of the tautomerization reaction.

The time $\tau_{99.9\%}$ necessary to reach 99.9% of the equilibrium concentration of the G*·T and G·T* DNA base mispairs in the system of reversible first-order forward ($k_f$) and reverse ($k_r$) reactions can be calculated by the formula (Atkins, 1998):

$$\tau_{99.9\%} = \frac{\ln 10^3}{k_f + k_r},$$  \hspace{1cm} (3)

and the lifetime $\tau$ of the G·T* complex is obtained as $1/k_r$.

The equilibrium constants were calculated using the standard equation (Atkins, 1998):

$$K = e^\frac{-\Delta G}{RT},$$  \hspace{1cm} (4)

where $\Delta G$ – the relative Gibbs free energy of the product – the G·T* DNA base mispair, $T$ – the absolute temperature and $R$ – the universal gas constant.

To estimate the values of the rate constants $k_f$ and $k_r$:

$$k_f, r = \Gamma \frac{k_B T}{h} e^{-\frac{\Delta G_{f, r}}{RT}}$$  \hspace{1cm} (5)

we applied the standard TS theory (Atkins, 1998), in which quantum tunneling effect are accounted by the Wigner’s tunneling correction (Wigner, 1932) that is adequate for the DPT reactions (Arabi & Matta, 2011;
Brovarets’, Zhurakivsky, & Hovorun, 2014a, 2014b, 2014c, 2014d; Brovarets’ & Hovorun, 2014a, 2014c, 2014d):

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

where \( k_B \) – Boltzmann’s constant, \( h \) – Planck’s constant, \( \Delta G_{fr} \) – Gibbs free energy of activation for the DPT reaction in the forward (\( f \)) and reverse (\( r \)) directions, \( \nu \) – the magnitude of the imaginary frequency associated with the vibrational mode at the TS\( \Gamma \) that connects reactants and products.

Bader’s quantum theory “Atoms in molecules” (QTAIM) was applied to analyze the electron density distribution (Bader, 1990), using program package AIM-All (Keith, 2010) with all default options. Wave functions were obtained at the level of theory used for geometry optimization. The presence of a bond critical point (BCP) (Bader, 1990), namely the so-called (3,–1) BCP, and a bond path between the H-bond donor and acceptor, as well as the positive value of the Laplacian at this BCP (\( \Delta \rho > 0 \)), were considered as criteria for the H-bond formation.

The energies of the intermolecular H-bonds under the investigation of the scans of their energies were evaluated by the empirical Espinosa–Molins–Lecomte (EML) formula (Espinosa, Molins, & Lecomte, 1998; Mata, Alkorta, Espinosa, & Molins, 2011), which was first successfully applied for the estimation of the individual energetic contributions of the separate H-bonds in the two Watson–Crick DNA base pairs (Matta, Castillo, & Boyd, 2006):

\[ E_{HB} = 5 \cdot V(r), \]  

where \( V(r) \) – the value of a local potential energy density at the (3,–1) BCPs.

The energies of the O6H···O4 and N2H···O2 H-bonds in the TS\( G^*\) were estimated by the Nikolaienko–Bulavin–Hovorun formulas (Nikolaienko, Bulavin, & Hovorun, 2012):

\[ E_{O6H\cdot\cdot\cdotO4} = -3.09 + 239 \cdot \rho, \]  
\[ E_{N2H\cdot\cdot\cdotO2} = -2.03 + 225 \cdot \rho, \]

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

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

\[ E_{HB} = 0.33 \cdot \sqrt{\Delta \nu} - 40, \]  

where \( \Delta \nu \) – 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 was applied to minimize the effect of vibrational resonances (Brovarets’ & Hovorun, 2013c, 2014b, 2013e).

The period of the intermolecular vibrations \( T \) was calculated as:

\[ T = \frac{1}{\nu \cdot c}, \]  

where \( \nu \) – the frequency of vibrations, \( c \) – the speed of the light in vacuum.

The atomic numbering scheme for the purine and pyrimidine bases is conventional (Saenger, 1984).

Obtained results and their discussion

Geometric and energetic characteristics of the G*·T and G·T* DNA base mispairs

The data given in the Tables 1 and 2 indicate that the G*·T base mispair possesses pairing scheme with the formation of the O6H···O4, N3H···N1, and N2H···O2 H-bonds, and the G·T* base mispair (\( \Delta G = 99 \) and \( \Delta E = 1.31 \) kcal mol\(^{-1} \) obtained at the MP2/cc-pVTZ//MP2/6–311++G(d,p) level of QM theory in the continuum with \( \varepsilon = 4 \) – of the O4H···O6, N1H···N3, and N2H···O2 H-bonds. The TS\( G^*\) (\( \Delta G_{TS} = 3.45 \) and \( \Delta E_{TS} = 6.30 \) kcal mol\(^{-1} \) obtained at the MP2/cc-pVTZ/B3LYP/6–311++G(d,p) level of QM theory in the continuum with \( \varepsilon = 4 \); \( \nu_i = 1123.9i/1178.7i/708.2i/800.2i \) cm\(^{-1} \)) is stabilized by the N1-H-N3 covalent bridge and the two parallel O6H···O4 and N2H···O2 H-bonds (Tables 1 and 2, Figure 1).

The current study found that the electron-topological (the electron density \( \rho \) is in the acceptable range of values; the Laplacian of the electron density \( \Delta \rho \) at the (3,–1) BCP of the bond is positive), geometric (the \( d_{H\cdot\cdot\cdotN1/N3/N2} \) and \( d_{H\cdot\cdot\cdotO6/O4/O2} \) distances are less than the sum of the corresponding Bondi’s (Bondi, 1964) van der Waals radii; the AH H-bond donating groups elongate upon the formation of the conventional AH···B H-bonds; the angles of the H-bonding are obtuse), and spectroscopic (the frequency of the stretching vibrational mode \( \nu(AH) \) of the AH donor group is shifted toward the lower frequencies or toward the red end of the spectrum under the formation of all H-bonds) characteristics confirm the existence of the canonical H-bonds in the investigated complexes (Brovarets’ & Hovorun, 2013c, 2013e, 2014b; Brovarets’, Zhurakivsky, & Hovorun, 2013a; Ponomareva, Yurenko, Zhurakivsky, van Mourik, & Hovorun, 2012; Yurenko, Zhurakivsky, Samijlenko, & Hovorun, 2011) (for more details refer to Table 1).

Geometric characteristics of the G*·T and G·T* DNA mismatches, that are important for their recognition by the DNA replication machinery

The biologically important G*·T and G·T* DNA base mispairs, formed by the flexible (Govorun, Danchuk, & Mishchuk, 1992; Hovorun, Gorb, & Leszcynski, 1999;
Table 1. Electron-topological, structural, vibrational, energetic, and polar characteristics of the intermolecular H-bonds in the G*·T/G*·T* DNA base mispairs and TSG*·T ↔ G·T* obtained at the B3LYP/6-311++G(d,p) (DFT) and MP2/6-311++G(d,p) (MP2) levels of theory in vacuo ($\varepsilon = 1$) and in the continuum with a low dielectric constant ($\varepsilon = 4$).

| Base mispairs/TS | AH···B H-bond | $\rho^a$ | $\Delta \rho^b$ | 100$\%^c$ | $d_{\ell-H}^d$ | $\angle \text{AH} \cdots \text{B}^f$ | $\Delta l_{\ell-H}^g$ | $\Delta \nu^h$ | $E_{\text{HB}}^i$ | $\mu^j$ |
|------------------|---------------|---------|-------------|----------|-------------|------------------|----------------|---------------|----------------|-------|
|                  |               | DFT MP2 | DFT MP2     | DFT MP2  | DFT MP2     | DFT MP2          | DFT MP2      | DFT MP2       | DFT MP2       | DFT MP2 |
| $\varepsilon = 1$|                |         |             |          |             |                  |               |               |               |        |
| G*·T             | O6H···O4      | .040    | .039        | .126     | .127        | 2.723            | 2.715         | 1.742         | 1.736         | 171.3  |
|                  | N3H···N1      | .040    | .042        | .091     | .096        | 6.17             | 5.73          | 1.843         | 1.814         | 175.9  |
|                  | N2H···O2      | .022    | .020        | .078     | .074        | 5.82             | 5.27          | 3.022         | 3.047         | 177.3  |
|                  | O4H···O6      | .062    | .060        | .147     | .150        | 2.05             | 2.02          | 2.582         | 2.582         | 173.1  |
|                  | N1H···N3      | .040    | .043        | .100     | .104        | 6.30             | 5.99          | 2.859         | 2.839         | 172.0  |
|                  | N2H···O2      | .023    | .023        | .082     | .081        | 5.65             | 5.44          | 2.990         | 2.989         | 174.0  |
|                  | TS$_{G*\rightarrow T}$ | .076    | .083        | .155     | .150        | 2.20             | 1.97          | 2.520         | 2.495         | 174.1  |
|                  | O6H···O4      | .037    | .033        | .122     | .111        | 6.22             | 5.15          | 2.806         | 2.854         | 176.3  |
| $\varepsilon = 4$|                |         |             |          |             |                  |               |               |               |        |
| G*·T             | O6H···O4      | .044    | .046        | .131     | .137        | 2.81             | 2.50          | 2.690         | 2.681         | 170.0  |
|                  | N3H···N1      | .038    | .039        | .091     | .096        | 6.20             | 5.84          | 2.905         | 2.873         | 175.5  |
|                  | N2H···O2      | .040    | .043        | .098     | .103        | 6.24             | 5.89          | 2.866         | 2.856         | 172.9  |
|                  | O4H···O6      | .055    | .051        | .143     | .141        | 2.59             | 2.63          | 2.617         | 2.615         | 171.5  |
|                  | N1H···N3      | .040    | .043        | .098     | .103        | 6.24             | 5.89          | 2.866         | 2.856         | 172.9  |
|                  | N2H···O2      | .025    | .025        | .087     | .088        | 5.91             | 5.81          | 2.966         | 2.971         | 174.0  |
|                  | TS$_{G*\rightarrow T}$ | .083    | .088        | .151     | .146        | 2.33             | 2.12          | 2.498         | 2.482         | 172.9  |
|                  | O6H···O4      | .036    | .032        | .120     | .107        | 5.50             | 5.51          | 2.816         | 2.873         | 176.5  |

$^a$The electron density at the (3,−1) BCP (a.u.).
$^b$The Laplacian of the electron density at the (3,−1) BCP (a.u.).
$^c$The ellipticity at the (3,−1) BCP.
$^d$The distance between the A and B atoms in the AH···B H-bond (Å).
$^e$The distance between the H and B atoms in the AH···B H-bond (Å).
$^f$The H-bond angle (degree).
$^g$Elongation of the donating group AH upon H-bonding (Å).
$^h$The red shift (positive value) or the blue shift (negative value) of the stretching vibrational mode of the H-bonded AH group (cm$^{-1}$).
$^i$The H-bond energy estimated by Iogansen's (Iogansen, 1999) or Nikolenko-Bulavin-Hovorun (Nikolenko et al., 2012) (marked with an asterisk) formulas (kcal mol$^{-1}$).
$^j$The dipole moment of the complex (D).
Table 2. Electron-topological, structural, and polar characteristics of the intermolecular H-bonds in the nine key points (KPs) revealed along the IRC of the G*T+→G*T+ tautomerization via the DPT obtained at the B3LYP/6-311++G(d,p) (DFT) and MP2/6-311++G(d,p) (MP2) levels of theory in vacuo ($\varepsilon=1$) and in the continuum with a low dielectric constant ($\varepsilon=4$).

| KPs, Bohr | AH···B H-bond | $\Delta\rho$ | $100\nu$ | $\Delta\sigma_{A\cdot B}$ | $\Delta\sigma_{H\cdot A}$ | $\rho_{A\cdot H}$ | $\rho_{H\cdot A}$ | $\rho\Delta\sigma_{1}$ | $\rho\Delta\sigma_{2}$ |
|----------|----------------|-------------|-----------|--------------------------|-----------------------|----------------|----------------|----------------|----------------|
| KT 1     | OOH···O4       | 0.40        | 0.44      | 0.39                      | 0.46                   | 0.12           | 0.13           | 0.12           | 0.13           |
|          | N3H···N1       | 0.40        | 0.38      | 0.32                      | 0.39                   | 0.09           | 0.09           | 0.09           | 0.09           |
| KT 2     | OOH···O4       | 0.72        | 0.79      | 0.80                      | 0.86                   | 0.13           | 0.16           | 0.16           | 0.16           |
|          | N3H···N1       | 1.10        | 1.09      | 1.10                      | 1.10                   | 0.00           | 0.00           | 0.00           | 0.00           |
| KT 3     | OOH···O4       | 0.36        | 0.36      | 0.32                      | 0.31                   | 0.12           | 0.12           | 0.12           | 0.12           |
|          | N3H···N1       | 0.75        | 0.82      | 0.80                      | 0.84                   | 0.15           | 0.15           | 0.15           | 0.15           |
| KT 4     | OOH···O4       | 0.37        | 0.36      | 0.33                      | 0.32                   | 0.15           | 0.15           | 0.15           | 0.15           |
|          | N3H···N1       | 0.37        | 0.36      | 0.33                      | 0.32                   | 0.15           | 0.15           | 0.15           | 0.15           |
| KT 5     | OOH···O4       | 0.79        | 0.85      | 0.84                      | 0.89                   | 0.14           | 0.14           | 0.14           | 0.14           |
|          | N3H···N1       | 1.10        | 1.11      | 1.10                      | 1.10                   | 0.00           | 0.00           | 0.00           | 0.00           |
| KT 6     | OOH···O4       | 0.38        | 0.37      | 0.33                      | 0.32                   | 0.12           | 0.12           | 0.12           | 0.12           |
|          | N3H···N1       | 0.65        | 0.67      | 0.78                      | 0.82                   | 0.07           | 0.07           | 0.07           | 0.07           |
| KT 7     | OOH···O4       | 0.34        | 0.35      | 0.31                      | 0.31                   | 0.11           | 0.11           | 0.11           | 0.11           |
|          | N3H···N1       | 0.63        | 0.65      | 0.73                      | 0.78                   | 0.10           | 0.10           | 0.10           | 0.10           |
| KT 8     | OOH···O4       | 0.33        | 0.34      | 0.30                      | 0.30                   | 0.10           | 0.10           | 0.10           | 0.10           |
|          | N3H···N1       | 0.61        | 0.63      | 0.71                      | 0.76                   | 0.10           | 0.10           | 0.10           | 0.10           |
| KT 9     | OOH···O4       | 0.62        | 0.55      | 0.60                      | 0.51                   | 0.14           | 0.14           | 0.14           | 0.14           |
|          | N3H···N1       | 0.40        | 0.40      | 0.43                      | 0.43                   | 0.10           | 0.10           | 0.10           | 0.10           |

Note: For footnote definitions see Table 1.
Nikolaenko, Bulavin, & Hovorun, 2011) G and T DNA bases in the canonical and mutagenic tautomeric forms, and the transition state of their mutual tautomeric transformation $\text{TS}_{G^*\cdot T^*\rightarrow G\cdot T^*}$ are planar structures at the DFT level of theory ($C_s$ symmetry; $C6\text{N1(G)}\text{N3C4(T)}=4.5/0.4/1^\circ$ in vacuum and $5.6/0.8/1^\circ$ in the continuum with $\varepsilon=4$ for the $G^*\cdot T^*/G\cdot T^*/G\cdot C^*$, respectively), whereas they become propeller-like structures at the MP2 level of theory ($C_2\text{v}$ symmetry; $<C6\text{N1} \text{G} \text{C4} \text{T}=27.9/22.4/29.4^\circ$ in vacuum and $27.1/24.2/30.1^\circ$ in the continuum with $\varepsilon=4$, respectively) (Figures 1, 15 and S14).

Another geometric requirement for the successful incorporation of the incorrect base pair into the DNA double helix is the mimicry of its glycosidic sizes (Goodman, 1997). So, the comparison of the glycosidic parameters of the $G^*\cdot T^*$ and $G\cdot T^*$ DNA base mispairs, namely the distance between the H9 and H1 glycosidic protons R(H9-H1) and the $\angle N9\text{HH}$ and $\angle N1\text{HH}$ glycosidic angles, with those for the canonical Watson–Crick base pairs indicates that they are close to the G-C DNA base pair (Table 4), in such a way adopting a structure that closely mimics that of a natural base pairs and utilizing an appropriate H-bonding potential. This data suggest that the $G^*\cdot T$ and $G\cdot T^*$ DNA base mispairs from the geometrical point of view would be successfully incorporated into the DNA double helix by the DNA polymerase during DNA biosynthesis, that agree well with the previously expressed assumptions (Danilov et al., 2005; Topal & Fresco, 1976).

**Thermodynamic and dynamic stability of the $G^*\cdot T$ and $G\cdot T^*$ DNA base mispairs**

From the data presented in Table 3 we can come to the conclusion that the $G^*\cdot T$ and $G\cdot T^*$ DNA base mispairs are thermodynamically (Brovarets et al., 2014a, 2014d) stable structures, since the values of their Gibbs free energies of the interaction between the bases within the base pair are less than zero. The intermolecular H-bonds in the $G^*\cdot T$ and $G\cdot T^*$ DNA base mispairs make a significant contribution into the electronic energy of the interaction between the bases, that corroborates well with the data obtained in previous observations (Brovarets et al., 2014a, 2014b, 2014c, 2014d; Brovarets & Hovorun, 2014a) (Table 3).

It should be noted that the electronic and Gibbs free energies of the interaction between the G and T* bases in the G-T* base mispair exceed the same values for the canonical G-C Watson–Crick base pair (Brovarets & Hovorun, 2013e) that can be considered as a unique case among those that were previously considered (Brovarets & Hovorun, 2013a, 2013b, 2013c, 2014a, 2014b; Brovarets et al., 2013b, 2013c, 2013d, 2014a, 2014b, 2014c, 2014d). Based on the obtained results the following order of stability can be derived both in vacuum and in the continuum with $\varepsilon=4$: A·T < G·T < G·C < G·T* (Table 3). In view of this, the G-T* DNA base mispair is not suitable candidate to cause spontaneous point mutations in DNA, since it acts as an impediment to work of the DNA polymerase, designed for a range of values of the canonical A·T and G-C DNA base pairs.

It is interestingly to note, that the G-T* base mispair is dynamically unstable structure in vacuum (the zero-point energy $E_{ZPPE}=1511.8 \text{ cm}^{-1}$ or 4.32 kcal mol$^{-1}$ of the corresponding vibrational mode, which frequency becomes imaginary at the $TS_{G^*\cdot T\rightarrow G\cdot T^*}$, exceeds the value of the reverse barrier $\Delta\Delta E=1334.7 \text{ cm}^{-1}$ or 3.82 kcal mol$^{-1}$), but it acquires dynamical stability in the continuum with $\varepsilon=4$ (the zero-point energy $E_{ZPPE}=1485.9 \text{ cm}^{-1}$ or 4.25 kcal mol$^{-1}$ of the corresponding vibrational mode is less than the value of the reverse barrier $\Delta\Delta E=1744.5 \text{ cm}^{-1}$ or 4.99 kcal mol$^{-1}$) (Table 5).

However, the lifetime of the formed G-T* base mispair involving T* mutagenic tautomizer is by three orders less than the time required for the replication machinery to forcibly dissociate a base pair into the monomers during DNA replication (several ns (Brovarets’, Kolomietz’, et al., 2012)) both in vacuum ($\tau=2.41\cdot10^{-12}$ s) and in the continuum with $\varepsilon=4$ ($\tau=6.39\cdot10^{-12}$ s) (Table 5). These results highlight that the short-lived G-T* base mispair plays the role of the supplier of the G·T base
mispair, that is the G·T mismatch with Watson–Crick architecture decays into the G* and T monomers at the dissociation by the DNA polymerase.

This, in addition to all of the above, clearly shows that exactly the G*·T base mispair, associated by three cooperative intermolecular O6H···O4 (9.50), N3H···N1 (7.11), and N2H···O2 (4.01) H-bonds (the energy of the each H-bond obtained in the weakly polar surrounding with $\varepsilon = 4$ is presented in kcal·mol$^{-1}$ in brackets after each of them), is experimentally registered in the base pair recognition pocket of the DNA polymerase (Bebenek et al., 2011) (see detailed discussion below).

It was found that all six low-frequency (Shishkin, Pelmenschikov, Hovorun, & Leszcynski, 2000) intermolecular vibrations (23.2, 36.2, 64.5, 85.1, 91.3, and 104.5 cm$^{-1}$ obtained at the MP2/6–311+G(d,p) level of QM theory in the continuum with $\varepsilon = 4$) are able to develop during the lifetime of the G·T* base mispair ($\tau = 6.39 \cdot 10^{-12}$ obtained at the MP2/aug-cc-pVTZ/MP2/6–311++G(d,p) level of QM theory in the continuum with $\varepsilon = 4$), since their periods $T (1.44 \cdot 10^{-12}, 9.23 \cdot 10^{-13}, 5.17 \cdot 10^{-13}, 3.92 \cdot 10^{-13}, 3.65 \cdot 10^{-13}$, and $3.19 \cdot 10^{-13}$ s) are less than this time interval.

The time necessary to reach 99.9% of the equilibrium concentration between the G*·T and G·T* DNA base mispairs ($\tau_{99.9\%} = 4.95 \cdot 10^{-11}$ s (Table 5)) was established to be less by eight orders than the time given by the replication machinery to incorporate one nucleotide into the synthesized DNA double helix ($10^{-3}$ s (Brovarets’, Kolomiets’, et al., 2012; Kornberg & Baker, 1992)).

Taking into account the above arguments, we summarized that the extremely stable G·T* DNA base mispair “slips away from the hands” of the replication machinery into the G*·T DNA base mispair with the lower energy, that further dissociates into the G* and T monomers.

### Table 3. Interbase interaction energies (in kcal mol$^{-1}$) for the G*·T/G·T* base mispairs and A·T/G·C Watson–Crick DNA base pairs obtained at the B3LYP/6–311++G(d,p) (DFT) and MP2/6–311++G(d,p) (MP2) levels of theory in vacuo ($\varepsilon = 1$) and in the continuum with a low dielectric constant ($\varepsilon = 4$).

| Base pairs | $\Delta E_{\text{int}}$ | $\Sigma E_{\text{HB}}$ | $\Sigma E_{\text{HB}}/|\Delta E_{\text{int}}|$,% | $\Delta G_{\text{int}}$ |
|------------|------------------|------------------|------------------|------------------|
|            | DFT              | MP2              | DFT              | MP2              | DFT              | MP2              | DFT              | MP2              |
| $\varepsilon = 1$ |
| G*·T       | 19.79            | 18.92            | 17.35            | 17.73            | 87.7             | 93.7             | 7.09             | 5.46             |
| G·T*       | 33.40            | 31.50            | 20.48            | 20.28            | 61.3             | 64.4             | 20.66            | 18.26            |
| A·T$^b$    | 14.92            | 14.55            | 12.97            | 12.39            | 86.9             | 85.1             | 1.43             | 2.49             |
| G·C$^b$    | 29.28            | 27.87            | 17.79            | 17.44            | 60.8             | 62.6             | 15.97            | 14.17            |
| $\varepsilon = 4$ |
| G*·T       | 16.66            | 19.90            | 20.91            | 20.62            | 125.5            | 103.6            | 3.85             | 6.32             |
| G·T*       | 36.99            | 23.76            | 17.19            | 18.22            | 46.5             | 76.7             | 24.11            | 10.19            |
| A·T$^b$    | 11.37            | 10.99            | 12.60            | 11.50            | 110.8            | 104.6            | .49              | .79              |
| G·C$^b$    | 21.44            | 20.38            | 17.76            | 17.38            | 82.8             | 85.3             | 7.82             | 6.54             |

Note: $\Delta E_{\text{int}}$ – the BSSE corrected electronic interaction energy; $\Sigma E_{\text{HB}}$ – the total energy of the intermolecular H-bonds; $\Delta G_{\text{int}}$ – the BSSE corrected Gibbs free energy of interaction (T = 298.15 K).

bData are taken from Brovarets’ and Hovorun (2013e, 2014b).

### Table 4. Selected geometrical parameters of the studied G*·T/G·T* base mispairs and A·T/G·C Watson–Crick base pairs obtained at the B3LYP/6–311++G(d,p) (DFT) and MP2/6–311++G(d,p) (MP2) levels of theory in vacuo ($\varepsilon = 1$) and in the continuum with a low dielectric constant ($\varepsilon = 4$).

| Geometrical parameters | G*·T | G·T* | A·T$^b$ | G·C$^b$ |
|------------------------|------|------|---------|---------|
|                        | DFT  | MP2  | DFT     | MP2     | DFT     | MP2     | DFT     | MP2     |
| $\varepsilon = 1$ |
| R(H–H)                 | 10.291 | 10.214 | 10.202 | 10.140 | 10.132 | 9.970 | 10.213 | 10.149 |
| $\alpha_1$             | 51.5  | 51.6  | 50.6    | 50.8    | 54.2    | 55.2    | 53.1    | 52.9    |
| $\alpha_2$             | 51.1  | 51.6  | 52.2    | 52.9    | 54.7    | 57.1    | 55.1    | 55.4    |
| $\varepsilon = 4$ |
| R(H–H)                 | 10.335 | 10.254 | 10.219 | 10.146 | 10.194 | 9.974 | 10.213 | 10.123 |
| $\alpha_1$             | 50.4  | 50.6  | 50.2    | 50.4    | 53.4    | 54.7    | 52.6    | 52.7    |
| $\alpha_2$             | 50.0  | 50.5  | 51.9    | 52.6    | 53.9    | 56.6    | 55.2    | 55.2    |

Notes: R(H–H) – distance between the H1 and H9 glycosidic protons, Å; $\alpha_1$ ($\angle$N9HH) and $\alpha_2$ ($\angle$N1HH) – glycosidic angles, degree.

bData are taken from Brovarets’ and Hovorun (2013e, 2014b).
**Foreseen reaction pathway of the G*-T*↔G·T* tautomerization via the DPT**

The spacings of the χ-like crossings on the calculated profiles of the $\rho$O6H/HO4, $\rho$N1H/HN3, $\Delta\rho$O6H/HO4 and $\Delta\rho$N1H/HN3 values of the intermolecular H-bonds along the IRC of the tautomerization (Figures 5, 6, 11, S4, S5, and S10) evidence that this tautomerization reaction occurs through the asynchronous concerted mechanism. Furthermore, we have established the nine key points (KPs), enabling us to divide the reaction pathway of the tautomerization into three different regions: the reactant (from $-4.35/-4.67/-5.78/-7.04$ to $-2.29/-3.11/-4.44/-5.31$ Bohr), transition state (from $-2.29/-3.11/-4.44/-5.31$ to $1.97/1.59/1.58/1.15$ Bohr), and product regions (from $1.97/1.59/1.58/1.15$ to $5.36/5.90/6.64/6.91$ Bohr), divided by the KP 2, coinciding with the first extremum (maximum) of the first derivative of the electronic energy with respect to the IRC (dE/dIRC) or the minimum of the reaction force (Politzer, Murray, & Jaque, 2013), and the KP 8, coinciding with the second extremum (minimum) of the dE/dIRC or the maximum of the reaction force (Politzer et al., 2013) (Figures 2, 3, S1 and S2). It should be noted that all intermolecular H-bonds revealed in the nine key points, the detailed physico-chemical characteristics of which are presented in Table 2, meet all requirements for the canonical H-bonding (see Brovarets’ et al., 2013a; Ponomareva et al., 2012; Yurenko et al., 2011 and bibliography therein).

Obtained data suggest that the energy equal to $4.84/5.66/4.81/4.73$ kcal·mol$^{-1}$, representing $79.3/77.3/77.0/37.9\%$ of the TS$_{G^*-T^*}$ electronic energy relatively to the starting G*-T DNA base mispair, is spent at the reactant region on the mutual adjustment of the G* and T DNA bases within the G*-T base mispair getting the donor and acceptor atoms closer to each other in order to initiate the DPT reaction at the TS region, whereas the energy equal to $3.43/4.13/5.53/5.93$ kcal·mol$^{-1}$, representing $56.2/56.3/80.5/47.5\%$ of the TS$_{G^*-T^*}$ electronic energy relatively to the starting G*-T DNA base mispair, is spent at the product region on the structural relaxation of the KP 8 to the final G-T DNA base mispair. These data suggest that almost equal amount of the electronic energy is spent at the reaction region just before the beginning of the chemical reaction and at the product region immediately after the ending of the chemical reaction.

The obtained reaction path enable us to track in detail the atomistic mechanism of the G*-T*↔G·T* DPT tautomerization. This reaction starts from the transfer of the proton localized at the N3 nitrogen atom of the T base in the G*-T DNA base mispair (the KP 1) along the N3H···N1 H-bond in the direction of the N1 nitrogen atom of the G* enol tautomer, reaching the KP 2 ($\Delta\rho$N1···H = 0), at which the H-N3 covalent bond of the T base is significantly weakened and the N1···H H-bond becomes the N1-H covalent bond. Remarkably, exactly at this point the chemical identity of the G* and T bases is lost and in fact the chemical reaction is initiated. Further, bypassing the KP 3 ($\rho$N1···H = $\rho$H···N3), characterized by the equivalent N1-H/H-N3 covalent bonds involved in the N1-H-N3 covalent bridge, the complex reaches the TS$_{G^*-T^-·G·T^*}$ (the KP 4). After this the base mispair transforms into the KP 5 ($\Delta\rho$H···N3 = 0), at which the H-N3 covalent bond becomes the H···N3 H-bond thereby arriving at the G'·T' zwitterionic non-stable intermediate. Then after the first proton is already transferred, the second mobile proton, localized at the O6+ oxygen atom of the G' protonated base, begins to move towards the O4− oxygen atom of the T deprotonated base that is accompanied by the breaking of the old O6-H covalent bond at the KP 6 ($\Delta\rho$H···O4 = 0), by the formation of the equivalent O6-H/H-O4 covalent bonds

Figure 2. Profiles of the electronic energy $\Delta E$ along the IRC of the G*-T*↔G·T* tautomerization via the DPT obtained at the MP2 level of theory (a) in vacuo and (b) in the continuum with a low dielectric constant ($\varepsilon = 4$).
Evolution of the major physicochemical properties of the G*·T and G·T* DNA base mispairs along the IRC of the G*·T↔G·T* DPT tautomerization

It should be mentioned that we did not reveal principal differences at the analysis of the graphs of the electron-topological characteristics obtained at the DFT and MP2 levels of QM theory (Figures 2–15 and S1–S14).

The calculated electronic energy profile of the DPT tautomerization is asymmetric with shoulder (Figures 2 and S1). It is noteworthy that sharp peak on this profile that corresponds to the TS_{G*·T→G·T*} is observed at the MP2 level of theory in the continuum with ε = 4 (Figure 2(b)).

On the profile of the first derivative of the electronic energy with respect to the IRC (dE/dIRC) there are one maximum (7.28/8.96/7.50/7.35 kcal·(mol·Bohr)^{-1}) at the KP2 and two minima: −3.85/−3.93/−6.1/−9.8 kcal·(mol·Bohr)^{-1} at the KP5 and −3.76/−4.97/−3.89/−5.09 kcal·(mol·Bohr)^{-1} at the KP8 (Figures 3 and S2). The graph of the changes of the dipole moment μ has an U-like shape with plateau changing at this within the range of values 3.29–5.77/4.35–7.20/2.99–4.98/3.94–6.66 D along the IRC of the G*·T↔G·T* tautomerization (Figure 4 and S3).

It was established that the curves of the electron density \( ρ \) (0.040–0.331/0.038–0.324/0.039–0.328/0.039–0.320 a.u.), the Laplacian of the electron density \( Δρ \) (−2.327–1.168/−2.265–1.167/−2.328–1.172/−2.245–1.70 a.u.), and the distance \( d_{ΔH/HB} \) between the hydrogen and electronegative A or B atoms (1.89/1.839/1.994–1.862/0.987/1.810/0.994–1.839 Å) for the N1-H bond intersect with those for the N3-H bond at the KP3, whereas the graph for the O6-H bond intersects with those for the O4-H bond at the KP7, indicating that their values are equalized at the points of intersection (Figures 5, 6, 11, S4, S5, and S10). The profile of the ellipticity ε for the N1-H bond \( ε_{N1-H} \) intersects with those for the N3-H bond \( ε_{N3-H} \) between the KPs 2 and 3 (.932–.663/0.026–0.026/0.029–0.060/0.029–0.059), while the graphs of the \( ε_{O4-H} \) and \( ε_{O6-H} \) values cross with each other exactly at the KP7 (.012–.027/0.013–.028/0.012–.023/0.013–.026) (Figures 7 and S6).

Our QM simulations, based on the studies of the energetic profiles of the intermolecular H-bonds along the IRC (Figures 8 and S7), revealed that they are cooperative (Mishchuk et al., 2000) both in the G*·T (\( d_{O6H--O4}/d_{N3H--N1}/d_{N2H--O2} = 1.00/1.02/0.27; 1.00/1.04/0.38; 1.00/1.28/0.07; 1.00/2.04/1.35 \)) and G·T* (\( d_{O4H--O6}/d_{N1H--N3}/d_{N2H--O2} = 1.00/0.76/0.01; 1.00/0.72/1.61; 1.00/0.88/2.25; 1.00/0.88/0.30 \) obtained at the DFT (ε = 1); DFT(ε = 4); MP2(ε = 1); MP2(ε = 4) levels of theory, respectively) DNA base mispairs.

The current study found that the upper O6H···O4 H-bond from the side of the major groove (11.05–53.41 kcal·mol^{-1}) and the middle N3H···N1 H-bond (9.81–38.83 kcal·mol^{-1}) exist within the range from the KP1 to KP6 and from the KP1 to KP2, gaining at this their maximum values at the KPs 6 and 2, respectively, whereas the upper O4H···O6 H-bond (20.10–54.74 kcal·mol^{-1}) from the side of the major groove and
The nature of the transition mismatches with Watson–Crick architecture

Figure 4. Profiles of the dipole moment $\mu$ along the IRC of the $G^*\cdot T \leftrightarrow G\cdot T^*$ tautomerization via the DPT obtained at the MP2 level of theory (a) in vacuo and (b) in the continuum with a low dielectric constant ($\varepsilon = 4$).

Figure 5. Profiles of the electron density $\rho$ at the (3,−1) BCPs of the covalent and hydrogen bonds along the IRC of the $G^*\cdot T \leftrightarrow G\cdot T^*$ tautomerization via the DPT obtained at the MP2 level of theory (a) in vacuo and (b) in the continuum with a low dielectric constant ($\varepsilon = 4$).

Figure 6. Profiles of the Laplacian of the electron density $\Delta\rho$ at the (3,−1) BCPs of the covalent and hydrogen bonds along the IRC of the $G^*\cdot T \leftrightarrow G\cdot T^*$ tautomerization via the DPT obtained at the MP2 level of theory (a) in vacuo and (b) in the continuum with a low dielectric constant ($\varepsilon = 4$).
the middle N1H···N3 H-bond (10.22–37.76/10.03–36.18/11.70–40.12 kcal·mol$^{-1}$) exist within the diapason from the KP 8 to KP 9 and from the KP 5 to KP 9, acquiring their maximum values at the KPs 8 and 5, respectively (Figures 8 and S7).

Another important finding is that the third N2H···O2 H-bond exists continuously throughout the reaction pathway of the G*·T ↔ G·T* tautomerization, while its electron-topological, energetic, and geometric characteristics vary marginally: ρ (0.022–0.039/0.021–0.037/0.020–0.033/0.019–0.032 a.u.), Δρ (0.078–0.124/0.074–0.122/0.072–0.113/0.068–0.109 a.u.), ε (0.052–0.058/0.054–0.060/0.049–0.055/0.052–0.058), $E_{N2H···O2}$ (4.68–10.44/4.40–9.95/4.44–8.87/4.18–8.29 kcal·mol$^{-1}$), $d_{N2H}$ (1.014–1.031/1.015–1.029/1.016–1.028/1.017–1.026 Å), and $d_{N2H···O2}$ (1.767–2.006/1.782–2.028/1.824–2.041/1.846–2.065 Å) (Figures 5, 6, 7, 8, 11, S4, S5, S6, S7, and S10).

The results of the present study show the “breathing” of the G*·T DNA base mispair in the course of the tautomerization, in particular, significant compression is observed at the TS region, as is evidenced by the changes in the distance between the O6 and O4 oxygen atoms ($d_{O6···O4}$) (2.425–2.719/2.426–2.690/2.410–2.718/2.414–2.668 Å), the N1 and N3 nitrogen atoms ($d_{N1···N3}$) (2.614–2.885/2.614–2.902/2.582–2.853/2.579–2.874 Å), and the N2 nitrogen and O2 oxygen atoms ($d_{N2···O2}$) (2.797–3.018/2.614–2.902/2.582–2.853/2.579–2.874 Å) involved in the intermolecular H-bonds (Figures 10 and S9). Moreover, the changes of the distance $R(H_{N9}···H_{N1})$ between the H1 and H9 glycosidic hydrogens (10.026–10.293/10.049–10.335/9.960–10.218/9.987–10.278 Å) were also found to cause this effect (Figures 13 and S12). The variations of the $\angle O6H···O4$ (171.1–174.3/169.8–173.1/171.0–173.8/169.5–173.0°),
\[ \angle \text{N1H} \cdots \text{N3} (172.1-177.8/175.2-178.4/178.4-178.0/178.4-178.0/178.4-178.0^\circ), \text{ and } \angle \text{N2H} \cdots \text{N2} (174.0-178.4/175.2-178.4/178.4-178.0\text{H-bond angles, the } \alpha_1 (\angle \text{N9HH}) (50.06-51.8/50.2-50.8/50.8-51.6/49.9-50.6^\circ) \text{ and } \alpha_2 (\angle \text{N1HH}) (50.7-52.2/50.0-51.9/51.0-52.9/50.0-52.5^\circ) \text{ glycosidic angles, and the } \angle \text{C6N1(G)N3C4(T)} \text{ dihedral angle (3.4-7.9/26.3-31.9/27.4-32.7^\circ) have been established (Figures 12, 14, 15, S11, S13, and S14).}

The scans of the NBO charges of the hydrogen atoms involved in the N3HI \cdots N1, O6HII \cdots O4 and N2HIII \cdots O2 H-bonds do not intersect with each other along the IRC of the tautomerization and can be arranged in the following order by the value of their charges: \( q_{\text{N3H} \cdots \text{N1}} (0.459-0.519/0.462-0.518/0.477-0.506 \text{ e}) > q_{\text{O6H} \cdots \text{O4}} (0.442-0.462/0.441-0.459/0.429-0.449/0.429-0.447 \text{ e}) > q_{\text{N2H} \cdots \text{O2}} (0.421-0.442/0.438-0.406/0.428-0.407/0.424 \text{ e}) \) (Figures 9 and S8).

\section*{What tautomeric form possesses the G-T mismatch with Watson–Crick geometry in the recognition pocket of the high-fidelity DNA polymerase?}

In order to answer this biologically important question, we have compared the experimental structural data obtained by X-ray analysis (Bebenek et al., 2011) with similar theoretical data presented in this paper for the G* \cdots T and G \cdots T* base mispairs in the continuum with \( \varepsilon = 4 \) at the MP2/6–311++G(d,p) level of theory (Table 6).

Satisfactory coincidence of the experimental data with theoretical one leaves no doubt that the enzymatically competent G \cdots T DNA mismatch exactly possesses electroneutral, energetically favorable G* \cdots T tautomeric form and is stabilized by three intermolecular O6H \cdots O4, N1H \cdots N3, and N2H \cdots O2 H-bonds. The corresponding anionic (deprotonated) G' \cdots T or G \cdots T' forms are excluded: we have previously shown (Brovarets' et al., 2010) that they have significantly non-planar structure, contradicting to the experimental data (Bebenek et al., 2011), which clearly demonstrate the near-planar structure of the G-T base mispair with Watson–Crick geometry.

Unfortunately, the numerical structural characteristics of the enzymatically competent A \cdots C base mispair with Watson–Crick geometry, that was experimentally registered in the recognition pocket of the DNA polymerase, have not been provided in the work (Wang et al., 2011). However, based on the previously obtained theoretical results (Brovarets' & Hovorun, 2013c), it can be argued by analogy with the G \cdots T base pair, that exactly the A \cdots C* tautomer with Watson–Crick geometry presents an enzymatically competent A \cdots C base mispair, which is experimentally recorded.

\section*{Concluding remarks}

In this paper we have discovered the physico-chemical picture of the G* \cdots T \leftrightarrow G \cdots T* tautomerization via the DPT in vacuum and in the continuum with a low dielectric constant (\( \varepsilon = 4 \)), typical for the hydrophobic interfaces of specific protein–nucleic acid interactions, within the framework of the QM/QTAIM approach. The calculations revealed that the G* \cdots T DNA base mispair joined by the cooperative O6H \cdots O4, N1H \cdots N3, and N2H \cdots O2 H-bonds tautomerizes into the G \cdots T* base mispair (\( \Delta G = 0.99 \text{ and } \Delta E = 1.31 \text{ kcal mol}^{-1} \) in the continuum with \( \varepsilon = 4 \)) stabilized by the cooperative O6 \cdots HO4, N1H \cdots N3, and N2H \cdots O2 H-bonds through the asynchronous concerted DPT without any intermediates along two antiparallel O6H \cdots O4 and N1H \cdots N3 H-bonds and is accompanied by the electronic and structural reconstruction maintaining at this its Watson–Crick-like geometry.
Figure 10. Profiles of the distance $d_{A\cdots B}$ between the electronegative A and B atoms of the $AH\cdots B$ H-bonds along the IRC of the $G^*\cdot T\leftrightarrow G\cdot T^*$ tautomerization via the DPT obtained at the MP2 level of theory (a) in vacuo and (b) in the continuum with a low dielectric constant ($\varepsilon = 4$).

Figure 11. Profiles of the distance $d_{A/H\cdots B}$ between the hydrogen and electronegative A or B atoms of the $AH\cdots B$ H-bonds along the IRC of the $G^*\cdot T\leftrightarrow G\cdot T^*$ tautomerization via the DPT obtained at the MP2 level of theory (a) in vacuo and (b) in the continuum with a low dielectric constant ($\varepsilon = 4$).

Figure 12. Profiles of the angle $\angle AH\cdots B$ of the $AH\cdots B$ H-bonds along the IRC of the $G^*\cdot T\leftrightarrow G\cdot T^*$ tautomerization via the DPT obtained at the MP2 level of theory (a) in vacuo and (b) in the continuum with a low dielectric constant ($\varepsilon = 4$).
Figure 13. Profiles of the distance $R(H_9-H_1)$ between the H1 and H9 glycosidic protons along the IRC of the $G^*\cdot T\leftrightarrow G\cdot T^*$ tautomerization via the DPT obtained at the MP2 level of theory (a) in vacuo and (b) in the continuum with a low dielectric constant ($\varepsilon = 4$).

Figure 14. Profiles of the $\alpha_1$ ($\angle N9HH$) and $\alpha_2$ ($\angle N1HH$) glycosidic angles along the IRC of the $G^*\cdot T\leftrightarrow G\cdot T^*$ tautomerization via the DPT obtained at the MP2 level of theory (a) in vacuo and (b) in the continuum with a low dielectric constant ($\varepsilon = 4$).

Figure 15. Profile of the $\angle C6N1(G)N3C4(T)$ dihedral angle along the IRC of the $G^*\cdot T\leftrightarrow G\cdot T^*$ tautomerization via the DPT obtained at the MP2 level of theory (a) in vacuo and (b) in the continuum with a low dielectric constant ($\varepsilon = 4$).
Table 5  Selected physico-chemical characteristics of the G*T ↔ G*T* tautomerization via the DPT obtained at the different levels of QM theory in vacuo (ε=1) and in the continuum with a low dielectric constant (ε=4).

| Levels of QM theory | ε  | ΔG^a | ΔE^b | ΔΔG^T_S^c | ΔΔE^T_S^d | ΔΔG^e | ΔΔE^f | ν^g | E_zPE^h | τ^i | τ_99.9% | K^k |
|--------------------|----|------|------|-------------|-------------|------|------|----|---------|----|---------|-----|
| MP2/6-311++G(2df,pd)//B3LYP/6-311++G(d,p) | 1  | 1.16 | 1.14 | 2.33        | 5.31        | 1.17 | 4.17 | 1460.1 | 3018.9 | 4.32 | 1509.5 | 5.43·10^{-13} | 4.38·10^{-12} | 0.14 |
|                      | 4  | 0.71 | 1.10 | 3.42        | 6.40        | 2.71 | 5.30 | 1854.8 | 2978.7 | 4.26 | 1489.3 | 6.86·10^{-12} | 4.84·10^{-11} | 0.30 |
| MP2/6-311++G(3df,2pd)//B3LYP/6-311++G(d,p) | 1  | 1.27 | 1.24 | 2.76        | 5.74        | 1.49 | 4.50 | 1572.9 | 3018.9 | 4.32 | 1509.5 | 9.34·10^{-13} | 7.69·10^{-12} | 0.12 |
|                      | 4  | 0.75 | 1.15 | 3.86        | 6.85        | 3.11 | 5.70 | 1993.8 | 2978.7 | 4.26 | 1489.3 | 1.34·10^{-11} | 9.63·10^{-11} | 0.28 |
| MP2/cc-pVTZ/B3LYP/6-311++G(d,p)         | 1  | 1.40 | 1.38 | 2.47        | 5.44        | 1.07 | 4.07 | 1422.8 | 3018.9 | 4.32 | 1509.5 | 4.54·10^{-13} | 3.82·10^{-12} | 0.09 |
|                      | 4  | 0.92 | 1.31 | 3.58        | 6.57        | 2.66 | 5.25 | 1838.0 | 2978.7 | 4.26 | 1489.3 | 6.33·10^{-12} | 4.81·10^{-11} | 0.21 |
| MP2/6-311++G(d,p)         | 1  | 1.76 | 2.34 | 4.34        | 6.91        | 2.58 | 4.57 | 1598.7 | 3023.5 | 4.32 | 1511.8 | 8.60·10^{-12} | 7.53·10^{-11} | 0.05 |
|                      | 4  | 1.69 | 2.01 | 4.97        | 7.82        | 3.28 | 5.81 | 2032.1 | 2971.8 | 4.25 | 1485.9 | 2.55·10^{-11} | 2.22·10^{-10} | 0.06 |
| MP2/6-311++G(2df,pd)//MP2/6-311++G(d,p) | 1  | 0.55 | 1.13 | 2.46        | 5.03        | 1.91 | 3.90 | 1363.7 | 3023.5 | 4.32 | 1511.8 | 7.77·10^{-12} | 1.81·10^{-11} | 0.40 |
|                      | 4  | 0.73 | 1.05 | 3.25        | 6.10        | 2.52 | 5.05 | 1766.9 | 2971.8 | 4.25 | 1485.9 | 7.12·10^{-12} | 5.07·10^{-11} | 0.29 |
| MP2/6-311++G(3df,2pd)//MP2/6-311++G(d,p) | 1  | 0.62 | 1.20 | 2.91        | 5.48        | 2.29 | 4.28 | 1497.7 | 3023.5 | 4.32 | 1511.8 | 5.29·10^{-12} | 3.59·10^{-11} | 0.35 |
|                      | 4  | 0.75 | 1.07 | 3.69        | 6.54        | 2.94 | 5.47 | 1914.8 | 2971.8 | 4.25 | 1485.9 | 1.45·10^{-11} | 1.04·10^{-10} | 0.28 |
| MP2/cc-pVTZ//MP2/6-311++G(d,p) | 1  | 0.84 | 1.42 | 2.66        | 5.23        | 1.83 | 3.82 | 1334.7 | 2971.8 | 4.32 | 1511.8 | 2.41·10^{-12} | 1.78·10^{-11} | 0.24 |
|                      | 4  | 0.99 | 1.31 | 3.45        | 6.30        | 2.46 | 4.99 | 1744.5 | 2971.8 | 4.25 | 1485.9 | 6.39·10^{-12} | 4.95·10^{-11} | 0.19 |

*The relative Gibbs free energy of the G*T* DNA base mispair (ΔG_{G*T*}=0 kcal mol⁻¹; T=298.15 K) (kcal mol⁻¹).
*The relative electronic energy of the G*T* DNA base mispair (ΔE_{G*T*}=0 kcal mol⁻¹) (kcal mol⁻¹).
*The Gibbs free energy of activation for the forward reaction of tautomerisation (T=298.15 K) (kcal mol⁻¹).
*The activation electronic energy for the forward reaction of tautomerisation (kcal mol⁻¹).
*The Gibbs free energy of activation for the reverse reaction of tautomerisation (T = 298.15 K) (kcal mol⁻¹).
*The activation electronic energy for the reverse reaction of tautomerisation (kcal mol⁻¹).
*The frequency of the vibrational mode of the tautomerized complex, which becomes imaginary in the TS_{G*T*} of tautomerization, obtained at the level of geometry optimization (cm⁻¹).
*The zero-point vibrational energy associated with this normal mode (cm⁻¹).
*The lifetime of the G*T* DNA base mispair (s).
*The time necessary to reach 99.9% of the equilibrium concentration of the G*T and G*T* base mispairs (s).
*The equilibrium constant of the G*T ↔ G*T* tautomerization via the DPT.
Table 6. Comparison of the geometrical parameters of the intermolecular H-bonds in the G·T transition mismatch with Watson–Crick geometry, obtained experimentally by X-ray analysis and theoretically by ab initio method (MP2/6-311++G (d,p) level of theory in the continuum with a low dielectric constant ($\varepsilon=4$)).

| A···B H-bond | Experiment (Bebenek et al., 2011) | Theory G*·T | Theory G·T* |
|--------------|----------------------------------|-------------|-------------|
| O6H···O4     | 2.66                             | 2.681       | 2.938       |
| N1···HN3     | 2.97                             | 2.873       | 2.836       |
| N2H···O2     | 3.06                             | 3.061       | 3.461       |

As a result, for the first time we have obtained the scans of the energetic, electron-topological, geometrical, polar, and NBO parameters along the IRC. The current findings add substantially to our understanding of the atomistic details of the G*·T* ↔ G·T* DPT tautomerization enabling us to distinguish the nine key points that illustrate in detail the progress of the tautomerization.

By providing a comprehensive investigation and analysis of the barrier heights and kinetic parameters of this tautomerization, it was found that the thermodynamically stable G·T* DNA base mispair is dynamically unstable structure in vacuum ($\tau=2.41\cdot10^{-12}$ s) and dynamically stable structure in the continuum with $\varepsilon=4$ ($\tau=6.39\cdot10^{-12}$ s), which lifetime enables all six low-frequency intermolecular vibrations to develop; however, it is by three orders less than the period of time necessary for the replication machinery to forcibly dissociate a base pair into the monomers during DNA replication (several ns).

Returning to the question posed in the title of this study, it is now possible to state that the G·T* base mispair is too short-lived structure to cause transition mutations in the course of the DNA replication. Taken together, these results suggest that the replication machinery as a part of the replisome deals only with the G*·T base mispair, effectively dissociating it on the G* and T monomers (at the next act of replication the G* enol mutagenic tautomer complementary binds with the T keto canonical tautomer, thereby inducing the appropriate point mutations), in such a way substantially altering the Löwdin’s scheme (Löwdin, 1963, 1966) of the fixation of the replication point errors in DNA, that suggests that both the G*·T and G·T* DNA mismatches are equivalent participants in the mutation process.

It is logically to suggest that the enzymatically competent G*·T DNA base mispair can be formed in two ways in the base pair recognition pocket of the high-fidelity DNA polymerase during the DNA biosynthesis: as a result of the complementary interaction between the G* and T bases (G* + T → G·T*) or through the transformation of the formed G·T* DNA base mispair (G + T* → G·T* → G*·T). These data mean that the short-lived G·T* DNA base mispair serves as a mediated provider of the G*·T transition mismatch during DNA replication and therefore it is important from the biological point of view.

Supplementary material
The supplementary material for this paper is available online at http://dx.doi.10.1080/07391102.2014.924879.

Supporting information
Sweeps 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,−1) BCPs of the intrapair H-bonds, the NBO charges of the hydrogen atoms involved in the intermolecular H-bonds, the glycosidic angles, the distance between the glycosidic hydrogens, and the dihedral angle along the IRC of the G*·T* ↔ G·T* tautomerization via the DPT obtained at the B3LYP/6–311++G(d,p) level of theory in vacuum and in the continuum with $\varepsilon=4$. This material is available free of charge via the Internet at http://www.tandfonline.com/toc/tbsd.

Acknowledgments
The authors express their sincere gratitude to Prof. Cherif F. Matta (Mount Saint Vincent University, Halifax, Canada) for fruitful discussion of the work, valuable suggestions, and comments.

Funding
This work was partially supported by the State Fund for Fundamental Research (SFFR) of Ukraine within the Ukrainian-Japanese project № F 52.2/001 for 2013–2014 years and by the Science and Technology Center in Ukraine (STCU) within the project № 5728 for 2012–2014 years. O.O.B. was supported by the Grant of the President of Ukraine to support scientific research of young scientists for 2014 year from the State Fund for Fundamental Research of Ukraine (project № GP/F56/074) and by the Grant of the Cabinet of Ministers of Ukraine for collectives of young researchers for carrying out applied scientific research, scientific and technological development in the priority areas of science and technology (№ 123). 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.

References
Arabi, A. A., & Matta, C. F. (2011). Effects of external electric fields on double proton transfer kinetics in the formic acid dimer. Physical Chemistry Chemical Physics, 13, 13738–13748.
Brovarets’ O. O., & Hovorun, D. M. (2013c). The physical-chemical essence of the purine pyrimidine transition mismatches with Watson–Crick geometry in DNA: A C* vs A* C. A QM and QTAIM atomistic understanding. Journal of Biomolecular Structure & Dynamics. doi:10.1080/07391102.2013.852133

Brovarets’, O. O., & Hovorun, D. M. (2013d). Prototropic tautomerism and basic molecular principles of hypoxanthine mutagenicity: An exhaustive quantum-chemical analysis. Journal of Biomolecular Structure and Dynamics, 31, 913–936.

Brovarets’, O. O., & Hovorun, D. M. (2013e). 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. doi:10.1080/07391102.2013.822929

Brovarets’, O. O., & Hovorun, D. M. (2014a). DPT tautomerization of the G·A syn and A*·G* anti DNA mismatches: A QM/QTAIM combined atomistic investigation. Physical Chemistry Chemical Physics, 16, 9074–9085.

Brovarets’, O. O., & Hovorun, D. M. (2014b). Can tautomerization of the A·T Watson–Crick base pair via double proton transfer provoke point mutations during DNA replication? A comprehensive QM and QTAIM analysis. Journal of Biomolecular Structure & Dynamics, 32, 127–154.

Brovarets’, O. O., & Hovorun, D. M. (2014c). How the long G·G* Watson–Crick DNA base mispair comprising keto and enol tautomers of the guanine tautomeres? The results of the QM/QTAIM quantum-chemical study. Physical Chemistry Chemical Physics, doi:10.1039/C4CP01241K

Brovarets’, O. O., & Hovorun, D. M. (2014d). Does the G·G* syn DNA mismatch containing canonical and rare tautomers of the guanine tautomerise through the DPT? A QM/QTAIM microstructural study. Molecular Physics. doi:10.1080/00268976.2014.927079

Brovarets’, O.O., Kolomiets, I.M., & Hovorun, D.M. (2012). Elementary molecular mechanisms of the spontaneous point mutations in DNA: A novel quantum-chemical insight into the classical understanding. In T. Tada (Ed.), Quantum chemistry – Molecules for innovations (pp. 59–102). Rijeka: In Tech Open Access.

Brovarets’, O. O., Yurenko, Y. P., Dubey, I. Y., & Hovorun, D. M. (2012). Can DNA-binding proteins of replisome tautomerize nucleotide bases? Ab initio model study. Journal of Biomolecular Structure & Dynamics, 29, 1101–1109.

Brovarets’, O. O., Yurenko, Y. P., & Hovorun, D. M. (2013a). Intermolecular CH···O/N H-bonds in the biologically important pairs of natural nucleobases: A thorough quantum-chemical study. Journal of Biomolecular Structure & Dynamics. doi:10.1080/07391102.2013.799439

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. (2013b). The physico-chemical “anatomy” of the tautomerization 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). 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. (2013d). The physico-chemical mechanism of the tautomerization 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. (2014a). DPT tautomerisation of the wobble guanine-thymine DNA base mispair is not mutagenic: QM and QTAIM arguments. Journal of Biomolecular Structure and Dynamics. doi:10.1080/07391102.2014.897259

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

Brovarets’, O. O., Zhurakivsky, R. O., & Hovorun, D. M. (2014c). Structural, energetic and tautomeric properties of the T·T°/T°·T DNA mismatch involving mutagenic tautomer of thymine: A QM and QTAIM insight. Chemical Physics Letters, 592, 247–255.

Brovarets’, O. O., Zhurakivsky, R. O., & Hovorun, D. M. (2014d). A QM/QTAIM microstructural analysis of the tautomerisation via the DPT of the hypoxanthine-adenine nucleobase pair. Molecular Physics. doi:10.1080/00268976.2013.877170

Cossi, M., Rega, N., Scalmani, G., & Barone, V. (2003). Energetics, 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 rare base pairs: The molecular mechanism of the spontaneous substitution mutations conditioned by tautomerism of bases. Chemical Physics Letters, 412, 285–293.

Danilov, V. I., Van Mourik, T., Kurita, N., Wakabayashi, H., Tsukamoto, T., & Hovorun, D. M. (2009). On the mechanism of the mutagenic action of 5-bromouracil: A DFT study of uracil and 5-bromouracil in a water cluster. The Journal of Physical Chemistry A, 113, 2233–2235.

Dewar, M. J. S., & Storch, D. M. (1985). Alternative view of enzyme reactions. Proceedings of the National Academy of Sciences of the United States of America, 82, 2225–2229.

Drake, J. W. (1991). Spontaneous mutation. Annual Review of Genetics, 25, 125–146.

Dunning, T. H., Jr. (1989). Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. The Journal of Chemical Physics, 90, 1007–1023.

Espinosa, E., Molins, E., & Lecomte, C. (1998). Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. Chemical Physics Letters, 285, 170–173.

Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Schultz, R. A., & Ellenberger, T. (2006). DNA repair and mutagenesis. Washington, DC: ASM Press.

Frisch, M. J., Head-Gordon, M., & Pople, J. A. (1990). Semi-direct algorithms for the MP2 energy and gradient. Chemical Physics Letters, 166, 281–289.

Frisch, M. J., Pople, J. A., & Binkley, J. S. (1984). Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets. The Journal of Chemical Physics, 80, 3265–3269.

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., & Cheeseman, J. R., … Pople, J. A. (2010). Gaussian 09 (Revision B.01). Wallingford, CT: Gaussian.

Furmanchuk, A., Isayev, O., Gorb, L., Shishkin, O. V., Hovorun, D. M., & Leszczyński, J. (2011). Novel view on the mechanism of water-assisted proton transfer in the DNA bases: Bulk water hydration. Physical Chemistry Chemical Physics, 13, 4311–4317.

Garcia-Moreno, B. E., Dwyer, J. J., Gittis, A. G., Lattman, E. E., Spencer, D. S., & Stites, W. E. (1997). Experimental measurement of the effective dielectric in the hydrophobic core of a protein. Biophysical Chemistry, 64, 211–224.

Goodman, M. F. (1997). Hydrogen bonding revisited: Geometric selection as a principal determinant of DNA replication fidelity. Proceedings of the National Academy of Sciences of the United States of America, 94, 10493–10495.

Govorun, D. N., Danchuk, V. D., & Mishchuk, Y. R. (1992). AM1 calculation of the nucleic acid bases structure and vibrational spectra. Journal of Molecular Structure, 267, 99–103.

Gutowski, M., Van Lente, J. H., Verbeek, J., Van Duijneveldt, F. B., & Chalasinski, G. (1986). The basis set superposition error in correlated electronic structure calculations. Chemical Physics Letters, 124, 370–375.

Harirhan, P. C., & Pople, J. A. (1973). The influence of polarization functions on molecular orbital hydrogenation energies. Theoretica Chimica Acta, 28, 213–222.

Hovorun, D. M., Gorb, L., & Leszczyński, J. (1999). From the nonplanarity of the amino group to the structural nonrigidity of the molecule: A post-Hartree-Fock ab initio study of 2-aminomimidazole. International Journal of Quantum Chemistry, 75, 245–253.

Hratchian, H. P., & Schlegel, H. B. (2004). Accurate reaction paths using a Hessian based predictor–corrector integrator. The Journal of Chemical Physics, 120, 9918–9924.

Hratchian, H. P., & Schlegel, H. B. (2005a). Finding minima, transition states, and following reaction pathways on ab initio potential energy surfaces. In C. E. Dykstra, G. Frenking, K. S. Kim, & G. Scuseria (Eds.), Theory and applications of computational chemistry (pp. 195–249). Amsterdam: Elsevier.

Hratchian, H. P., & Schlegel, H. B. (2005b). Using Hessian updating to increase the efficiency of a Hessian based predictor-corrector reaction path following method. Journal of Chemical Theory and Computation, 1, 61–69.

Hwang, H., & Taylor, J.-S. (2005). Evidence for Watson–Crick and not Hoogsteen or wobble base pairing in the selection of nucleotides for insertion opposite pyrimidines and a thymine dimer by yeast DNA Pol η. Biochemistry, 44, 4850–4860.

Iogansen, A. V. (1999). Direct proportionality of the hydrogen bonding energy and the intensification of the stretching ν(XH) vibration in infrared spectra. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 55, 1585–1612.

Keith, T. A. (2010). AIMAll (Version 10.07.01). Retrieved from aim.tkgristmill.com

Kendall, R. A., Dunning, T. H., Jr, & Harrison, R. J. (1992). Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. The Journal of Chemical Physics, 96, 6796–6806.

Kool, E. T. (2002). Active site tightness and substrate fit in DNA replication. Annual Review of Biochemistry, 71, 191–219.
Samijlenko, S. P., Yurenko, Y. P., Stepanyugin, A. V., & Hovorun, D. M. (2012). Tautomeric equilibrium of uracil and thymine in model protein–nucleic acid contacts. Spectroscopic and quantum chemical approach. The Journal of Physical Chemistry B, 114, 1454–1461.

Shishkin, O. V., Pelmenschikov, A., Hovorun, D. M., & Leszczynski, J. (2000). Theoretical analysis of low-lying vibrational modes of free canonical 2′-deoxyribonucleosides. Chemical Physics, 260, 317–325.

Sordo, J. A. (2001). On the use of the Boys–Bernardi function counterpoise procedure to correct barrier heights for basis set superposition error. Journal of Molecular Structure: THEOCHEM, 537, 245–251.

Sordo, J. A., Chin, S., & Sordo, T.L. (1988). On the counterpoise correction for the basis set superposition error in large systems. Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta), 74, 104–110.

Tirado-Rives, J., & Jorgensen, W. L. (2008). Performance of B3LYP density functional methods for a large set of organic molecules. Journal of Chemical Theory and Computation, 4, 297–306.

Topal, M. D., & Fresco, J. R. (1976). Complementary base pairing and the origin of substitution mutations. Nature, 263, 285–289.

Von Borstel, R. C. (1994). Origins of spontaneous base substitutions. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 307, 131–140.

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 of the United States of America, 108, 17644–17648.

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

Wigner, E. (1932). Über das Überschreiten von Potentialschw-ellen bei chemischen Reaktionen [Crossing of potential thresholds in chemical reactions]. Zeitschrift für Physikalische Chemie, B19, 203–216.

Yurenko, Y. P., Zhurakivsky, R. O., Samijlenko, S. P., & Hovorun, D. M. (2011). Intramolecular CH···O hydrogen bonds in the AI and BI 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.