PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

DPT tautomerisation of the G·A_{syn} and A*·G*$_{syn}$ DNA mismatches: A QM/QTAIM combined atomistic investigation

Ol'ha O. Brovarets' and Dmytro M. Hovorun✉

[a]Department of Molecular and Quantum Biophysics, Institute of Molecular Biology and Genetics, National Academy of Sciences of Ukraine, 150 Akademika Zabolotnoho Str., 03680 Kyiv, Ukraine

[b]Department of Molecular Biotechnology and Bioinformatics, Institute of High Technologies, Taras Shevchenko National University of Kyiv, 2-h Akademika Hlushkova Ave., 03022 Kyiv, Ukraine

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

Abstract. By applying a combined QM and QTAIM atomistic computational approach we have established for the first time that the G·A_{syn} and A*·G*$_{syn}$ DNA mismatches (rare tautomers are marked with an asterisk), causing spontaneous transversions with substantially various probability, radically differ from each other in their ability to tautomerise through the double proton transfer (DPT). The A*·G*$_{syn}$ mismatch tautomerises quite easily ($\Delta\Delta G_{TS} \approx 4 \cdot kT$ at room temperature) into the A·G*$_{syn}$ mismatch through the asynchronous concerted mechanism, whereas the G·A_{syn} base mispair does not tautomerise *via* the DPT at all, since there is no local minimum corresponding to the tautomerised G*·A*$_{syn}$ mismatch on the potential energy surface. It was established that the A·G*$_{syn}$ base mispair is dynamically unstable H-bonded complex with an extremely low lifetime of $2.17 \cdot 10^{-13}$ s. Consequently, the obtained results allow us to believe that spontaneous or forced dissociation of both the G·A_{syn} and A*·G*$_{syn}$ DNA mismatches by the DNA-polymerase occurs with the preservation of the tautomeric status of the bases.

Keywords. Rare tautomers of the adenine and guanine · *Syn*- and *anti*-conformations· Sweeps of the energetic, electron-topological, geometric, polar and NBO parameters along the IRC · The double proton transfer · Cooperativity of the H-bonds · B3LYP and MP2 levels of QM theory · QTAIM analysis

Introduction

The penetration into the underlying physico-chemical mechanisms that provide the fantastically high precision of the enzymatic DNA biosynthesis in the living cell is an important fundamental problem in structural biology and related disciplines, which is now far from its final solution [1,2,3,4,5,6,7,8]. It is now found for certain that the root cause of spontaneous mutations, in particular point mutations, that inevitably accompany this very important biological process, responsible for the reproduction of the genetic information from generation to generation, is the formation of the mispairs of nucleotide bases adopting the geometry that closely mimics the structure of the Watson-Crick base pairs in the base-pair recognition pocket of the high-fidelity replicational DNA polymerase [9,10,11,12]. Meanwhile, the asymmetry of the spontaneous point mutations in DNA, i.e., the dependence of their frequency of occurrence on the appurtenance of the bases to the template DNA strand or to the incoming deoxynucleotide triphosphate, is one of their least understood features [13,14]. It was experimentally registered (see ref. [15] and bibliography presented there) that the adenine·guanine (A·G) and G·A transversions demonstrate noticeable asymmetry: mistakes happen more often when the A is the base of the incoming deoxynucleotide triphosphate, while the G – of the template strand (here and below in the X·Y mispair the X base belonging to the template strand is in left and the Y base of the incoming deoxynucleotide triphosphate - in the right).

In our recent works [16,17] we have postulated for the first time that the incorrect purine-purine G·A_{syn} DNA base pair [18,19,20] is responsible for the highly probable G·A transversion, while the A*·G*$_{syn}$ base mispair [16,17] – for the low probable A·G transversion (see Scheme 1). The assumption has also been made that both these mispairs are formed in the base-pair recognition pocket of the replicational DNA polymerase through a common intermediate - the so-called long A·G/G·A Watson-Crick base pair, which is stabilized by the N6H···O6 and N1H···N1 hydrogen bonds (H-bonds) and the N2H···HC2 dihydrogen bond [21].

Thus, the A*·G*→A*·G*$_{syn}$ conformational transition is preceded by the double proton transfer (DPT) tautomerisation of the A·G base pair [21], that takes off the steric constraints for its implementation; the G·A_{syn} base mispair is formed through the conformational transformation of the G·A base mispair without changing of its tautomeric status [22]. Within this model, the asymmetry of the transversions is explained by the significantly lower Gibbs free energy of the G·A_{syn} base pair than of the A*·G*$_{syn}$ base pair [16,17]. It should be remarked that this simplest physico-chemical model is based on the experimentally observed regularity - in the recognition pocket of the DNA polymerase only the base in the incoming deoxynucleotide triphosphate can perform *anti→syn* conformational switch; while for the template base this switch is prohibited [10,13,14].

At the same time the biologically important question remains unanswered - what is the atomistic mechanism of the DPT tautomerisation of the G·A_{syn} and A*·G*$_{syn}$ base mispairs causing spontaneous transversions in DNA? An answer to this question provides information about the cooperativity of the intermolecular H-bonds that stabilize these base pairs and also allows us to understand the way in which the dissociation of the G·A_{syn} and A*·G*$_{syn}$ DNA base mispairs by the polymerase machinery occurs – with or without the alteration of the tautomeric status [23,24,25] of the bases that are part of these pairs.

This work is devoted to the searching for the answers to the just formulated questions and discussion of the possible biological consequences that can be derived from them. We have managed to demonstrate that the studied G·A_{syn} and A*·G*$_{syn}$ DNA base mispairs exhibit unique physico-chemical properties: while the first base pair does not tautomerise due to the absence of the local minimum corresponding to the G*·A*$_{syn}$ base mispair on its potential energy surface, the A*·G*$_{syn}$ base mispair tautomerises with the constant of tautomerisation, greatest among all investigated earlier incorrect base pairs [16,17].

It was found for the first time that whereas all three investigated G·A_{syn}, A*·G*$_{syn}$ and A·G*$_{syn}$ base mispairs are thermodynamically stable structures, since the values of their Gibbs free energies of interaction are less than zero, the A*·G*$_{syn}$ base mispair is dynamically unstable complex with an extremely low lifetime of $2.17\cdot10^{-13}$ s, that is by several orders less than the time spent by the DNA-polymerase machinery to forcibly dissociate a base pair into the monomers (several ns [8]).

In the result of the investigation it was found that the A*·G*$_{syn}$↔A·G*$_{syn}$ Löwdin's [23,27,28] tautomerisation *via* the DPT is 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.

It was established based on the profiles of the H-bond energies that the A*·G*$_{syn}$ base mispair is stabilized by the cooperative [29,30,31] O6H···N6 (9.71) and N1H···N7 (6.99) H-bonds, while the A·G*$_{syn}$ base mispair – by the cooperative N6H···O6 (7.23) and N7H···N1 (9.59 kcal·mol^{-1}) H-bonds.

At the same time, a new approach to the definition of the cooperativity of the intermolecular H-bonds in the H-bonded complexes that do not principally tautomerise through the DPT because of the lack of the local minimum corresponding to the tautomerised complex has been proposed for the first time. By using this approach it was found that the N6H···O6 (5.08) and N1H···N7 (6.12 kcal·mol^{-1}) H-bonds in the G·A_{syn} base mispair are cooperative.

Computational Details

Gaussian'09 suite of programs has been used to provide all quantum-chemical calculations [32]. Geometries and harmonic vibrational frequencies of the G·A_{syn}, A*·G^*_{syn} and A·G^*_{syn} DNA base mispairs and the transition state ($TS_{A^*\cdot G^*syn \leftrightarrow A\cdot G^*syn}$) of their mutagenic tautomerisation *via* the DPT were obtained employing Density Functional Theory (DFT) [33] with the B3LYP hybrid functional [34], which includes Becke's three-parameter exchange functional (B3) [35] combined with Lee, Yang and Parr's (LYP) correlation functional [36] in connection with Pople's 6-311++G(d,p) basis set in vacuum. The DFT method has been recommended in the literature for the description of the tautomerisation phenomena of the H-bonded nucleobase pairs in the ground electronic state [37,38,39,40] and for the study of the vibrations of the constituents of the nucleic acids [41,42,43,44,45,46,47], since it has shown a good balance between computational cost and accuracy and therefore can be considered as the shortest way to MP2 results [27,28,48,49,50]. A scaling factor of 0.9668 [51,52] has been used in the present work at the B3LYP quantum-mechanical (QM) level of theory to correct the harmonic frequencies of all the studied structures. We then performed single point energy calculations using correlated MP2 level of theory [53] with the 6-311++G(2df,pd) and 6-311++G(3df,2pd) Pople's [54,55,56] and cc-pVTZ and cc-pVQZ Dunning's cc-type [57,58] basis sets for the B3LYP/6-311++G(d,p) geometries to consider electronic correlation effects as accurately as possible.

The correspondence of the stationary points to the G·A_{syn}, A*·G^*_{syn} and A·G^*_{syn} base mispairs (see Scheme 1) or to the $TS_{A^*\cdot G^*syn \leftrightarrow A\cdot G^*syn}$, located by means of **S**ynchronous **T**ransit-guided **Q**uasi-**N**ewton (STQN) method [59,60], 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.

Since the stationary points (A*·G^*_{syn}, A·G^*_{syn} base mispairs and the $TS_{A^*\cdot G^*syn \leftrightarrow A\cdot G^*syn}$) were located, the IRC path was established by following the IRC in the forward and reverse directions from the TS using the Hessian-based predictor-corrector (HPC) integration algorithm [61,62,63] with tight convergence criteria. These calculations eventually ensure that the proper reaction pathway leads either to the expected reactants or products.

The electronic interaction energies E_{int} have been computed at the MP2/6-311++G(2df,pd)//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, neglecting the deformational energy. In each case the interaction energy was corrected for the basis set superposition error (BSSE) [64,65] through the counterpoise procedure [66,67].

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

$G=E_{el}+E_{corr}$, (1)

where E_{el} – the electronic energy, E_{corr} – the thermal correction.

The lifetime τ of the A·G*$_{syn}$ base mispair was calculated using the formula $1/k_r$, where the value of the reverse k_r rate constant for the A*·G*$_{syn}$↔A·G*$_{syn}$ tautomerisation reaction was obtained as [68,69,70,71]:

$$k_r = \Gamma \cdot \frac{k_B T}{h} e^{-\frac{\Delta\Delta G_r}{RT}}, \quad (2)$$

where $\Gamma = 1 + \frac{1}{24}\left(\frac{h\nu_i}{k_B T}\right)^2$ – Wigner's tunneling correction [72], k_B – Boltzmann's constant, T = 298.15 K – absolute temperature, h – Planck's constant, $\Delta\Delta G_r$ – the Gibbs free energy of activation for the DPT reaction in the reverse direction, R – universal gas constant, ν_i – the magnitude of the imaginary frequency at the $TS_{A^*\cdot G^*syn \leftrightarrow A\cdot G^*syn}$.

Bader's quantum theory "Atoms in molecules" (QTAIM) was applied to analyse the electron density distribution [73], using program package AIMAll [74] with all the default options. Wave functions were obtained at the level of theory used for geometry optimisation. The presence of a bond critical point (BCP) [73], 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 (Δρ≥0), were considered as criteria for the H-bond formation [75,76,77].

The energies of the intermolecular H-bonds E_{HB} in the A*·G*$_{syn}$ and A·G*$_{syn}$ base mispairs were evaluated by the empirical Iogansen's formula [78]:

$E_{HB} = 0.33 \cdot \sqrt{\Delta\nu - 40}$, (3)

where Δν - the magnitude of the redshift (relative to the free molecule) of the stretching mode of the H-bonded groups. The partial deuteration of the CH, NH, OH and NH_2 groups involved in the H-bonding was applied to eliminate the effect of vibrational resonances [77].

The energy of the N6H···O6 H-bond in the $TS_{A^*\cdot G^*syn \leftrightarrow A\cdot G^*syn}$ was estimated by the Nikolaienko-Bulavin-Hovorun formula [79]:

$E_{N6H\cdots O6}$=-2.03+225·ρ, (4)

where ρ is the electron density at the (3,-1) BCP of the N6H···O6 H-bond.

The energies of all intermolecular H-bonds E_{HB} under the investigation of the sweeps of their energies were evaluated by the empirical Espinosa-Molins-Lecomte (EML) formula [80,81] based on the electron density distribution at the (3,-1) BCPs of the H-bonds:

E_{HB}=0.5·V(r), (5)

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

To determine the deformation energy necessary to apply for the investigated base mispairs to acquire the Watson-Crick sizes, we have adjusted the glycosidic angles at the G, A* and A bases in the G·A_{syn}, A*·G^*_{syn} and A·G^*_{syn} base pairs to those corresponding to the A and G bases in the A·T and G·C Watson-Crick base pairs, respectively, while at the A_{syn}, G^*_{syn} and G^*_{syn} bases - to those corresponding to the T and C bases and the distance between the appropriate glycosidic hydrogens and then frozen them using "opt=modredundant" key word.

To determine the cooperativity of the intermolecular H-bonds in the G·A_{syn} base mispair, that does not tautomerise *via* the DPT, we have realized a new approach. The lengths of the N6H and N1H atomic groups – the donors of the N6H···O6 and N1H···N7 H-bonds - has been forcedly stretched (each of them separately) with the step=0.01 Å and then the energy E_{HB} calculated by the EML formula [80,81] has been acquired at the each step at the (3,-1) BCPs of the H-bonds. As a result we obtained the dependency of the energy E_{HB} on the lengths N6H/N1H, that allows us to determine the cooperativity of the N6H···O6 and N1H···N7 H-bonds in the G·A_{syn} base mispair.

Results and Discussion

We have divided this section into three paragraphs aiming to simplify the perception of the material, concretely, the first one is dedicated to the thermodynamical/dynamical stability of the investigated G·A_{syn}, A*·G^*_{syn} and A·G^*_{syn} base mispairs and to their physico-chemical parameters; in the second part the physico-chemical mechanism of the A*·G^*_{syn}↔A·G^*_{syn} tautomerisation *via* the DPT would be discussed; and finally in the third part the evolution of the selected physico-chemical characteristics of the base mispairs and of the intermolecular specific contacts stabilizing them along the IRC of the A*·G^*_{syn}→A·G^*_{syn} DPT tautomerisation would be covered. Obtained by us results are numerically represented in Tables 1-3 and graphically depicted in Scheme 1 and Figures 1-2, S1-S6.

1. Thermodynamical and dynamical stability of the G·A_{syn}, A*·G^*_{syn} and A·G^*_{syn} DNA base mispairs and their main physico-chemical characteristics

1.1. The A*·G^*_{syn} and A·G^*_{syn} base mispairs

Firstly, we have analysed the thermodynamical and dynamical (vibrational) stability of the investigated A*·G^*_{syn} and A·G^*_{syn} base mispairs, containing A base in the *anti*- and G base in the *syn*-orientations, or structural and energetical characteristics of the interbase interactions stabilizing them.

It was found that the planar A*·G^*_{syn} (ΔG=ΔE=0.00), stabilized by the O6H···N6 (9.71) and N1H···N7 (6.99) H-bonds, and the planar A·G^*_{syn} (ΔG=1.89 and ΔE=2.20 kcal·mol^{-1}), stabilized by the

N6H···O6 (7.23) and N7H···N1 (9.59 kcal·mol^{-1}) H-bonds, structures with C_s symmetry and planar heterocycles of the flexible [41,83,84] A and G bases, are thermodynamically stable, since their commensurable Gibbs free energies of interaction are less than zero (ΔG_{int}=-11.47 (A*·G*$_{syn}$) and -12.96 (A·G*$_{syn}$) kcal·mol^{-1}) (Tables 1-3, Fig. 2). In contrast to the A·A* [21], A·A*$_{syn}$ [31] and A·G [71] base mispairs, we have not established electron-topological and spectroscopic manifestations of the third specific C2H···HC8 contact as in the A*·G*$_{syn}$, A·G*$_{syn}$ and TS$_{A*·G*syn↔A·G*syn}$ stationary structures, so along the IRC of the A*·G*$_{syn}$↔A·G*$_{syn}$ tautomerisation *via* the DPT. Moreover, the electronic and Gibbs free interaction energies for the A*·G*$_{syn}$ (ΔG_{int}=-11.47 and ΔE_{int}=-23.00) and A·G*$_{syn}$ (ΔG_{int}=-12.96 and ΔE_{int}=-26.38) base mispairs are close, but less than those for the G·C Watson-Crick DNA base pair (ΔG_{int}=-15.97 and ΔE_{int}=-29.28 kcal·mol^{-1} obtained at the MP2/6-311++G(2df,pd)//B3LYP/6-311++G(d,p) level of QM theory) [27]. Notably, the intermolecular H-bonds in the A*·G*$_{syn}$ and A·G*$_{syn}$ base mispairs consists more than half (72.6 and 63.8 %, respectively) of their electronic interaction energies, making a major contribution into their stabilization. The $\Sigma E_{HB}/|\Delta E_{int}|$ electronic-energy relationship can be considered as an invariant characteristic of all H-bonded nucleobase pairs [48,27,28,77], but not exceptionally of these base pairs.

A computational survey of the potential energy surface of the planar A*·G*$_{syn}$ base mispair reveals that it tautomerises into the planar A·G*$_{syn}$ base mispair through the slightly non-planar TS$_{A*·G*syn↔A·G*syn}$ (C_1 symmetry; ∠C6N6(A)C6C5(G$_{syn}$)=11.3º), stabilized by the N1-H-N7 covalent bridge and the N6H···O6 H-bond (12.62 kcal·mol^{-1}), with imaginary frequency ν_i=940.7*i* cm^{-1} (Tables 1 and 2; Fig. 2).

Transition state optimizations and intrinsic reaction coordinate (IRC) calculations for the A*·G*$_{syn}$↔A·G*$_{syn}$ tautomerisation *via* the DPT showed that there is additional TS at the IRC=-5.23 Bohr connecting the A*·G*$_{syn}$ base mispair (IRC=-8.75 Bohr) and the A^+·G^-_{syn} zwitterionic intermediate (IRC=-3.76 Bohr) (Figs. 2 and S1). We established that the localized A^+·G^-_{syn} zwitterionic intermediate is formal and cannot be considered as thermodynamically stable structure, since the value of the Gibbs free energy for the reverse barrier of the A*·G*$_{syn}$→A^+·G^-_{syn} transition are negative (-1.31 kcal·mol^{-1} calculated at the MP2/6-311++G(2df,pd)//B3LYP/6-311++G(d,p) level of theory). This means that there is a single TS (IRC=0.00 Bohr) for the A*·G*$_{syn}$↔A·G*$_{syn}$ tautomerisation *via* the DPT, that governs the reaction rate and therefore the A·G*$_{syn}$ product selectivity (Fig. 2).

By comparing the value of the zero-point energy (3.85 kcal·mol^{-1}) of the corresponding vibrational mode, which frequency becomes imaginary at the TS$_{A*·G*syn↔A·G*syn}$, with the electronic activation energy for the reverse reaction of the A*·G*$_{syn}$→A·G*$_{syn}$ DPT tautomerisation (2.08 kcal·mol^{-1}), it was revealed that the A·G*$_{syn}$ base mispair is dynamically unstable structure [27,28,85]. These

observations do not depend on the used QM levels of theory (for more details see Table 3). Moreover, the Gibbs free energy of activation for the reverse reaction of the $A^*{\cdot}G^*_{syn} \rightarrow A{\cdot}G^*_{syn}$ DPT tautomerisation is negligibly small (<*kT*). This results in extremely short lifetime of the $A{\cdot}G^*_{syn}$ base mispair that is equal to $2.17{\cdot}10^{-13}$ s (Table 3). Therefore any of the 6 low-frequency intermolecular vibrations [86] (9.9, 21.7, 63.2, 73.1, 103.7 and 111.5 cm^{-1}) can develop during this period of time, since their periods are noticeably greater than this time interval. This observation additionally indicates that the $A{\cdot}G^*_{syn}$ base mispair is not dynamically stable structure and the tautomeric status of the bases included in the $A^*{\cdot}G^*_{syn}$ base mispair does not change upon its spontaneous or forced dissociation by the DNA-polymerase machinery: in both cases this process occurs according to the scheme $A^*{\cdot}G^*_{syn} \xrightarrow{dissociation} A^*+G^*_{syn}$.

1.2. The $G{\cdot}A_{syn}$ base mispair

Lets now discuss the results of the investigation of the $G{\cdot}A_{syn}$ tautomerisation *via* the DPT into the $G^*{\cdot}A^*_{syn}$ base mispair, involving the G base in the *anti*- and the A base in the *syn*-orientations respectively to the sugar residue. The $G{\cdot}A_{syn}$ base mispair is non-planar (∠N1C6(G)C6C5(A_{syn})=-21.0º) structure with C_1 symmetry (Scheme 1). The electronic and Gibbs free energies of interaction between the G and A_{syn} bases in the $G{\cdot}A_{syn}$ base mispair constitute -17.00 and -2.80 $kcal{\cdot}mol^{-1}$, respectively.

Providing the calculations for the tautomerised $G^*{\cdot}A^*_{syn}$ base mispair and the TS of the $G{\cdot}A_{syn} \leftrightarrow G^*{\cdot}A^*_{syn}$ tautomerisation, it was revealed that there is no potential-energy minima associated with these structures. This means that the $G{\cdot}A_{syn}$ base mismatch, stabilized solely by the two N6H···O6 (5.08) and N1H···N7 (6.12 $kcal{\cdot}mol^{-1}$) H-bonds, which total energy makes 65.9 % of the electronic interaction energy, does not tautomerise at all and the tautomerised $G^*{\cdot}A^*_{syn}$ base mismatch almost instantly converts into the $G{\cdot}A_{syn}$ base mismatch without any barrier.

In order to clarify such important physico-chemical characteristic of the intermolecular H-bonds as their cooperativity, we have proposed a new approach based on the initial forced stretching of the N6H and N1H atomic groups – donors of the N6H···O6 and N1H···N7 H-bonds, respectively, with subsequent sequential fixation of their length and geometry optimization (Fig. 1). As a result, we found out that the N6H···O6 and N1H···N7 H-bonds are cooperative (Fig. 1).

We established that comparably small amount of the electronic deformational energy ΔE_{def} [16,17,21] should be applied for the $G{\cdot}A_{syn}$ base mispair (R(H-H)=10.399 Å; α_1=51.6º; α_2=38.5º) to aquire the glycosidic sizes of the A·T (R(H-H)=10.132 Å; α_1=54.2º; α_2=54.7º) (ΔE_{def}=3.00 $kcal{\cdot}mol^{-1}$) and G·C (R(H-H)=10.213 Å; α_1=53.1º; α_2=55.1º) (ΔE_{def}=3.61 $kcal{\cdot}mol^{-1}$ obtained at the MP2/6-

311++G(2df,pd)//B3LYP/6-311++G(d,p) level of theory) Watson-Crick DNA base pairs, that represents an imperative property for the structural isosterisity with Watson-Crick base pairs and faithful DNA replication (see ref. [21] and bibliography there). Notably, we revealed the formation of third C8H···N2 H-bond at the adjusting of the geometry of the G·A_{syn} base mispair to the sizes of the A·T (0.48 kcal·mol^{-1}) and G·C (0.41 kcal·mol^{-1}) Watson-Crick DNA base pairs.

2. Atomistic mechanism of the A*·G*$_{syn}$↔A·G*$_{syn}$ tautomerisation *via* the DPT

In order to penetrate into the essence of the energetical, structural, polar, electron-topological and NBO reorganizations of the A*·G*$_{syn}$ base mispair and intermolecular interactions stabilizing it along the IRC we have carried out 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,-1) BCPs of the intrapair H-bonds, the NBO charges of the hydrogen atoms involved in the tautomerisation, the glycosidic angles and the distance between the glycosidic hydrogens at each step along the IRC of the A*·G*$_{syn}$↔A·G*$_{syn}$ tautomerisation. In such a way we obtained the scanning or, in other words, sweeps or profiles [27,28,30,31,48]) of these characteristics *vs.* the IRC (Figs. S1-S6), revealing surprising physico-chemical details of the A*·G*$_{syn}$↔A·G*$_{syn}$ DPT tautomerisation.

Analysing the spans of the χ-like conjunctions on the $d_{O6H/HN6}$, $d_{N1H/HN7}$, ρ and Δρ profiles of the intermolecular H-bonds (Figs. S3a, S3b and S4b), we have arrived at the conclusion that the A*·G*$_{syn}$↔A·G*$_{syn}$ tautomerisation proceeds through the *asynchronous concerted mechanism*. Additionally, basing on the changes of the electron density and geometry of the intermolecular H-bonds along the IRC of the A*·G*$_{syn}$↔A·G*$_{syn}$ tautomerisation, we have allocated 9 key points, that characterise the course of this reaction [69,70,71] (Figs. 2, S3a, S3b and S4b).

Thus, the A*·G*$_{syn}$↔A·G*$_{syn}$ DPT tautomerisation is initiated by the sequential migration of the first proton at the O6 oxygen atom of the G*$_{syn}$ enol tautomer in the A*·G*$_{syn}$ base mispair (the key point 1; IRC=-8.75 Bohr) along the O6H···N6 H-bond in the direction of the N6 nitrogen atom of the A* imino tautomer. This transition passes through the key points 2 ($\Delta\rho_{N6\cdots H}$=0; IRC=-5.48 Bohr), 3 ($\rho_{N6\text{-}H}$=$\rho_{H\text{-}O6}$; IRC=-5.23 Bohr), that coincides with the aforementioned formal TS of the A*·G*$_{syn}$→A^+·G^-_{syn} transition, 4 ($\Delta\rho_{H\cdots O6}$=0; IRC=-5.09 Bohr) and is accompanied by the breakage of the H-O6 covalent bond and the formation of the new N6-H covalent bond. Consequently, the reaction is arrived at the A^+·G^-_{syn} non-stable intermediate. Then, the second mobile proton, localized at the $N1^+$ nitrogen atom of the A^+ protonated base starts out to move towards the $N7^-$ nitrogen atom of the G_{syn}^- deprotonated base

across the key points 5 ($\Delta\rho_{H\cdots N7}$=0; IRC=-0.19 Bohr), 6 - the $TS_{A^*\cdot G^*syn\leftrightarrow A\cdot G^*syn}$ (IRC=0.00 Bohr), 7 ($\rho_{N1\text{-}H}=\rho_{H\text{-}N7;}$ IRC=0.06 Bohr) and 8 ($\Delta\rho_{N1\cdots H}$=0; IRC=0.31 Bohr), eventually reaching the $A\cdot G^*_{syn}$ base mispair (the key point 9; IRC=3.73 Bohr). At this the rupture of the N1-H covalent bond and the formation of the H-N7 covalent bond are observed.

It is interesting to note that the values of the electron density ρ at the (3,-1) BCPs of the H-bonds in the nine key points involving stationary structures, usually treated as a measure of these interactions [75,76,77], range from 0.038 a.u. up to the 0.167 a.u. (Tables 1 and 2). The values of the Laplacian of the electron density Δρ at the (3,-1) BCPs are positive for all intrapair H-bonds and lie within a wide range from 0.000 a.u. up to the 0.160 a.u. (Tables 1 and 2), demonstrating that the H-bonds are attractive closed-shell ineractions. The value of the ellipticity ε varies in the range $(1.56\div5.73)\cdot10^{-2}$.

The classical geometrical criteria are satisfied for all canonical H-bonds in the nine key points along the IRC: the $d_{H\cdots N6/N1/N7}$ (1.261÷1.855 Å) and $d_{H\cdots O6}$ (1.285÷1.714 Å) distances are less than the sum of corresponding Bondi's [87] van der Waals radii (2.75 and 2.72 Å, respectively). The values of the angle of the H-bonding (161.0÷179.0º) are also the trends of the H-bonds (Tables 1 and 2). These observations indicate that the classical geometrical criteria are fulfilled along the IRC, even in the non-stationary points.

The established nine key points allow us to provide a partitioning of the pathway of the $A^*\cdot G^*_{syn}\rightarrow A\cdot G^*_{syn}$ DPT tautomerisation into three distinct regions: the reactant, TS and product regions, separated by the key points 2 (location of the reaction force minimum) and 8 (location of the reaction force maximum), in which the extrema of the first derivative of the electron energy with respect to the IRC - dE/dIRC – are reached [88] (Figs. 2 and S1).

It was established that *the reactant region*, where structural changes such as bond stretching, angle bending and mutual adjustment of the bases within the $A^*\cdot G^*_{syn}$ base pair take place preparing the complex for the DPT tautomerisation, is located between the key points 1 and 2 (-8.75÷-5.48 Bohr). Examination of the barrier heights shows that the electronic energy necessary to bring the donor and acceptor atoms as close as possible to each other and 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 2.31 $kcal\cdot mol^{-1}$, representing 51.7 % of the TS electronic energy. The DPT reaction actually occurs at *the TS region*, limited by the key points 2 and 8 (-5.48÷0.31 Bohr), where the A* and G^*_{syn} bases lose their chemical individuality and where the structural and electronic rearrangements occur accompanying bond breakages and formations. *The product region*, where the structural relaxation of the reaction complex starts to reach the final $A\cdot G^*_{syn}$ base mispair, is located

between the key points 8 and 9 (0.31÷3.73 Bohr), respectively. It was estimated, that comparably small amount of the electronic energy (2.47 kcal·mol^{-1}) releases during the transformation of the key point 8 to the key point 9, that is the energy difference between the key points 8 and 9, representing 55.2 % of the TS energy. Observed data evidence that almost equal amount of energy is spent as at the rebuilding and reorganization of the A* and G*$_{syn}$ bases within the A*·G*$_{syn}$ base pair actually before the initiation of the chemical reaction, so at the relaxation of the key point 8 to the final A·G*$_{syn}$ base mispair.

In the course of the investigation we established the electronic deformational energies of the A*·G*$_{syn}$ (R(H-H)=10.411 Å; α_1=50.3º; α_2=37.5º) and A·G*$_{syn}$ (R(H-H)=10.211 Å; α_1=53.0º; α_2=39.6º) base mispairs [16,17,21], which are inversely proportional to the probability of their incorporation into the DNA double helix: ΔE_{def} (A·T)=3.18 and ΔE_{def} (G·C)=3.72 for the A*·G*$_{syn}$ base mispair and ΔE_{def} (A·T)=4.68 and ΔE_{def} (G·C)=5.54 kcal·mol^{-1} for the A·G*$_{syn}$ base mispair obtained at the MP2/6-311++G(2df,pd)//B3LYP/6-311++G(d,p) level of QM theory. The presented above values indicate that these DNA base mispairs most closely resemble the geometry of the A·T Watson-Crick base pair and that the substantial part of the deformational energy is spent on the adjusting of the glycosidic angles.

3. Evolution of the selected physico-chemical characteristics of the DNA base mispairs and of the intermolecular specific contacts stabilizing them along the IRC of the A*·G*$_{syn}$→A·G*$_{syn}$ DPT tautomerisation

Notably, the electron energy profile of the A*·G*$_{syn}$↔A·G*$_{syn}$ DPT tautomerisation is double-well (Fig. S1a) and zero vibrational level of the vibrational mode in the A·G*$_{syn}$ base mispair (2690.4 cm^{-1}), which frequency becomes imaginary in the TS$_{A^*\cdot G^*syn\leftrightarrow A\cdot G^*syn}$ of the DPT tautomerisation, lies above the barrier of the DPT tautomerisation.

It is important to note that the O6H···N6 and N1H···N7 H-bonds in the A*·G*$_{syn}$ base mispair exist between the key points 1 and 2 and between the key points 1 and 5, reaching their maximal energies at the key points 2 and 5, for which $\Delta\rho$=0, respectively. The N6H···O6 and N7H···N1 H-bonds in the A·G*$_{syn}$ base mispair exist within the range from the key point 4 to the key point 9 and from the key point 8 to the key point 9, reaching their maximal energies at the key points 4 and 8, for which $\Delta\rho$=0, respectively (Fig. S3d).

Analysis of the dependencies of the H-bond energies on the IRC presented in Figure S3d allows us to reach a clear conclusion that the O6H···N6/N6H···O6 H-bonds and the N1H···N7/N7H···N1 H-bonds are cooperative with each other [27,28,29] and mutually strengthen each other in the A*·G*$_{syn}$ ($dE_{N1H\cdots N7}/dE_{O6H\cdots N6}$=0.92 at the IRC=-8.75 Bohr) and A·G*$_{syn}$ ($dE_{N6H\cdots O6}/dE_{N7H\cdots N1}$=0.21 at the IRC=3.73 Bohr) base mispairs, respectively.

The graphs of the electron density ρ and the Laplacian of the electron density Δρ at the (3,-1) BCPs of the N6-H and H-O6 bonds intersect exactly at the key point 3, while of the N1-H and H-N7 bonds – at the key point 7 (Figs. S3a and S3b). Moreover, the shapes of these profiles are similar to each other, indicating a strong correlation between them [77]. In contrast, the graph of the ellipticity ε at the (3,-1) BCP of the N7-H bond crosses with the graphs for the N1-H and N6-H bonds between the key points 5 and 6 (Fig. S3c) and with the O6-H bond – nearby the key point 8, indicating the extreme sensitivity of the intermolecular bonds to the dynamical behavior of the base pair and the modulation of their energies by the low-frequency intermolecular vibrations of the base pair that tautomerises [86]. The values of the ρ (0.038÷0.312 a.u.), Δρ (-2.003÷0.161 a.u.) and ε (0.012÷0.057) parameters lie within a wide range of values and are in line with the results, presented in our recent works [27,28,30,31].

We have established significant alterations of the absolute value of the dipole moment μ of the studied base pairs within the range of values 6.42÷10.79 D along the IRC of the A*·G*$_{syn}$↔A·G*$_{syn}$ tautomerisation (Fig. S2). It is evident from the profile of the dipole moment that the tautomerised A·G*$_{syn}$ base mispair is more polar than the starting complex – the A*·G*$_{syn}$ base mispair and a significant increasing of the polarity of the system is observed in the vicinity of the TS$_{A*·G*syn↔A·G*syn}$, that can be explained by the zwitterionic structure of the base pair.

The obtained sweeps of the NBO charges [89] of the hydrogen atoms localized between the N6 and O6, N1 and N7 atoms along the IRC of the A*·G*$_{syn}$↔A·G*$_{syn}$ tautomerisation *via* the DPT are presented in Fig. S5. Interestingly, the graph for the NBO charge of the hydrogen atom involved in the O6H···N6 H-bond lies above the graph for the N1H···N7 H-bond at the reactant region and then, changing places at the TS region, the graph for the N6H···O6 H-bond starts to be settled down under the graph for the N7H···N1 H-bond at the product region. The sharp peaks on the sweeps of the NBO charges are observed nearby the key points 4 and 5-8, reflecting the proton movements (Fig. S5).

We revealed the compression, or so-called "breathing", of the starting A*·G*$_{syn}$ base mispair throughout the tautomerisation process (Figs. 2, S4a and S6a). It is believed that this phenomenon occurs due to the contraction of the distance between the O6 oxygen and N6 nitrogen atoms at the key point 3 (by 0.289 Å) and between the N1 and N7 nitrogen atoms at the key point 6 (by 0.304 Å) (Fig. S4a). Moreover, the H-bond angles and glycosidic parameters vary at the compression of the A*·G*$_{syn}$ base mispair along the IRC of the A*·G*$_{syn}$↔A·G*$_{syn}$ tautomerisation, that to a large extent is observed at the TS region: ∠O6HN6 (161.1÷173.5º); ∠N1HN7 (172.9÷179.0º) (Fig. S4a); R(H-H) (9.966÷10.399 Å); α_1 (∠N9H(A)H(G$_{syn}$)) (50.4÷53.8º) and α_2 (∠N9H(G$_{syn}$)H(A)) (36.4÷40.2º) (Fig. S6).

Concluding Remarks

B3LYP and MP2 levels of QM theory, QTAIM analysis, NBO theory and the methodology of the sweeps of the energetical, electron-topological, geometrical, polar and NBO parameters were used to study the structure, barrier heights, thermodynamics, kinetics, electron-topological, polar and NBO properties of the $A^*{\cdot}G^*_{syn} \leftrightarrow A{\cdot}G^*_{syn}$ tautomerisation *via* the DPT.

It was found that the $A^*{\cdot}G^*_{syn}$ DNA base mispair is stabilized by the cooperative O6H···N6 (9.71) and N1H···N7 (6.99) H-bonds, while the $A{\cdot}G^*_{syn}$ base mispair – by the cooperative N6H···O6 (7.23) and N7H···N1 (9.59 kcal·mol^{-1}) intermolecular H-bonds.

By elaborating a novel approach to the definition of the cooperativity of the intermolecular H-bonds in the H-bonded complexes, which in principle can not tautomerise, it was found that the N6H···O6 (5.08) and N1H···N7 (6.12 kcal·mol^{-1}) intermolecular H-bonds in the $G{\cdot}A_{syn}$ base mispair are cooperative.

The 9 key points were detected and completely investigated along the IRC of the $A^*{\cdot}G^*_{syn} \leftrightarrow A{\cdot}G^*_{syn}$ tautomerisation *via* the DPT.

We have established for the first time that the $G{\cdot}A_{syn}$ and $A^*{\cdot}G^*_{syn}$ DNA base mispairs causing spontaneous transversions radically differ from each other in their ability to tautomerise through the DPT. The $A^*{\cdot}G^*_{syn}$ base mispair (C_s) tautomerises into the $A{\cdot}G^*_{syn}$ base mispair (C_s) *via* the asynchronous concerted DPT through the non-planar transition state (C_1), whereas the $G{\cdot}A_{syn}$ base mismatch (C_1) does not tautomerise at all, since there is no local minimum corresponding to the tautomerised $G^*{\cdot}A^*_{syn}$ base mispair on its electronic (potential) energy surface. The $A^*{\cdot}G^*_{syn}$ base mispair tautomerises quite easily: the Gibbs free energy of activation for the forward reaction of the $A^*{\cdot}G^*_{syn} \rightarrow A{\cdot}G^*_{syn}$ DPT tautomerisation constitutes only $4{\cdot}kT$ at room temperature.

It was established that whereas all three investigated $G{\cdot}A_{syn}$, $A^*{\cdot}G^*_{syn}$ and $A{\cdot}G^*_{syn}$ DNA base mispairs are thermodynamically stable structures, since the values of their Gibbs free energies of interaction are less than zero, the $A^*{\cdot}G^*_{syn}$ base mispair is dynamically unstable H-bonded complex with an extremely low lifetime ($2.17{\cdot}10^{-13}$ s) and any of its 6 low-frequency intermolecular vibrations can develop during this span of time. Consequently, the dissociation of both the $G{\cdot}A_{syn}$ and $A^*{\cdot}G^*_{syn}$ DNA base mispairs occurs spontaneously or forcedly by the DNA-polymerase machinery, with the preservation of the tautomeric status of the bases, that form them.

Obtained results allow us to make the following biologically important assumptions. The dissociation of the $G{\cdot}A_{syn}$ and $A^*{\cdot}G^*_{syn}$ base mispairs, responsible for the asymmetric transversions, is carried out by the DNA polymerase machinery without changing of the tautomeric status of the bases that form them, i.e. $G{\cdot}A_{syn} \rightarrow G + A_{syn}$ and $A^*{\cdot}G^*_{syn} \rightarrow A^* + G^*_{syn}$. In the last case it means that the

A·G*$_{syn}$ mismatch "escapes from the hands" of the replication machinery due to its immediate transformation into the A*·G*$_{syn}$ base mispair. This finding is important, in our point of view, in terms of the theory of the spontaneous point mutations, namely - of the process of their emergence and consolidation in the DNA biosynthesis in a living cell.

Electronic Supplementary Material. 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 electron density, the Laplacian of the electron density, the ellipticity and the energy at the (3,-1) BCPs of the intrapair H-bonds, the distances and the angle of the intermolecular H-bonds, the NBO charges of the hydrogen atoms involved in the intermolecular interactions and the glycosidic parameters (angles and distance between the glycosidic hydrogens) of the base mispair along the IRC of the A*·G*$_{syn}$↔A·G*$_{syn}$ tautomerisation *via* the DPT obtained at the B3LYP/6-311++G(d,p) level of QM theory in vacuum are available free of charge *via* the Internet at http://rsc.org.

Acknowledgments. 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 SFFR of Ukraine (project № GP/F56/074). 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

1. Kornberg A., Baker T. A. *DNA replication.* New York, NY: W. H. Freeman, 1992.
2. Friedberg E.C., Walker G.C., Siede W., Wood R.D., Schultz R.A., Ellenberger T. *DNA repair and mutagenesis*. Washington D.C.: ASM Press, 2006.
3. *DNA Repair* (Ed. by Kruman I.). Rijeka: In Tech Open Access, 2011.
4. Johnson K.A. DNA replication caught in the act, *Nature*, **487**, 177-178.
5. Nakamura N., Zhao Y., Yamagata Y., Hua Y., Yang W., Watching DNA polymerase η make a phosphodiester bond, *Nature*, 2012, **487**, 196-201.
6. Kirmizialtin S., Nguyen V., Johnson K.A., Elber R. How conformational dynamics of DNA polymerase select correct substrates: experiments and simulations, *Structure*, 2012, **20**, 618-627.
7. Freudenthal B.D., Beard W.A., Shock D.D., Wilson S.H. Observing a DNA polymerase choose right from wrong, *Cell*, 2013, **154**, 157-168.
8. Brovarets' O.O., Kolomiets' I.M., Hovorun D.M. Elementary molecular mechanisms of the spontaneous point mutations in DNA: a novel quantum-chemical insight into the classical understanding. In: Tada T. (ed.), Quantum chemistry – molecules for innovations. Rijeka: In Tech Open Access, 2012, pp 59–102.
9. Goodman M. F. Hydrogen bonding revisited: geometric selection as a principal determinant of DNA replication fidelity, *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 10493-10495.
10. Kool E. T. Active site tightness and substrate fit in DNA replication, *Annu. Rev. Biochem.*, 2002, **71**, 191–219.
11. Bebenek K., Pedersen L.C., Kunkel T.A. Replication infidelity *via* a mismatch with Watson-Crick geometry, *Proc. Natl. Acad. Sci. USA*, 2011, **108**, 1862-1867.
12. Wang W., Hellinga H.W., Beese L.S. Structural evidence for the rare tautomer hypothesis of spontaneous mutagenesis, *Proc. Natl. Acad. Sci. USA*, 2011, **124**, 17644–17648.
13. Poltev V.I., Bruskov V.I. Molecular mechanisms of spontaneous transversions and transitions, *Mol. Biol.,* 1977, **11,** 661–670.
14. Poltev V.I., Shulyupina N.V., Bruskov V.I. Fidelity of nucleic acid biosynthesis. Comparison of computer modeling results with experimental data, *Mol. Biol., **32***, 1998, 233-240.
15. Kretulskie A.M., Spratt T.E. Structure of purine-purine mispairs during misincorporation and extension by *Escherichia coli* DNA polymerase I, *Biochemistry*, 2006, **45**, 3740-3746.
16. Brovarets' O. O., Hovorun D. M. Quantum-chemical investigation of the elementary molecular mechanisms of pyrimidine-purine transversions, *Ukr. Biokhim. Zh.,* 2010, **82**, 55-60.
17. Kochina O. S., Yurenko Ye. P., Hovorun D. M. Does non-empirical quantum chemistry allow understanding nature of purine-purine mismatches formation by high fidelity DNA polymerases? *Biopolym. Cell*, 2007, **23**, 167-171.
18. Brown T., Hunter W. H., Kneale G., Kennard O. Molecular structure of the G·A base pair in DNA and its implications for the mechanism of transversion mutations, *Proc. Natl. Acad. Sci. USA,* 1986, **83**, 2402-2406.
19. Schlick T., Arora K., Beard W.A., Wilson S.H. Perspective: pre-chemistry conformational changes in DNA polymerase mechanisms, *Theor. Chem. Acc.*, 2012, **131**, 1287.
20. Prasad B. R., Kamerlin S. C. L., Florian J., Warshel A. Prechemistry barriers and checkpoints do not contribute to fidelity and catalysis as long as they are not rate limiting, *Theor. Chem. Acc.*, 2012, **131**, 1288.

21. Brovarets' O.O., Zhurakivsky R.O., Hovorun D.M. 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, *J. Comput. Chem.*, 2014, **35**, 451-466.
22. Brovarets' O. O. Physico-chemical nature of the spontaneous and induced by the mutagens transitions and transversions (PhD thesis). Taras Shevchenko National University of Kyiv: Kyiv, 2010.
23. Kosenkov D., Kholod Y., Gorb L., Shishkin O., Hovorun D.M., Mons M., Leszczynski J. *Ab initio* kinetic simulation of gas-phase experiments: tautomerization of cytosine and guanine, *J. Phys. Chem. B*, 2009, **113**, 6140-6150.
24. Furmanchuk A., Isayev O., Gorb L., Shishkin O.V., Hovorun D.M., Leszczynski J. Novel view on the mechanism of water-assisted proton transfer in the DNA bases: bulk water hydration, *Phys. Chem. Chem. Phys.*, 2011, **13**, 4311-4317.
25. Platonov M.O., Samijlenko S.P., Sudakov O.O., Kondratyuk I.V., Hovorun D.M. To what extent can methyl derivatives be regarded as stabilized tautomers of xanthine? *Spectrochim. Acta A Mol. Biomol. Spectrosc.,* 2005, **62,** 112–114.
26. Löwdin P.-O. Proton tunneling in DNA and its biological implications, *Rev. Mod. Phys.*, 1963, **35**, 724-732.
27. Brovarets' O.O., Hovorun D.M. Can tautomerisation of the A·T Watson-Crick base pair *via* double proton transfer provoke point mutations during DNA replication? A comprehensive QM and QTAIM analysis, *J. Biomol. Struct. Dynam.,* 2014, **32**, 127-154.
28. Brovarets' O.O., Hovorun D.M. 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, *J. Biomol. Struct. Dynam.,* 2013. DOI: 10.1080/07391102.2013.822829.
29. Mishchuk Ya.R., Potyagaylo A.L., Hovorun D.M. Structure and dynamics of 6-azacytidine by MNDO/H quantum-chemical method, *J. Mol. Struct.*, 2000, **552**, 283-289.
30. Brovarets' O.O., Zhurakivsky R.O., Hovorun D.M. The physico-chemical "anatomy" of the tautomerisation through the DPT of the biologically important pairs of hypoxanthine with DNA bases: QM and QTAIM perspectives, *J. Mol. Model.*, 2013, **19**, 4119-4137.
31. Brovarets' O.O., Zhurakivsky R.O., Hovorun D.M. 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, *J. Mol. Model.,* 2013, **19**, 4223-4237.
32. Frisch M.J., Trucks G.W., Schlegel H.B., Scuseria G.E., Robb M.A., Cheeseman J.R., Pople J.A. et. al. (2010) GAUSSIAN 09 (Revision B.01), Gaussian Inc. Wallingford, C.T.
33. Parr R.G., Yang W. *Density-functional theory of atoms and molecules.* Oxford University Press: Oxford, 1989.
34. Tirado-Rives J., Jorgensen W.L. Performance of B3LYP Density Functional Methods for a large set of organic molecules, *J. Chem. Theory. Comput.*, 2008, **4**, 297–306.
35. Becke A.D. Density-functional exchange-energy approximation with correct asymptotic behavior, *Phys. Rev. A*, 1988, **38**, 3098-3100.
36. Lee C., Yang W., Parr R.G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, *Phys. Rev. B*, 1988, **37**, 785-789.
37. Brovarets' O. O., Hovorun D. M. Quantum-chemical investigation of tautomerization ways of Watson-Crick DNA base pair guanine-cytosine, *Ukr. Biokhim. Zh.,* 2010, **82**, 55-60.
38. Brovarets' O.O., Hovorun D.M. How stable are the mutagenic tautomers of DNA bases? *Biopolym. Cell*, 2010, **26**, 72–76.

39. Brovarets' O.O., Hovorun D.M. Stability of mutagenic tautomers of uracil and its halogen derivatives: The results of quantum-mechanical investigation, *Biopolym. Cell*, 2010, **26**, 295–298.
40. Brovarets' O.O., Hovorun D.M. Intramolecular tautomerization and the conformational variability of some classical mutagens – cytosine derivatives: Quantum chemical study, *Biopolym. Cell,* 2011, **27***,* 221–230.
41. Hovorun D.M., Gorb L., Leszczynski J. From the nonplanarity of the amino group to the structural nonrigidity of the molecule: A post-Hartree–Fock *ab initio* study of 2-aminoimidazole, *Int. J. Quantum. Chem.*, 1999, **75**, 245–253.
42. Pelmenschikov A., Hovorun D.M., Shishkin O.V., Leszczynski J. A density functional theory study of vibrational coupling between ribose and base rings of nucleic acids with ribosyl guanosine as a model system, *J Chem. Phys.,* 2000, **113**, 5986–5990.
43. Samijlenko S.P., Krechkivska O.M., Kosach D.A., Hovorun D.M. Transitions to high tautomeric states can be induced in adenine by interactions with carboxylate and sodium ions: DFT calculation data, *J. Mol. Struct.*, 2004, **708**, 97-104.
44. Brovarets' O.O., Zhurakivsky R.O., Hovorun D.M. Is there adequate ionization mechanism of the spontaneous transitions? Quantum-chemical investigation, *Biopolym. Cell*, 2010, **26**, 398–405.
45. Brovarets' O.O., Hovorun D.M. IR Vibrational spectra of H-bonded complexes of adenine, 2-aminopurine and 2-aminopurine^{+} with cytosine and thymine: Quantum-chemical study, *Opt. Spectrosc.*, 2011, **111,** 750–757.
46. Samijlenko S.P., Yurenko Y.P., Stepanyugin A.V., Hovorun D.M. Tautomeric equilibrium of uracil and thymine in model protein–nucleic acid contacts. Spectroscopic and quantum chemical approach, *J. Phys. Chem. B,* 2011, **114**, 1454-1461.
47. Brovarets' O.O., Hovorun D.M. Prototropic tautomerism and basic molecular principles of hypoxanthine mutagenicity: An exhaustive quantum-chemical analysis, *J. Biomol. Struct. Dynam.*, 2013, **31**, 913-936.
48. Brovarets' O. O., Hovorun D. M. The physicochemical essence of the purine·pyrimidine transition mismatches with Watson-Crick geometry in DNA: A·C* *versa* A*·C. A QM and QTAIM atomistic understanding, *J. Biomol. Struct. Dynam.,* 2013. DOI: 10.1080/07391102.2013.852133.
49. Matta C.F. How dependent are molecular and atomic properties on the electronic structure method? Comparison of Hartree-Fock, DFT, and MP2 on a biologically relevant set of molecules, *J. Comput. Chem.*, 2010, **31**, 1297–1311.
50. Lozynski M., Rusinska-Roszak D., Mack H.-G. Hydrogen bonding and density functional calculations: The B3LYP approach as the shortest way to MP2 results, *J. Phys. Chem. A*, 1998, **102,** 2899–2903.
51. Brovarets' O.O., Zhurakivsky R.O., Hovorun D.M. The physico-chemical mechanism of the tautomerisation *via* the DPT of the long Hyp*·Hyp Watson-Crick base pair containing rare tautomer: a QM and QTAIM detailed look, *Chem. Phys. Lett.*, 2013, **578**, 126-132.
52. Brovarets' O.O., Zhurakivsky R.O., Hovorun D.M. Structural, energetic and tautomeric properties of the T·T*/T*·T DNA mismatch involving mutagenic tautomer of thymine: A QM and QTAIM insight, *Chem. Phys. Lett.*, 2014, **592**, 247-255.
53. Frisch M.J., Head-Gordon M., Pople J.A. Semi-direct algorithms for the MP2 energy and gradient, *Chem Phys. Lett.*, 1990, **166,** 281-289.
54. Frisch M.J., Pople J.A., Binkley J.S. Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets, *J. Chem. Phys.,* 1984, **80**, 3265–3269.
55. Hariharan P.C., Pople J.A. The influence of polarization functions on molecular orbital hydrogenation energies, *Theor. Chem. Accounts. Theor. Comput. Model.,* 1973, **28**, 213–222.

56. Krishnan R., Binkley J.S., Seeger R., Pople J.A. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions, *J. Chem. Phys.,* 1980, **72**, 650–654.
57. Dunning Jr. T.H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, *J. Chem. Phys.,* 1989, **90**, 1007–1023.
58. Kendall R.A., Dunning Jr.T.H., Harrison R.J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions, *J. Chem. Phys.,* 1992, **96**, 6796–6806.
59. Peng C., Schlegel H.B. Combining synchronous transit and quasi-Newton methods to find transition states, *Isr. J. Chem.*, 1993, **33**, 449–454.
60. Peng C., Ayala P.Y., Schlegel H.B., Frisch M.J. Using redundant internal coordinates to optimize equilibrium geometries and transition states, *J. Comput. Chem.*, 1996, **17**, 49–56.
61. Hratchian H.P., Schlegel H.B. Accurate reaction paths using a Hessian based predictor-corrector integrator, *J. Chem. Phys.,* 2004, **120**, 9918-9924.
62. Hratchian H.P., Schlegel H.B. Finding minima, transition states, and following reaction pathways on *ab initio* potential energy surfaces. In: Dykstra C.E., Frenking G., Kim K.S., Scuseria G. (eds.) Theory and applications of computational chemistry: The first 40 years, Elsevier, Amsterdam, 2005, pp. 195-249.
63. Hratchian H.P., Schlegel H.B. Using Hessian updating to increase the efficiency of a Hessian based predictor-corrector reaction path following method, *J. Chem. Theory Comput.*, 2005, **1,** 61-69.
64. Boys S.F., Bernardi F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors, *Mol. Phys.*, 1970, **19**, 553-566.
65. Gutowski M., Van Lenthe J.H., Verbeek J., Van Duijneveldt F.B., Chalasinski G. The basis set superposition error in correlated electronic structure calculations, *Chem. Phys. Lett.,* 1986, **124**, 370–375.
66. Sordo J.A., Chin S., Sordo T.L. On the counterpoise correction for the basis set superposition error in large systems, *Theor. Chim. Acta*, 1988, **74**, 101-110.
67. Sordo J.A. On the use of the Boys–Bernardi function counterpoise procedure to correct barrier heights for basis set superposition error*, J. Mol. Struct. (THEOCHEM)*, 2001, **537**, 245-251.
68. Atkins P.W. *Physical chemistry,* Oxford University Press, Oxford, 1998.
69. Brovarets' O.O., Hovorun D.M. Atomistic understanding of the C·T mismatched DNA base pair tautomerization *via* the DPT: QM and QTAIM computational approaches, *J. Comput. Chem.*, 2013, **34**, 2577-2590.
70. Brovarets' O.O., Hovorun D.M. Atomistic nature of the DPT tautomerisation of the biologically important C•C* DNA base mispair containing amino and imino tautomers of the cytosine: A QM and QTAIM approach, *Phys. Chem. Chem. Phys.*, 2013, **15**, 20091-20104.
71. Brovarets' O.O., Zhurakivsky R.O., Hovorun D.M. Does the tautomeric status of the adenine bases change under the dissociation of the A*•A_{syn} Topal-Fresco DNA mismatch? A combined QM and QTAIM atomistic insight, *Phys. Chem. Chem. Phys.*, 2014, **16**, 3715-3725.
72. Wigner E. Über das Überschreiten von Potentialschwellen bei chemischen Reaktionen [Crossing of potential thresholds in chemical reactions], *Zeit. Physik. Chem.* B, 1932, **19**, 203–216.
73. Bader R.F.W. *Atoms in molecules: A quantum theory,* Oxford University Press, Oxford, 1990.
74. Keith T.A. (2011) AIMAll (Version 11.12.19) Retrieved from http://aim.tkgristmill.com.
75. Ponomareva A.G., Yurenko Y.P., Zhurakivsky R.O., van Mourik T., Hovorun D.M. Complete conformational space of the potential HIV-1 reverse transcriptase inhibitors d4U and d4C. A quantum chemical study, *Phys. Chem. Chem. Phys.*, 2012, **14,** 6787-6795.

76. Ponomareva A.G., Yurenko Y.P., Zhurakivsky R.O., van Mourik T., Hovorun D.M. Structural and energetic properties of the potential HIV-1 reverse transcriptase inhibitors d4A and d4G: a comprehensive theoretical investigation, *J. Biomol. Struct. Dynam.,* 2013. DOI: 10.1080/07391102.2013.789401.
77. Brovarets' O.O., Yurenko Y.P., Hovorun D.M. Intermolecular CH…O/N H-bonds in the biologically important pairs of natural nucleobases: a thorough quantum-chemical study, *J. Biomol. Struct. Dynam.,* 2013 DOI: 10.1080/07391102.2013.799439.
78. Iogansen A.V. Direct proportionality of the hydrogen bonding energy and the intensification of the stretching *ν*(XH) vibration in infrared spectra, *Spectrochim. Acta A Mol. Biomol. Spectrosc.,* 1999, **55**, 1585–1612.
79. Nikolaienko T.Y., Bulavin L.A., Hovorun D.M. (2012). Bridging QTAIM with vibrational spectroscopy: The energy of intramolecular hydrogen bonds in DNA-related biomolecules, *Phys. Chem. Chem. Phys.,* 2012, **14**, 7441–7447.
80. Espinosa E., Molins E., Lecomte C. Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities, *Chem. Phys. Lett.*, 1998, **285**, 170–173.
81. Mata I., Alkorta I., Espinosa E., Molins E. Relationships between interaction energy, intermolecular distance and electron density properties in hydrogen bonded complexes under external electric fields, *Chem Phys. Lett.,* 2011, **507**, 185–189.
82. Saenger W. *Principles of nucleic acid structure*, Springer, New York, 1984.
83. Govorun D.N., Danchuk V.D., Mishchuk Ya.R., Kondratyuk I.V., Radomsky N.F., Zheltovsky N.V. AM1 calculation of the nucleic acid bases structure and vibrational spectra, *J. Mol. Struct.*, 1992, **267**, 99–103.
84. Nikolaienko T.Y., Bulavin L.A., Hovorun D.M. How flexible are DNA constituents? The quantum-mechanical study, *J. Biomol. Struct. Dynam.*, 2011, **29**, 563-575.
85. Brovarets' O.O., Yurenko Y.P., Dubey I.Ya., Hovorun D.M. Can DNA-binding proteins of replisome tautomerize nucleotide bases? *Ab initio* model study, *J. Biomol. Struct. Dynam.*, 2012, **29**, 1101–1109.
86. Shishkin O.V., Pelmenschikov. A., Hovorun D.M., Leszczynski J. Theoretical analysis of low-lying vibrational modes of free canonical 2′-deoxyribonucleosides, *Chem. Phys.*, 2000, **260**, 317–325.
87. Bondi A.J. Van der Waals volumes and radii, *J. Phys. Chem.,* 1964, **68**, 441-451.
88. Politzer P., Murray J.S., Jaque P. Perspectives on the reaction force constant, *J. Mol. Model.*, 2013, **19**, 4111-4118.
89. Weinhold F., Landis C. *Valency and bonding. A natural bond orbital donor-acceptor perspective.* Cambridge: Cambridge University Press, 2005.

Table 1 Electron-topological, structural, vibrational and energetical characteristics of the intermolecular H-bonds revealed in the A*·G*$_{syn}$, A·G*$_{syn}$, G·A$_{syn}$ and TS$_{A*·G*syn↔A·G*syn}$ obtained at the B3LYP/6-311++G(d,p) level of theory in vacuum

Complex	AH···B H-bond	ρ[a]	$\Delta\rho$[b]	$100\cdot\varepsilon$[c]	$d_{A\cdots B}$[d]	$d_{H\cdots B}$[e]	∠AH···B[f]	Δd_{AH}[g]	$\Delta\nu$[h]	E_{HB}[i]
A*·G*$_{syn}$	O6H···N6	0.057	0.103	4.56	2.645	1.664	161.0	0.048	905.4	9.71
	N1H···N7	0.038	0.094	5.72	2.894	1.855	179.5	0.028	489.3	6.99
A·G*$_{syn}$	N6H···O6	0.042	0.135	2.64	2.745	1.714	172.4	0.030	520.3	7.23
	N7H···N1	0.051	0.093	5.73	2.801	1.739	175.2	0.051	884.3	9.59
TS$_{A*·G*syn↔A·G*syn}$	N6H···O6	0.065	0.155	2.33	2.614	1.553	172.4	-	-	12.62*
G·A$_{syn}$	N6H···O6	0.029	0.104	2.85	2.886	1.874	169.9	0.016	276.9	5.08
	N1H···N7	0.032	0.087	5.50	2.958	1.926	175.6	0.022	383.8	6.12

[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 (H-bond donor) and B (H-bond acceptor) electronegative atoms of the H-bonds, Å
[e]The distance between the H and B atoms of the H-bonds, Å
[f]The H-bond angle, degree
[g]The elongation of the H-bond donating group AH upon H-bonding, Å
[h]The redshift of the stretching vibrational mode ν(AH), cm^{-1}
[i]The energy of the H-bonds, calculated by Iogansen's [78] and Nikolaienko-Bulavin-Hovorun (marked with an asterisk) [79] formulas, kcal mol^{-1}

Table 2 Electron-topological and structural characteristics of the intermolecular bonds revealed in the 9 key points and the polarity of the latters along the IRC of the A*·G*$_{syn}$↔A·G*$_{syn}$ DPT tautomerisation obtained at the B3LYP/6-311++G(d,p) level of theory in vacuum

Complex	AH···B H-bond/ A-H-B covalent bond	ρ	$\Delta\rho$	$100\cdot\varepsilon$	$d_{A\cdots B}$	$d_{H\cdots B}$	∠AH···B	μ
Key point 1 (-8.75 Bohr):	O6H···N6	0.057	0.103	4.56	2.645	1.664	161.0	6.43
A*·G*$_{syn}$	N1H···N7	0.038	0.094	5.72	2.894	1.855	179.5	
Key point 2 (-5.48 Bohr):	O6H···N6	0.117	0.000	3.66	2.464	1.384	164.1	7.37
$\Delta\rho_{N6\cdots H}$=0	N1H···N7	0.052	0.103	5.04	2.776	1.723	176.6	
Key point 3 (-5.23 Bohr):	O6-H-N6	0.162	-0.275	3.29	2.461	1.261	166.0	8.63
ρ_{N6-H}=ρ_{H-O6}	N1H···N7	0.053	0.099	4.94	2.773	1.713	176.7	
Key point 4 (-5.09 Bohr):	N6H···O6	0.135	0.000	1.56	2.465	1.285	166.4	9.39
$\Delta\rho_{H\cdots O6}$=0	N1H···N7	0.054	0.097	4.88	2.771	1.707	176.8	
Key point 5 (-0.19 Bohr):	N6H···O6	0.067	0.152	2.31	2.610	1.542	172.4	10.23
$\Delta\rho_{H\cdots N7}$=0	N1H···N7	0.112	0.000	3.50	2.588	1.418	173.9	
Key point 6 (0.00 Bohr):	N6H···O6	0.065	0.155	2.33	2.614	1.553	172.4	9.49
$TS_{A*\cdot G*syn\leftrightarrow A\cdot G*syn}$								
Key point 7 (0.06 Bohr):	N6H···O6	0.065	0.156	2.34	2.615	1.556	172.4	9.24
ρ_{N1-H}=ρ_{H-N7}	N1-H-N7	0.154	-0.217	2.98	2.584	1.293	174.7	
Key point 8 (0.31 Bohr):	N6H···O6	0.062	0.160	2.36	2.618	1.568	172.5	8.39
$\Delta\rho_{N1\cdots H}$=0	N7H···N1	0.111	0.000	4.12	2.588	1.419	175.0	
Key point 9 (3.73 Bohr):	N6H···O6	0.042	0.135	2.64	2.745	1.714	172.4	7.75
A·G*$_{syn}$	N7H···N1	0.051	0.093	5.73	2.801	1.739	175.2	

Notes: For footnote definitions see Table 1. μ – the dipole moment of the complex, D.

Table 3 Energetic and kinetic characteristics of the A*·G*$_{syn}$↔A·G*$_{syn}$ tautomerisation *via* the DPT *in vacuo* (T=298.15 K) obtained at the different levels of QM theory for the geometry calculated at the B3LYP/6-311++G(d,p) level of QM theory

Level of QM theory	ΔG^a	ΔE^b	$\Delta\Delta G_{TS}{}^c$	$\Delta\Delta E_{TS}{}^d$	$\Delta\Delta G^e$	$\Delta\Delta E^f$		τ^g
						kcal·mol^{-1}	cm^{-1}	
MP2/6-311++G(2df,pd)	1.83	2.14	2.03	4.21	0.20	2.08	726.4	$1.25\cdot10^{-13}$
MP2/6-311++G(3df,2pd)	1.71	2.02	2.16	4.35	0.45	2.33	815.5	$1.92\cdot10^{-13}$
MP2/cc-pVTZ	2.67	2.98	3.07	5.25	0.40	2.27	795.2	$1.75\cdot10^{-13}$
MP2/cc-pVQZ	1.89	2.20	2.42	4.60	0.52	2.40	840.2	$2.17\cdot10^{-13}$

[a]The Gibbs free energy of the A·G*$_{syn}$ base mispair relatively to the A*·G*$_{syn}$ base mispair ($\Delta G_{A^*\cdot G^*syn}$=0.00), kcal·mol^{-1}

[b]The electronic energy of the A·G*$_{syn}$ base mispair relatively to the A*·G*$_{syn}$ base mispair ($\Delta E_{A^*\cdot G^*syn}$=0.00), kcal·mol^{-1}

[c]The Gibbs free energy of activation for the forward reaction of the A*·G*$_{syn}$→A·G*$_{syn}$ tautomerisation, kcal·mol^{-1}

[d]The activation electronic energy for the forward reaction of the A*·G*$_{syn}$→A·G*$_{syn}$ tautomerisation, kcal·mol^{-1}

[e]The Gibbs free energy of activation for the reverse reaction of the A*·G*$_{syn}$→A·G*$_{syn}$ tautomerisation, kcal·mol^{-1}

[f]The activation electronic energy for the reverse reaction of the A*·G*$_{syn}$→A·G*$_{syn}$ tautomerisation

[g]The lifetime of the A·G*$_{syn}$ base mispair, s

The frequency of the vibrational mode in the A·G*$_{syn}$ base mispair, which becomes imaginary in the $TS_{A^*\cdot G^*syn\leftrightarrow A\cdot G^*syn}$ of the A*·G*$_{syn}$↔A·G*$_{syn}$ DPT tautomerisation, is equal to 2690.4 cm^{-1} and the zero-point vibrational energy associated with this normal mode is equal to 3.85 kcal·mol^{-1} or 1345.2 cm^{-1} obtained at the B3LYP/6-311++G(d,p) level of QM theory.

The Gibbs free energy of the G·A$_{syn}$ base mispair relatively to the A*·G*$_{syn}$ base mispair ($\Delta G_{A^*\cdot G^*syn}$=$\Delta E_{A^*\cdot G^*syn}$=0.00) is equal to $\Delta G_{G\cdot Asyn}$=-8.91 kcal·mol^{-1} and the electronic energy – is equal to $\Delta E_{G\cdot Asyn}$=-9.70 kcal·mol^{-1} obtained at the MP2/6-311++G(2df,pd)//B3LYP/6-311++G(d,p) level of QM theory *in vacuo*.

Graphical Abstract.

The $A^*{\cdot}G^*_{syn} \leftrightarrow A{\cdot}G^*_{syn}$ DPT tautomerisation proceeds through the asynchronous concerted mechanism; the $G{\cdot}A_{syn}$ DNA base mispair does not tautomerise at all.