O.O. BROVARETS,[1,2] D.M. HOVORUN[1,2]

[1] Institute of Molecular Biology and Genetics, Nat. Acad. of Sci. of Ukraine
(150, Academician Zabolotnyi Str., Kyiv 03680, Ukraine; e-mail: dhovorun@imbg.org.ua)
[2] Institute of High Technologies, Taras Shevchenko National University of Kyiv
(2h, Academician Glushkov Ave., Kyiv 03022, Ukraine)

HOW DO LONG IMPROPER PURINE-PURINE PAIRS OF DNA BASES ADAPT THE ENZYMATICALLY COMPETENT CONFORMATION? STRUCTURAL MECHANISM AND ITS QUANTUM-MECHANICAL GROUNDS

UDC 577.3

Aimed at the answer the biophysically important question posed in the title of this article, we have first investigated, at the MP2/aug-cc-pVDZ//B3LYP/6–311++G(d,p) level of quantum-mechanical theory, the structural, energetic, and dynamic features of the acquisition of the enzymatically competent conformation by the incorrect $A^*\cdot A$(WC), $G\cdot A$(WC), $A^*\cdot G^*$(WC), and $G^*\cdot G$(WC) DNA base mispairs with the Watson–Crick (WC) geometry – active players on the field of spontaneous point mutagenesis. It is first shown that the characteristic time of these non-dissociative $A^*\cdot A(WC)\leftrightarrow A^*\cdot A_{syn}$(TF), $G\cdot A(WC)\leftrightarrow G\cdot A_{syn}$, $A^*\cdot G^*(WC)\leftrightarrow A^*\cdot G^*_{syn}$, and $G\cdot G^*(WC)\leftrightarrow G\cdot G^*_{syn}$ conformational transitions is much less than the period of time that it is spent by a high-fidelity DNA-polymerase on the incorporation of one nucleotide into the DNA double helix.

Keywords: DNA biosynthesis, spontaneous transversion, large-amplitude rearrangement, conformational transition, Watson–Crick-like mispair, purine-purine DNA mismatch, MP2 and B3LYP, QTAIM.

1. Introduction

The critical analysis of the literature evidences that, at the present time, there are no generally accepted and internally non-contradictable microstructural ideas of which exactly DNA base mispairs play a role of transversions [1–9], not to mention in which way they adapt their geometry in the hydrophobic recognition pocket of a high-fidelity DNA-polymerase to the enzymatically competent, i.e., to the sizes of the classical adenine·thymine (A · T(WC)) and guanine · cytosine (G · C(WC)) Watson–Crick (WC) DNA base pairs [10, 11].

In our recent works [12–14], we have made assumption quite realistic from the biophysical point of view that the purine-purine mismatches, which play the pivotal role of transversions in the DNA biosynthesis, namely $A^*\cdot A_{syn}$(TF) [12], the so-called Topal–Fresco (TF) nucleobase pair, $G\cdot A_{syn}$ [13], $A^*\cdot G^*_{syn}$ [13] and $G\cdot G^*_{syn}$ [14], are formed in the recognition pocket of the high-fidelity DNA-polymerase during its conformational transformation from the open conformation into the enzymatically competent closed one in two successive stages (here and below, the rare, in particular mutagenic [15–17], tautomers of the DNA bases are marked with an asterisk, and the subscript "*syn*" designates the *syn*-orientation of the base relatively to the sugar residue [18], while the absence of the subscript denotes the *anti*-orientation of the base). Initially, in the open state of the high-fidelity DNA-polymerase, the so-called long $A^*\cdot A$(WC) [19], $G\cdot A$(WC)/$A\cdot G$(WC) [20] and $G\cdot G^*$(WC) [21] Watson–Crick DNA base pairs (on the left is the base that belongs to the template strand) are formed, which adapt then their geometry to the Watson–Crick architecture at the transition of the enzyme into its working conformation due to the $A^*\cdot A(WC)\rightarrow A^*\cdot A_{syn}$(TF), $G\cdot A(WC)\rightarrow G\cdot A_{syn}$, $A^*\cdot G^*(WC)\rightarrow A^*\cdot G^*_{syn}$, and $G\cdot G^*(WC)\rightarrow G\cdot G^*_{syn}$ conformational transitions.

At the same time, the quite logical question arises: "Why are not these transversions formed in the recog-

© O.O. BROVARETS, D.M. HOVORUN, 2015

nition pocket of the high-fidelity DNA polymerase at once, i.e., at the H-pairing of two purine bases, one of which (incoming one) is in the *syn*-orientation?" The answer lies in the structural and dynamic properties of the DNA-like [22] *syn*-conformers of purine 2′-deoxyribonucleosides [18]. It turns out that they are dynamically unstable structures with very short lifetimes, which does not allow them to form H-bonds in the *syn*-conformation with others bases. There is only one single output from this situation – to form a long Watson–Crick H-connected pair of the purine bases, each of which is in the *anti*-conformation, and then to switch one of these bases into the *syn*-conformation by the $anti \to syn$ conformational transition. Let us recall that such conformational transition in a hydrophobic recognition pocket of the high-fidelity DNA-polymerase is allowed from the steric consideration only for the base of the incoming (but not for the maternal!) nucleotide (here and below, the base of the incoming nucleotide in the designations of the considered H-bonded base pairs is always situated on the right, and the maternal base – on the left). Figuratively speaking, the base that would perform the $anti \to syn$ conformational transition should firstly "throw anchor" through the H-bonding with another (parent) purine base, by forming a long Watson–Crick base pair, and then it would eventually pass from the *anti*- to *syn*-conformation, which would be dynamically stable.

This work is intended to seek the answer to the biophysically important question formulated in the title of this article. By using quantum-mechanical calculations at the MP2/aug-cc-pVDZ//B3LYP/6–311++G(d,p) level of theory, we succeed, for the first time, to localize the transition states (TSs) of the $A \cdot A^*(WC) \leftrightarrow A^* \cdot A_{syn}(TF)$, $G \cdot A(WC) \leftrightarrow G \cdot A_{syn}$, $A^* \cdot G^*(WC) \leftrightarrow A^* \cdot G^*_{syn}$, and $G \cdot G^*(WC) \leftrightarrow G \cdot G^*_{syn}$ conformational transitions and to prove that they occur by a non-dissociative mechanism, i.e., without the breaking of all intermolecular H-bonds without any exception, stabilizing, base pairs. At this, the $N6H_2$ amino group of the A base and the O6H hydroxyl group of the G* base, which serve as "molecular joints" at the $anti \leftrightarrow syn$ conformational transition of the base of the incoming nucleotide, keep the intermolecular H-bonds with another base of the pair as donors of the H-bonding. It will be demonstrated that the time, during which these conformational transitions are carried out ($(1 \div 3) \cdot 10^{-7}$ s), is significantly less than the time ($\sim 8.3 \cdot 10^{-4}$ s [23]) used by the high-fidelity DNA-polymerase for the incorporation of a single nucleotide into the structure of the DNA double helix, but exceeds the period of time ($\sim 10^{-9}$ s [8]) spent by the DNA-polymerase for the forced dissociation of the H-bonded DNA base pairs into the isolated DNA bases by two orders.

2. Computational Methods

All calculations of the geometries and the harmonic vibration frequencies of the considered base mispairs, as well as the transition states of their conformational transitions, have been performed, by using Gaussian'09 package [24] at the B3LYP DFT/6–311++G(d,p) level of theory [25, 26], which was applied to analogous complexes and was verified to give accurate geometrical structures, normal mode frequencies, barrier heights, and characteristics of intermolecular H-bonds [27–29]. The scaling factor that is equal to 0.9668 has been applied in the present work to the correction of the harmonic frequencies of all studied base pairs [30–32]. We have confirmed the minima and TSs located by means of the synchronous transit-guided quasi-Newton method [33] on the potential energy landscape by the absence or presence, respectively, of the imaginary frequency in the vibrational spectra of the complexes.

In order to consider the electronic correlation effects as accurately as possible, we followed geometry optimizations with single-point energy calculations using MP2 functional [34] and a wide variety of basis sets, in particular, Pople's basis sets of valence triple-ζ quality [35, 36], as well as Dunning's cc-type basis sets [37], augmented with polarization and/or diffuse functions: 6–311++G(2df, pd) and aug-cc-pVDZ.

Reaction pathways have been established by following the intrinsic reaction coordinate (IRC) in the forward and reverse directions from each TS using the Hessian-based predictor-corrector integration algorithm [38] with tight convergence criteria. These calculations eventually ensure that the proper reaction pathway, connecting the expected reactants and products on each side of the TS, has been found [39, 40].

Electronic interaction energies E_{int} have been calculated at the MP2/6–311++G(2df, pd) level of the-

ory as the difference between the total energy of the base mispair and the energies of the isolated monomers. The Gibbs free energy of interaction has been obtained using a similar equation. In each case, the interaction energy was corrected for the basis set superposition error (BSSE) [41, 42] through the counterpoise procedure [43, 44].

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

$$G = E_{\rm el} + E_{\rm corr}, \tag{1}$$

where $E_{\rm el}$ – electronic energy, while $E_{\rm corr}$ – thermal correction. We applied the standard TS theory [45] to estimate the activation barriers of the conformational transitions.

The time $\tau_{99.9\%}$ necessary to reach 99.9% of the equilibrium concentrations of the reactant and the product in the system of reversible first-order forward ($k_{\rm f}$) and reverse ($k_{\rm r}$) reactions can be estimated by formula [45]:

$$\tau_{99.9\%} = \frac{\ln 10^3}{k_{\rm f} + k_{\rm r}}. \tag{2}$$

To estimate the values of rate constants $k_{\rm f}$ and $k_{\rm r}$,

$$k_{\rm f,r} = \Gamma \frac{k_{\rm B}T}{h} e^{-\frac{\Delta\Delta G_{\rm f,r}}{RT}}, \tag{3}$$

we applied the standard TS theory [45], in which the quantum tunneling effect is accounted by Wigner's tunneling correction [46], which was successfully applied to other biologically important reactions [47–49]:

$$\Gamma = 1 + \frac{1}{24}\left(\frac{h\nu_i}{k_{\rm B}T}\right)^2, \tag{4}$$

where $k_{\rm B}$ – Boltzmann's constant, h – Planck's constant, $\Delta\Delta G_{\rm f,r}$ – Gibbs free energy of activation for the reaction in the forward (f) and reverse (r) directions, ν_i – magnitude of imaginary frequency associated with the vibrational mode at TSs.

Bader's Quantum Theory of Atoms in Molecules (QTAIM) was used to analyze the electron density distribution [50–53]. The topology of the electron density was analyzed, by using program package AIMAll [54] with all default options. The presence of a bond critical point (BCP), namely the so-called (3,–1) BCP, and the bond path between a hydrogen donor and an acceptor, as well as the positive value of Laplacian at this BCP ($\Delta\rho > 0$), were considered as criteria for the H-bond formation [55, 56]. Wave functions were obtained at the level of theory used for the geometry optimization.

The energies of the weak C2H$\cdots$N2 H-bond [55–57] in the A$^*\cdot$G*(WC) base mispair and the attractive N1$\cdots$N6/N7 van der Waals contacts in the $\rm TS_{A^*\cdot A(WC)\leftrightarrow A^*\cdot A_{syn}(TF)}/TS_{G\cdot A(WC)\leftrightarrow G\cdot A_{syn}}$ transition states were calculated by the empirical Espinosa–Molins–Lecomte (EML) formula [57–59] based on the electron density distribution at the (3,–1) BCPs of the H-bonds:

$$E_{\rm C2H\cdots N2/N1\cdots N6/N7} = 0.5V(r), \tag{5}$$

where $V(r)$ – the value of local potential energy at the (3,–1) BCP.

The energies of all other conventional AH$\cdots$B H-bonds were evaluated by the empirical Iogansen formula [60]:

$$E_{\rm AH\cdots B} = 0.33\sqrt{\Delta\nu - 40}, \tag{6}$$

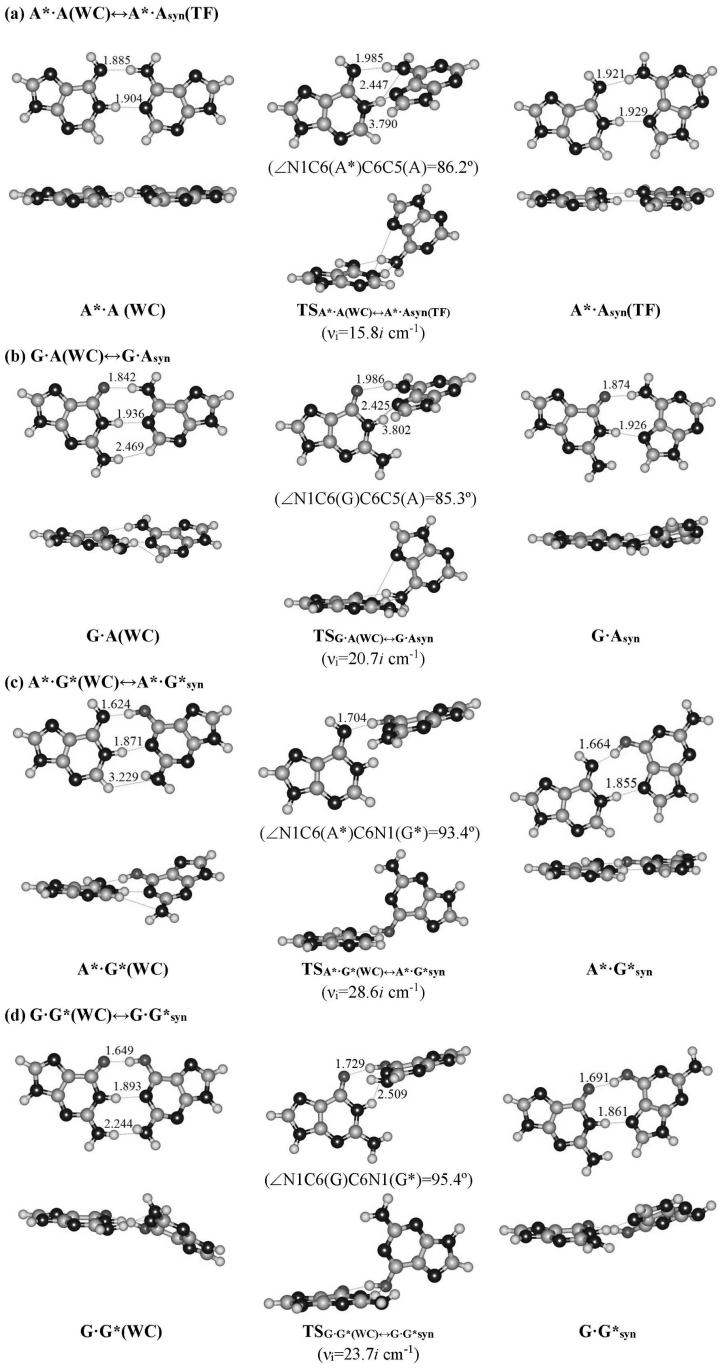
where $\Delta\nu$ – the magnitude of frequency shift of the stretching mode of the AH H-bonded group involved in the AH$\cdots$B H-bond relatively the unbound group. The partial deuteration was applied to minimize the effect of vibrational resonances [61, 62].

To determine the deformation energies necessary to be used for the investigated C$\cdot$C(WC), C$\cdot$T(WC) and T$\cdot$T*(WC) base mispairs to acquire the Watson–Crick sizes, we have adjusted the glycosidic angles at the base localized on the left within the base pair to those corresponding to the A and G bases in the A$\cdot$T(WC) and G$\cdot$C(WC) Watson–Crick base pairs, respectively, while at the base localized on the right – to those corresponding to the T and C bases, respectively, and the distance between the glycosidic hydrogens of the bases – to those of the A$\cdot$T(WC) and G$\cdot$C(WC) Watson–Crick base pairs and then frozen these parameters, by using the "opt = modredundant" key word.

The atomic numbering scheme for the DNA bases is conventional [63].

3. Results and Their Discussion

The obtained results are presented on Figure and in Tables 1–3. Their analysis allows us to reach the following conclusions.



Structures corresponding to the stationary points on the reaction pathways of the (*a*) $A^*\cdot A(WC) \leftrightarrow A^*\cdot A_{syn}(TF)$, (*b*) $G\cdot A(WC) \leftrightarrow G\cdot A_{syn}$, (*c*) $A^*\cdot G^*(WC) \leftrightarrow A^*\cdot G^*_{syn}$, and (*d*) $G\cdot G^*(WC) \leftrightarrow G\cdot G^*_{syn}$ *anti* $\leftrightarrow$ *syn* conversions through the large-scale adjustments of the bases relative to each other obtained at the B3LYP/6–311++G(d,p) level of theory. Dotted lines indicate AH$\cdots$B H-bonds and attractive N$\cdots$N van der Waals contacts (their lengths are presented in Å); dihedral angle characterizing non-planarity is indicated near each TS; carbon atoms are in light-blue, nitrogen – in dark-blue, hydrogen – in grey and oxygen – in red; ν_i – imaginary frequency

Table 1. **Electron-topological, structural, vibrational, and energetic characteristics of the intermolecular H-bonds in the purine-purine mispairs containing A and G nucleobases and TSs of their large-scale *anti* ↔ *syn* conformational reorganizations, energetic and polar characteristics of the latters obtained at the B3LYP/6–311++G(d,p) level of theory**

Base pair/TS	AH···B H-bond/N···N van der Waals contact	ρ^a	$\Delta\rho^b$	$100\varepsilon^c$	$d_{\mathrm{A\cdots B/N\cdots N}}{}^d$	$d_{\mathrm{H\cdots B}}{}^e$	Δd^f_{AH}	$\angle\mathrm{AH\cdots B}^g$	$\Delta\nu^h$	$E_{\mathrm{AH\cdots B/N\cdots N}}{}^i$	$\Delta\mathrm{G}^j$	μ^k
A*·A(WC)[19]	N6H···N6	0.035	0.091	7.19	2.918	1.885	0.028	176.3	491.5	7.01	0.00	1.75
	N1H···N1	0.034	0.087	6.76	2.943	1.904	0.030	179.0	474.6	6.88		
A*·A_{syn} (TF)[12]	N6H···N6	0.032	0.098	7.17	2.937	1.921	0.023	168.9	410.4	6.35	0.56	7.96
	N1H···N7	0.032	0.093	5.99	2.959	1.929	0.022	173.5	389.1	6.17		
$TS_{A^*\cdot A(WC)\leftrightarrow A^*\cdot A_{syn}(TF)}$	N6H···N6	0.028	0.082	3.34	2.960	1.985	0.021	157.5	362.0	5.92	8.09	5.46
	N1H···N6	0.011	0.033	40.58	3.317	2.447	0.004	143.3	63.7	1.61		
	N1···N6	0.002	0.008	67.50	3.790	–	–	–	–	0.39*		
G·A(WC)[20]	N6H···O6	0.032	0.109	3.79	2.866	1.842	0.019	176.6	336.2	5.68	0.00	5.21
	N1H···N1	0.032	0.084	6.64	2.972	1.936	0.024	178.9	429.1	6.51		
	N2H···HC2	0.004	0.014	33.40	3.153	2.469	0.0007	124.6	−0.5	0.68*		
G·A_{syn}[13]	N6H···O6	0.029	0.104	2.85	2.886	1.874	0.016	169.9	276.9	5.08	0.76	7.93
	N1H···N7	0.032	0.087	5.50	2.958	1.926	0.022	175.6	383.8	6.12		
$TS_{G\cdot A(WC)\leftrightarrow G\cdot A_{syn}}$	N6H···O6	0.024	0.083	3.30	2.942	1.986	0.015	154.7	249.1	4.77	8.39	6.71
	N1H···N6	0.011	0.035	36.44	3.296	2.425	0.003	143.3	3470.1	19.33		
	N1···N7	0.002	0.009	104.95	3.802	–	–	–	–	0.40*		
A*·G*(WC)[20]	O6H···N6	0.064	0.095	4.89	2.648	1.624	0.060	172.7	1126.9	10.88	0.00	5.26
	N1H···N1	0.037	0.091	6.52	2.905	1.871	0.028	173.1	490.6	7.01		
	C2H···N2	0.003	0.009	34.03	3.897	3.229	−0.00009	120.8	−3.2	0.42*		
A*·G^*_{syn}[13]	O6H···N6	0.057	0.103	4.56	2.645	1.664	0.048	161.0	905.4	9.71	−0.15	6.43
	N1H···N7	0.038	0.094	5.72	2.894	1.855	0.028	179.5	489.3	6.99		
$TS_{A^*\cdot G^*(WC)\leftrightarrow G^*_{syn}}$	O6H···N6	0.053	0.101	3.12	2.695	1.704	0.043	165.6	829.9	9.27	7.55	5.07
G·G*(WC)[21]	O6H···N6	0.050	0.138	2.48	2.646	1.649	0.036	171.2	687.1	8.39	0.00	8.45
	N1H···N1	0.035	0.091	6.64	2.920	1.893	0.021	172.4	385.6	6.14		
	N2H···N2	0.016	0.048	6.71	3.231	2.244	0.006	163.9	108.6	2.73		
G·G^*_{syn}[14]	O6H···O6	0.043	0.138	1.07	2.653	1.691	0.025	162.3	490.8	7.01	2.24	8.60
	N1H···N7	0.037	0.095	5.35	2.897	1.861	0.026	175.7	460.8	6.77		
$TS_{G\cdot G^*(WC)\leftrightarrow G\cdot G^*_{syn}}$	O6H···O6	0.041	0.124	0.95	2.699	1.729	0.026	164.1	486.0	6.97	9.17	7.39
	N1H···O6	0.008	0.029	202.83	3.299	2.509	0.001	134.5	317.2	5.49		

[a]The electron density at the (3,−1) BCP of the H-bond, a.u.; [b]The Laplacian of the electron density at the (3,−1) BCP of the H-bond, a.u.; [c]The ellipticity at the (3,−1) BCP of the H-bond: [d]The distance between the A (H-bond donor) and B (H-bond acceptor) atoms of the AH···B H-bond, Å; [e]The distance between the H and B atoms of the AH···B H-bond, Å; [f]The elongation of the H-bond donating group AH upon the AH···B H-bonding, Å; [g]The H-bond angle, degree; [h]The redshift of the stretching vibrational mode ν(AH) of the AH H-bonded group, cm^{-1}; [i]Energy of the H-bonds, calculated by Iogansen's [60] or Espinose–Molins–Lecomte (marked with an asterisk) [58, 59] formulas, $\mathrm{kcal\cdot mol}^{-1}$; [j]The relative Gibbs free energy of the complex obtained at the MP2/aug-cc-p VDZ//B3LYP/6–311++G(d,p) level of theory under normal conditions, $\mathrm{kcal\cdot mol}^{-1}$; [k]The dipole moment of the complex, D.

Table 2. **Energetic and kinetic characteristics of the $A^*\cdot A(WC) \leftrightarrow A^*\cdot A_{syn}(TF)$, $G\cdot A(WC) \leftrightarrow G\cdot A_{syn}$, $A^*\cdot G^*(WC) \leftrightarrow A^*\cdot G^*_{syn}$, $G\cdot G^*(WC) \leftrightarrow G\cdot G^*_{syn}$ *anti* $\leftrightarrow$ *syn* conformational transitions *via* the large-scale reorganizations of the bases relative to each other obtained at the different levels of theory for the geometry calculated at the B3LYP/6–311++G(d,p) level of theory**

Level of theory	ΔG^a	ΔE^b	$\Delta\Delta G_{TS}{}^c$	$\Delta\Delta E_{TS}{}^d$	$\Delta\Delta G^e$	$\Delta\Delta E^f$	$\tau_{99.9}{}^g$
$A^*\cdot A(WC) \leftrightarrow A^*\cdot A_{syn}(TF)$							
MP2/6–311++G(2df, pd)	0.54	1.22	7.50	7.50	6.96	6.29	1.01×10^{-7}
MP2/aug-cc-pVDZ	0.56	1.23	8.09	8.09	7.53	6.85	2.66×10^{-7}
$G\cdot A(WC) \leftrightarrow G\cdot A_{syn}$							
MP2/6–311++G(2df, pd)	0.73	0.55	7.80	8.29	7.07	7.75	1.32×10^{-7}
MP2/aug-cc-pVDZ	0.76	0.58	8.39	8.89	7.64	8.31	3.47×10^{-7}
$A^*\cdot G^*(WC) \leftrightarrow A^*\cdot G^*_{syn}$							
MP2/6–311++G(2df, pd)	−0.14	0.95	7.01	7.70	7.15	6.75	8.60×10^{-8}
MP2/aug-cc-pVDZ	−0.15	0.94	7.55	8.24	7.70	7.30	2.15×10^{-7}
$G\cdot G^*(WC) \leftrightarrow G\cdot G^*_{syn}$							
MP2/6–311++G(2df, pd)	1.84	3.06	8.30	9.70	6.46	6.64	5.85×10^{-8}
MP2/aug-cc-pVDZ	2.24	3.47	9.17	10.56	6.92	7.10	1.30×10^{-7}

[a]The Gibbs free energy of the product relative to the reactant of the reaction ($T = 298.15$ K), kcal $\cdot$ mol^{-1}; [b]The electronic energy of the product relative to the reactant of the reaction, kcal $\cdot$ mol^{-1}; [c]The Gibbs free energy barrier for the forward reaction, $\cdot$ mol^{-1}; [e]The Gibbs free energy barrier for the reverse reaction, kcal $\cdot$ mol^{-1}; [f]The electronic energy barrier for the reverse reaction, kcal $\cdot$ mol^{-1}; [g] the time necessary to reach 99.9% of the equilibrium concentration between the reactant and the product of the conformational transition, s.

The characteristic feature of all revealed transition states, connecting long Watson-Crick DNA base pairs and their enzymatically competent conformations, is their non-planarity, the low value of imaginary frequency (from $15.8i$ to $30.4i$ cm^{-1}), and the high polarity (5.07÷7.39 D) (Figure, Table 1). Moreover, all structures that correspond to the transition states are stabilized by the participation of the cooperative specific contacts, among which there are both H-bonds and attractive van der Waals contacts (Table 1). Their energetic characteristics enable us to identify them as specific interactions of a medium-strong strength. A comprehensive analysis of the nature of the conformational transitions of the long $A^*\cdot A(WC)$, $G\cdot A(WC)$, $A^*\cdot G^*(WC)$ and $G^*\cdot G(WC)$ Watson–Crick DNA base pairs into the enzymatically competent $A^*\cdot A_{syn}(TF)$, $G\cdot A_{syn}$, $A^*\cdot G^*_{syn}$, and $G\cdot G^*_{syn}$ conformations, respectively, shows that they are non-dissociative, i.e., they are not accompanied by the rupture of all H-bonds without exception. At these conformational transitions, at least one H-bond, in which an amino or hydroxyl group in the 6th position is involved, is not broken along the IRC. The comparison of the energetic characteristics presented in Table 2 with similar data for the *anti* $\leftrightarrow$ *syn* transition of the purine 2′-deoxyribonucleosides [18] clearly shows that the limiting stage of the acquisition of the enzymatically competent conformation by the long Watson–Crick DNA base pairs is precisely the conformational conversions of the pairs instead of nucleosides. The latter process as low energetic is only assisting the first one.

Despite the fact that the heterocycles of the nucleotide bases are rather soft structures for the out-of-plane bending [64, 65], they remain almost planar in the processes of conformational transformations of the pairs.

All conformational transitions without exception are rather rapid processes (Table 2) in comparison with the time that the high-fidelity DNA-polymerase spends for the incorporation of a single nucleotide into the structure of the DNA double helix that is synthesized ($\sim 8.3\cdot10^{-4}$ s [23]). Moreover, all final $A^*\cdot A_{syn}(TF)$, $G\cdot A_{syn}$, $A^*\cdot G^*_{syn}$ and $G\cdot G^*_{syn}$ com-

Table 3. **Structural and energetic characteristics of the incorrect pyrimidine-pyrimidine base pairs, responsible for the spontaneous transversions, and the classical A · T(WC) and G · C(WC) Watson–Crick DNA base pairs obtained at the MP2/6–311++G(2df, pd)//B3LYP/6–311++G(d,p) level of theory**

Base mispair	Reference	R($H_{N1/N9}$-$H_{N1/N9}$)[a], Å	α_1[b], °	α_2[c], °	ΔE_{def}[d], kcal · mol^{-1}	
					to A · T(WC)	to G · C(WC)
C · C*(WC)	[66]	8.086	60.3	59.5	8.57	8.76
C · T(WC)	[67]	8.215	59.7	57.0	8.67	8.87
T · T*(WC)	[68]	8.385	53.6	58.1	10.97	10.91
A · T(WC)	[10]	10.130	54.3	54.8	0.00	0.25
G · C(WC)	[11]	10.209	52.9	55.3	0.11	0.00

[a]The distance between the glycosidic protons at the N1/N9 atoms in the pyrimidine-pyrimidine and purine-pyrimidine base pairs, respectively; [b]Glycosidic angle for the base situated on the left within the base pair; [c]Glycosidic angle for the base situated on the right within the base pair; [d]The deformation energy necessary to apply for the DNA mismatch to acquire the sizes of the A · T(WC) and G · C(WC) Watson–Crick DNA base pairs.

plexes – products of the A*· A(WC) ↔ A*· A_{syn}(TF), G·A(WC) ↔ G · A_{syn}, A*· G*(WC) ↔ A*· G^*_{syn}, and G · G*(WC) ↔ G · G^*_{syn} conformational transformations – are dynamically stable complexes: their lifetimes ($\tau = (2 \div 7)10^{-8}$ s) enable all their low-frequency intermolecular vibrations to develop. Noteworthy, that the values of lifetime τ exceed, by one order, the time spent by the DNA-polymerase machinery for the forced dissociation of DNA pairs into isolated DNA bases ($\sim 10^{-9}$ s [7]).

Previously, we have demonstrated that all incorrect A*· A_{syn}(TF) [12], G · A_{syn} [13], A*· G^*_{syn} [13], and G · G^*_{syn} [14] DNA base pairs, that are involved in the origin of spontaneous point mutations, acquire quite easily the characteristic geometrical dimensions of the classical A · T(WC) [10] and G · C(WC) [11] Watson–Crick DNA base pairs during thermal fluctuations [7], which eventually guarantees their chemical incorporation into the structure of the DNA double helix that is synthesized. In this paper, we have filled the existing gap, by calculating the analogous data for the so-called short C · C*(WC) [66], C · T(WC) [67], and T · T*(WC) [68] Watson–Circk DNA base pairs playing the role of spontaneous transversions. It was revealed that these nucleobase pairs are characterized by a considerably greater energy than the one that should be spent to adjust their geometry to classical Watson–Crick sizes (Table 3). Nevertheless, we believe that these estimates are overstated, because it cannot be excluded that these stretched pyrimidine-pyrimidine base pairs would be additionally stabilized by the interactions with the recognition pocket of the high-fidelity DNA-polymerase. The reason for this assumption is structural feature of these pairs – the positions of their N3 and O2 atoms in the stretched state coincide with the locations of similar atoms in the A · T(WC), T · A(WC), G · C(WC) and C · G(WC) Watson–Crick DNA base pairs. It should be noted that, according to the QTAIM analysis, these base pairs in the stretched conformational state are stabilized by two weakened intermolecular H-bonds.

It is worth to mention that none of the purine-purine and pyrimidine-pyrimidine DNA base mispairs discussed in this paper, which are, figuratively speaking, "star actors" in their enzymatically competent conformations, does not create steric hindrances in the recognition pocket of the high-fidelity DNA-polymerase from the side of the DNA minor groove. We have arrived at this conclusion, by carefully comparing the van der Waals outlines of all incorrect base mispairs in their enzymatically competent conformations from the side of the DNA minor groove with similar summary outlines of all four canonical A · T(WC), T · A(WC), G · C(WC), and C · G(WC) Watson–Crick DNA base pairs.

4. Conclusions

Here, at the MP2/aug-cc-pVDZ//B3LYP/6–311++G(d,p) level of quantum-mechanical theory, the structural, energetic, and dynamical features of the acquisition of the enzymatically competent conformation by the incorrect A*· A(WC),

$G \cdot A(WC)$, $A^* \cdot G^*(WC)$ and $G \cdot G^*(WC)$ DNA base mispairs with Watson–Crick geometry have been determined for the first time. We have found that the characteristic time of these non-dissociative $A^* \cdot A(WC) \leftrightarrow A^* \cdot A_{syn}(TF)$, $G \cdot A(WC) \leftrightarrow G \cdot A_{syn}$, $A^* \cdot G^*(WC) \leftrightarrow A^* \cdot G^*_{syn}$, and $G \cdot G^*(WC) \leftrightarrow G \cdot G^*_{syn}$ *anti* $\leftrightarrow$ *syn* conformational transitions is much less than the time spent by the high-fidelity DNA-polymerase on the incorporation of one nucleotide into the DNA double helix. The possible biological applications of the obtained results and the promising lines of the future research were discussed in our previous works [69, 70] in details.

We dedicate this paper to the outstanding scientist, excellent teacher, and our dear colleague Academician Leonid A. Bulavin, communication with whom inspires to creativity, on the occasion of his 70th anniversary.

O.O.B. was supported by the grant of the President of Ukraine to support the scientific researches of young scientists for the year 2015 from the State Fund for Fundamental Research of Ukraine (project No. GP/F61/028), by the grant of the NAS of Ukraine for the realization of the scientific and research works of young scientists of the NAS of Ukraine for years 2015–2016, and by the Scholarship of the President of Ukraine for young scientists for the years 2014–2016. This work was performed, by using computational facilities of the joint computer cluster of SSI "Institute for Single Crystals" of the NAS of Ukraine and the Institute for Scintillation Materials of the NAS of Ukraine incorporated into the Ukrainian National Grid.

1. J.D. Watson and F.H.C. Crick, Cold Spring Harb. Symp. Quant. Biol. **18**, 123 (1953).
2. E.B. Freese, Brookhaven Symp. Biol. **12**, 63 (1959).
3. M.D. Topal and J.R. Fresco, Nature **263**, 285 (1976).
4. R.C. von Borstel, Mutat. Res. **307**, 131 (1994).
5. E.C. Friedberg, G.C. Walker, W. Siede, R.D. Wood, R.A. Schultz, and T. Ellenberger, *DNA Repair and Mutagenesis* (ASM Press, Washington, D.C., 2006).
6. O.S. Kochina, Ye.P. Yurenko, and D.M. Hovorun, Biopol. Cell **23**, 167 (2007).
7. O.O. Brovarets' and D.M. Hovorun, Ukr. Biochem. J. **82**, 55 (2010).
8. O.O. Brovarets', I.M. Kolomiets', and D.M. Hovorun, *Quantum Chemistry – Molecules for Innovations*, edited by T. Tada (In Tech, 2012), p. 59.
9. G. Rossetti, P.D. Dans, I. Gomez-Pinto, I. Ivani, G. Gonzalez, and M. Orozco, Nucleic Acids Res. **43**, 4309 (2015).
10. O.O. Brovarets' and D.M. Hovorun, J. Biomol. Struct. & Dynam. **32**, 127 (2014).
11. O.O. Brovarets' and D.M. Hovorun, J. Biomol. Struct. & Dynam. **32**, 1474 (2014).
12. O.O. Brovarets', R.O. Zhurakivsky, and D.M. Hovorun, Phys. Chem. Chem. Phys. **16**, 3715 (2014).
13. O.O. Brovarets' and D.M. Hovorun, Phys. Chem. Chem. Phys. **16**, 9074 (2014).
14. O.O. Brovarets' and D.M. Hovorun, Mol. Phys. **112**, 3033 (2014).
15. O.O. Brovarets' and D.M. Hovorun, Biopol. Cell **27**, 221 (2011).
16. A. Furmanchuk, O. Isayev, L. Gorb, O.V. Shishkin, D.M. Hovorun, and J. Leszczynski, Phys. Chem. Chem. Phys. **13**, 4311 (2011).
17. S.P. Samijlenko, Y.P. Yurenko, A.V. Stepanyugin, and D.M. Hovorun, J. Phys. Chem. B **114**, 1454 (2012).
18. R.O. Zhurakivsky and D.M. Hovorun, in *Physical Principles of Molecular Organization and Dynamics of Structural Biopolymers*, edited by Yu.P. Blagoi *et al.* (Karazin National Univ. of Kharkiv, Kharkiv, 2012), p. 71.
19. O.O. Brovarets', R.O. Zhurakivsky, and D.M. Hovorun, J. Mol. Model. **19**, 4223 (2013).
20. O.O. Brovarets', R.O. Zhurakivsky, and D. M. Hovorun, J. Comput. Chem. **35**, 451 (2014).
21. O.O. Brovarets' and D.M. Hovorun, Phys. Chem. Chem. Phys. **6**, 15886 (2014).
22. Y.P. Yurenko, R.O. Zhurakivsky, S.P. Samijlenko, and D.M. Hovorun, J. Biomol. Struct. & Dynam. **29**, 5 (2011).
23. S. Kirmizialtin, V. Nguyen, K.A. Johnson, and R. Elber, Structure **20**, 618 (2012).
24. M.J. Frisch, G.W. Trucks, H.B. Schlegel *et al.*, *GAUSSIAN 09 (Revision B.01)* (Gaussian Inc., Wallingford, 2010)
25. R.G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford Univ. Press, Oxford, 1989).
26. C. Lee, W. Yang, and R.G. Parr, Phys. Rev. B **37**, 785 (1988).
27. K.B. Wiberg, J. Comput. Chem. **25**, 1342 (2004).
28. C.F. Matta, J. Comput. Chem. **31**, 1297 (2010).
29. O.O. Brovarets', R.O. Zhurakivsky, and D.M. Hovorun, Chem. Phys. Lett. **578**, 126 (2013).
30. O.O. Brovarets' and D.M. Hovorun, Biopol. Cell **26**, 72 (2010).
31. O.O. Brovarets' and D.M. Hovorun, Biopol. Cell **26**, 295 (2010).
32. O.O. Brovarets', R.O. Zhurakivsky, and D.M. Hovorun, Biopol. Cell **26**, 398 (2010).
33. C. Peng, P.Y. Ayala, H.B. Schlegel, and M.J. Frisch, J. Comput. Chem. **17**, 49 (1996).
34. M.J. Frisch, M. Head-Gordon, and J.A. Pople, Chem. Phys. Lett. **166**, 281 (1990).
35. P.C. Hariharan and J.A. Pople, Theor. Chem. Acc. **28**, 213 (1973).
36. R. Krishnan, J.S. Binkley, R. Seeger, and J.A. Pople, J. Chem. Phys. **72**, 650 (1980).
37. R.A. Kendall, T.H. Dunning, Jr, and R.J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
38. H.P. Hratchian and H.B. Schlegel, in *Theory and Applications of Computational Chemistry: The First 40 Years*,

edited by C.E. Dykstra, G. Frenking, K.S. Kim, and G. Scuseria (Elsevier, Amsterdam, 2005), p. 195.
39. O.O. Brovarets', R.O. Zhurakivsky, and D.M. Hovorun, J. Biomol. Struct. & Dynam. **33**, 674 (2015).
40. O.O. Brovarets', R.O. Zhurakivsky, and D.M. Hovorun, J. Mol. Model. **19**, 4119 (2013).
41. S.F. Boys and F. Bernardi, Mol. Phys. **19**, 553 (1970).
42. M. Gutowski, J.H. Van Lenthe, J. Verbeek, F.B. Van Duijneveldt, and G. Chalasinski, Chem. Phys. Lett. **124**, 370 (1986).
43. J.A. Sordo, S. Chin, and T.L. Sordo, Theor. Chem. Acc. **74**, 101 (1988).
44. J.A. Sordo, J. Mol. Struct. **537**, 245 (2001).
45. P.W. Atkins, *Physical Chemistry* (Oxford Univ. Press, Oxford, 1998).
46. E. Wigner, Zeit. Physik. Chemie B **19**, 203 (1932).
47. O.O. Brovarets', R.O. Zhurakivsky, and D.M. Hovorun, Mol. Phys. **112**, 2005 (2014).
48. O.O. Brovarets' and D.M. Hovorun, J. Biomol. Struct. & Dynam. **33**, 28 (2015).
49. O.O. Brovarets' and D.M. Hovorun, J. Biomol. Struct. & Dynam. **33**, 925 (2015).
50. R.F.W. Bader, *Atoms in Molecules: A Quantum Theory* (Oxford Univ. Press, Oxford, 1990).
51. C.F. Matta, J. Comput. Chem. **35**, 1165 (2014).
52. C. Lecomte, E. Espinosa, and C.F. Matta, IUCrJ **2**, 161 (2015).
53. C.F. Matta, L. Huang, and L. Masaa, J. Phys. Chem. A **115**, 12451 (2011).
54. T.A. Keith, *AIMAll* (Version 10.07.01) (2010), available at: aim.tkgristmill.com.
55. O.O. Brovarets', Y.P. Yurenko, and D.M. Hovorun, J. Biomol. Struct. & Dynam. **32**, 993 (2015).
56. O.O. Brovarets', Y.P. Yurenko, and D.M. Hovorun, J. Biomol. Struct. & Dynam. **33**, 1624 (2014).
57. C.F. Matta, N. Castillo, and R.J. Boyd, J. Phys. Chem. B **110**, 563 (2006).
58. E. Espinosa, E. Molins, and C. Lecomte, Chem. Phys. Lett. **285**, 170 (1998).
59. I. Mata, I. Alkorta, E. Espinosa, and E. Molins, Chem. Phys. Lett. **507**, 185 (2011).
60. A.V. Iogansen, Spectrochim. Acta A Mol. Biomol. Spectrosc. **55**, 1585 (1999).
61. O.O. Brovarets' and D.M. Hovorun, Ukr. Biochem. J. **82**, 55 (2010).
62. O.O. Brovarets' and D.M. Hovorun, Ukr. Biochem. J. **82**, 51 (2010).
63. W. Saenger, *Principles of Nucleic Acid Structure* (Springer, New York, 1984).
64. D.N. Govorun, V.D. Danchuk, Ya.R. Mishchuk, I.V. Kondratyuk, N.F. Radomsky, and N.V. Zheltovsky, J. Mol. Struct. **267**, 99 (1992).
65. T.Y. Nikolaienko, L.A. Bulavin, and D.M. Hovorun, J. Biomol. Struct. & Dynam. **29**, 563 (2011).
66. O.O. Brovarets' and D.M. Hovorun, Phys. Chem. Chem. Phys. **15**, 20091 (2013).
67. O.O. Brovarets' and D.M. Hovorun, J. Comput. Chem. **34**, 2577 (2013).
68. O.O. Brovarets', R.O. Zhurakivsky, and D.M. Hovorun, Chem. Phys. Lett. **592**, 247 (2014).
69. O.O. Brovarets' and D.M. Hovorun, Phys. Chem. Chem. Phys. **17**, 15103 (2015).
70. O.O. Brovarets' and D.M. Hovorun, J. Biomol. Struct. & Dynam., DOI: 10.1080/07391102.2015.1046936 (2015).

Received 07.07.15

О.О. Броварець, Д.М. Говорун

ЯК ДОВГІ НЕПРАВИЛЬНІ ПУРИНОВО-ПУРИНОВІ ПАРИ ОСНОВ ДНК НАБУВАЮТЬ ЕНЗИМАТИЧНО-КОМПЕТЕНТНОЇ КОНФОРМАЦІЇ? СТРУКТУРНИЙ МЕХАНІЗМ ТА ЙОГО КВАНТОВО-МЕХАНІЧНЕ ОБҐРУНТУВАННЯ

Р е з ю м е

Прагнучи дати відповідь на біофізично важливе запитання, винесене у заголовок статті, ми вперше дослідили на квантово-механічному рівні теорії MP2/aug-cc-pVDZ//B3LYP/6–311++G(d,p) структурно-енергетичні та динамічні особливості набуття неправильними парами основ ДНК $A^*\cdot A(WC)$, $G\cdot A(WC)$, $A^*\cdot G^*(WC)$ і $G^*\cdot G(WC)$ – активними гравцями на полі спонтанного точкового мутагенезу – ензиматично-компетентної конформації. Вперше показано, що характерний час цих недисоціативних переходів $A^*\cdot A(WC)\leftrightarrow A^*\cdot A_{syn}(TF)$, $G\cdot A(WC)\leftrightarrow G\cdot A_{syn}$, $A^*\cdot G^*(WC)\leftrightarrow A^*\cdot G^*_{syn}$ та $G\cdot G^*(WC)\leftrightarrow G\cdot G^*_{syn}$ значно менший періоду часу, який витрачає високоточна ДНК-полімераза на інкорпорацію одного нуклеотиду у подвійну спіраль ДНК, що синтезується.

О.А. Броварец, Д.Н. Говорун

КАК ДЛИННЫЕ НЕПРАВИЛЬНЫЕ ПУРИНОВО-ПУРИНОВЫЕ ПАРЫ ОСНОВАНИЙ ДНК АДАПТИРУЮТ СВОЮ КОНФОРМАЦИЮ К ЭНЗИМАТИЧЕСКИ-КОМПЕТЕНТНОЙ? СТРУКТУРНЫЙ МЕХАНИЗМ И ЕГО КВАНТОВО-МЕХАНИЧЕСКОЕ ОБОСНОВАНИЕ

Р е з ю м е

Стремясь дать ответ на биофизически важный вопрос, вынесенный в заглавие статьи, мы впервые исследовали на квантово-механическом уровне теории MP2/aug-cc-pVDZ//B3LYP/6–311++G(d,p) структурно-энергетические и динамические особенности адаптации неправильными парами оснований ДНК $A^*\cdot A(WC)$, $G\cdot A(WC)$, $A^*\cdot G^*(WC)$ и $G^*\cdot G(WC)$ – активными игроками на поле спонтанного точечного мутагенеза – энзиматически компетентной конформации. Впервые показано, что характерное время этих недиссоциативных переходов $A^*\cdot A(WC)\leftrightarrow A^*\cdot A_{syn}(TF)$, $G\cdot A(WC)\leftrightarrow G\cdot A_{syn}$, $A^*\cdot G^*(WC)\leftrightarrow A^*\cdot G^*_{syn}$ и $G\cdot G^*(WC)\leftrightarrow$ $\leftrightarrow G\cdot G^*_{syn}$ значительно меньше периода времени, в течение которого высокоточная ДНК-полимераза инкорпорирует один нуклеотид в двойную спираль ДНК при ее биосинтезе.