RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

New structural hypostasis of the A·T and G·C Watson-Crick DNA base pairs caused by their mutagenic tautomerisation in a wobble manner: A QM/QTAIM prediction

Ol'ha O. Brovarets'[a,b] & Dmytro M. Hovorun[a,b,✉]

[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. E-mail: dhovorun@imbg.org.ua

Abstract. It is widely known that canonical DNA base pairs preserve Watson-Crick (WC) scheme of the pairing that enables accurate transfer of genetic information across generations. Herein, we present firstly discovered biologically important property of the classical A·T(WC) and G·C(WC) nucleobase pairs to tautomerise into the wobble (w) H-bonded mismatches – $A^*\cdot T_{\uparrow}$(w), $A\cdot T^*_{O2\uparrow}$(w), $A\cdot T^*_{\downarrow}$(w) and $G\cdot C^*_{\uparrow}$(w), $G^*\cdot C_{\downarrow}$(w), $G\cdot C^*_{\downarrow}$(w), $G^*\cdot C_{\uparrow}$(w), respectively, containing mutagenic tautomers of the bases (denoted by asterisks), and *vice versa*. QM and QATIM calculations show that these non-dissociative tautomerisation processes, that are quite fast in comparison with the rate of the DNA replication in the cell, are controlled by the highly stable and highly polar transition states – tight $A^+\cdot T^-$, $A^-\cdot T^+$, $G^+\cdot C^-$ and $G^-\cdot C^+$ ion pairs, – formed by the protonated and deprotonated bases. Estimation of the populations of the tautomerised states of the A·T(WC) ($6.1\cdot10^{-9}$–$1.5\cdot10^{-7}$) and G·C(WC) ($4.2\cdot10^{-11}$–$1.4\cdot10^{-9}$) base pairs in the continuum with ε=4, corresponding to the interface of the protein-nucleic acid interactions, points to their involvement in the processes of the nucleation of the spontaneous point replication errors in DNA arising with frequencies to around 10^{-11}–10^{-9} errors per nucleotide replicated. Presented conception about the mutagenic quantum wobbling of the canonical A·T(WC) and G·C(WC) nucleobase pairs as their intrinsic property resulting in their tautomerisation allows us to explain the increase in the frequency of point mutations induced by 2-aminopurine and also non-contradictory interpret the NMR experimental literature data that show slow exchange between the G·P(WC) Watson-Crick-like base pair with the corresponding wobble pair, which includes 6H,8H-3,4-dihydropyrimido[4,5-c][1,2]oxazin-7-one (P) mutagen. These findings significantly expand the horizons of the classical Watson-Crick tautomeric hypothesis of the spontaneous point mutagenesis. Collectively, our investigation revealed hitherto unknown ability of the canonical Watson-Crick DNA base pairs to switch into the wobble mismatches with mutagenic tautomers, clarifying the nature of genome instability. Moreover, this paper reveals new facets of the Watson–Crick tautomeric hypothesis of the spontaneous point mutagenesis arising at the DNA replication and significantly broadens the possibilities for rational design of the chemical mutagens with targeted action, which could be interesting for medicine, synthetic biology and biotechnology.

Keywords: A·T and G·C Watson-Crick DNA base pairs · DNA replication · Spontaneous point mutations · Induced mutations · Wobble-like conformation · Tautomeric transition · 2-aminopurine · 6H,8H-3,4-dihydropyrimido[4,5-c][1,2]oxazin-7-one · Mutagenicity

† Electronic supplementary information (ESI) available: (i) Computational details; (ii) Profiles of the glycosidic parameters of the Watson-Crick DNA base pairs along the IRC of the tautomerisation; (iii) Stationary structures on the reaction pathways of the tautomeric transformations of the A·T and G·C Watson-Crick DNA base pairs; (iv) Energetic profiles of the reaction pathways of the tautomeric transformations of the A·T and G·C Watson-Crick DNA base pairs; (v) Physico-chemical parameters of the specific intermolecular contacts in the investigated DNA base pairs and TSs of their wobbling tautomerisations; (vi) Energetic and kinetic characteristics of the tautomerisations of the DNA base pairs containing canonical bases and mutagenic analogue of cytosine P, and also of the conformational interconversions of the mirror-symmetric enantiomers of the DNA base mispairs; (vii) Interaction energies for the investigated DNA base pairs and TSs of their tautomeric wobbling; (viii) Energetic and kinetic characteristics of the wobbling tautomerisations of the Watson-Crick DNA base pairs; (ix) Ranges of the existence of the obtained patterns of the specific intermolecular interactions along the IRC of the Watson-Crick DNA base pairs tautomerisations; (x) Movies of the tautomeric conversions of the A·T and G·C Watson-Crick DNA base pairs *via* the sequential DPT into the wobble mismatches containing mutagenic tautomers. See DOI: 10.1039/x0xx00000x.

Although DNA is replicated in the cell with high precision, anyway errors occur[1,2]. In recent years Watson-Crick tautomeric hypothesis[3] that associates the nature of spontaneous point mutations arising in DNA with the transition of nucleotide bases from the canonical into the rare, mutagenic tautomeric forms[4,5] resulting in the formation of the base pairs by their participation, has received reliable experimental[6-10], in particular X-ray analysis[7,8], NMR relaxation dispersion measurements,[9,10] and theoretical[11-13] verifications. Nevertheless, these data do not allow us to answer on the sore point according the physico-chemical mechanisms enabling DNA bases in the canonical tautomeric form to acquire rare or mutagenic tautomeric form before the dissociation of the Watson-Crick (WC) nucleobase pairs into the monomers by the replication machinery in order to produce mispairs resulting in further misincorporations and as a result – the spontaneous point mutations at the DNA replication. At the same time, the clarification of the concrete physico-chemical mechanisms[11-16] that ensure transition of the nucleotide bases from the canonical into the rare tautomeric form at the DNA replication, i.e. at the forced dissociation of the Watson-Crick nucleobase pairs into the monomers by the replication machinery, remains its "narrow" place.

As of today, it has been definitely established that double proton transfer (DPT) along neighboring intermolecular H-bonds[17] as in the Watson-Crick[18-21] and wobble[22,23] DNA base pairs, so in the complexes of amino acid residues of the DNA-binding proteins with nucleotide bases[24,25] can't cope with this biologically important task.

In this paper we have pursued the goal to reveal unknown properties of the well-studied canonical A·T(WC) and G·C(WC) DNA base pairs, which are inherent to these H-bonded complexes itself and able to explain the intrinsic DNA mutability without involvement of the external agents.

All calculations of the geometries and harmonic vibrational frequencies of the considered base mispairs and transition states of their conversion have been performed using Gaussian'09 package[26] at the DFT (B3LYP)/6-311++G(d,p) level of theory[27-29], that has been applied for analogous systems and verified to give accurate geometrical structures, normal mode frequencies, barrier heights and characteristics of intermolecular H-bonds[20,21,30]. A scaling factor that is equal to 0.9668 has been applied in the present work for the correction of the harmonic frequencies of all studied base pairs[31-35]. We have confirmed the minima and TSs, located by means of Synchronous Transit-guided Quasi-Newton method[36], on the potential energy landscape by the absence or presence, respectively, of the imaginary frequency in the vibrational spectra of the complexes. We applied standard TS theory[37] to estimate the activation barriers of the tautomerisation reactions.

Reaction pathways have been monitored by following intrinsic reaction coordinate (IRC) in the forward and reverse directions from each TS using 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. We've have investigated the evolution

of the energetic and geometric characteristics of the H-bonds and base pairs along the reaction pathway establishing them at each point of the IRC[39-41].

In order to consider electronic correlation effects as accurately as possible, we followed geometry optimizations with single point (SP) energy calculations using MP2 level of theory[42] and 6-311++G(2df,pd) Pople's basis set of valence triple-ζ quality[43,44] and aug-cc-pVDZ Dunning's cc-type basis set[45], augmented with polarization and/or diffuse functions.

Physico-chemical parameters have been estimated by the known formulas of the physico-chemical kinetics[37] (for more details see Methods Section in ESI). Bader's quantum theory of Atoms in Molecules (QTAIM)[46,47] was applied to analyse the electron density distribution[48,49].

In the present paper we have considered the simplest physico-chemical model of the base mispairs in the base-pair recognition pocket of the high-fidelity DNA-polymerase, namely the H-bonded pairs of nucleotide bases in the continuum with ε=1/ε=4. In this case, we have relied on the results obtained in the work[50], in which the adequacy of this model was convincingly proved, at least at the study of the tautomerisation of the H-bonded pairs of nucleotide bases, where insignificance of the influence of the stacking and the sugar-phosphate backbone on the tautomerisation process has been demonstrated. Thereby, their impact can be neglected in the first approximation. In addition, the applied model can help to distinguish the lowest structural level, at which the tautomerisation effects can be observed, and to estimate the changes at the sequential complication of the model.

In this study we have chosen the simplest level of the base pairs that adequately reflects the processes occurring in real systems[51] without deprivation of the structurally functional properties of the bases in the composition of DNA. In this case, the value of the effective dielectric constant ε ($1 < \varepsilon < 4$), that is characteristic for the anhydrous molecular crystals, satisfactorily models the substantially hydrophobic recognition pocket of the DNA-polymerase machinery as a part of the replisome [19,21,25].

Mutagenic tautomeric wobbling of the canonical A·T Watson–Crick DNA base pair. In this paper we discovered for the first time a new physico-chemical property of the classical A·T(WC) DNA base pair to be in the fast (in comparison with the rate of the DNA replication in the cell[2]) tautomeric equilibrium with the short-lived wobble (w) $A^*\cdot T_{\uparrow}(w)$, $A\cdot T^*_{O2\uparrow}(w)$ and $A\cdot T^*_{\downarrow}(w)$ H-bonded mismatches, comprising mutagenic tautomers of nucleotide bases (here and below marked by an asterisk) (Figs. 1, 2, Table 1 and Fig. S1, Tables S1-S3)**.** It turned out that the A·T(WC) DNA base pair exists simultaneously in three other biologically important hypostasis – $A^*\cdot T_{\uparrow}(w)$, $A\cdot T^*_{O2\uparrow}(w)$ and $A\cdot T^*_{\downarrow}(w)$. Their forced dissociation into monomers by the DNA-polymerase machinery with necessity generates mutagenic tautomers of the DNA bases, which are themselves long-lived structures[52] and are able to cause spontaneous point mutations[11-13].

Common feature of all three A·T(WC)↔$A^*\cdot T_{\uparrow}(w)$, A·T(WC)↔$A\cdot T^*_{O2\uparrow}(w)$ and A·T(WC)↔$A\cdot T^*_{\downarrow}(w)$ routes of the quantum mutagenic tautomerisation *via* the sequential DPT accompanied with the structural rearrangement of the bases lies in the fact that they are controlled by the plane symmetrical, highly stable and

highly polar transition states – tight ion pairs formed by the protonated A^+ base and deprotonated T^- base – and occur by the non-dissociative mechanism (Figs. 1, 2 and Fig. S1). This non-dissociative quantum mechanism of the mutual transformation of the base pairs is accompanied by the significant alteration of their geometry (from Watson-Crick to wobble and *vice versa*) and succession of the patterns (from 9 to 10) of the specific intermolecular interactions including both AH···B H-bonds and loosened A-H-B covalent bridges along the IRC of tautomerisation (Figs. 1-3 and Table S5). Quite high stabilization energy (>100 kcal·mol^{-1}) of the transition states of the tautomerisation processes (Table S3) negates the direct participation of water molecules in them as a chemical agent and eliminates influence of the stacking interactions[53] on the course of the tautomerisation. Such point of view[11-13] is confirmed experimentally for similar G·T(w)↔G*·T(WC) tautomerisation[9]. C_s symmetry of the structures involved in all three tautomerisation processes remains unchanged; at this orientation of the methyl group of the T base also remains undisturbed. Notably, both tautomerised A*·$T_{\uparrow}$(w), A·T*$_{O2\uparrow}$(w) and A·T*$_{\downarrow}$(w) mismatches and transition states of their tautomerisation – A·T(WC)↔A*·$T_{\uparrow}$(w), A·T(WC)↔A·T*$_{O2\uparrow}$(w) and A·T(WC)↔A·T*$_{\downarrow}$(w) – are considerably more polar structures than the A·T(WC) DNA base pair (Table S1). These structures that are involved in the above described processes of tautomerisation are thermodynamically stable structures with negative value of the Gibbs free energy of interaction (ΔG_{int}<0) (Table S3).

The A·T(WC)↔A·T*$_{\downarrow}$(w) tautomerisation process, where two-stage transition of the one and the same proton occurs, is much easier than two others A·T(WC)↔A*·$T_{\uparrow}$(w) and A·T(WC)↔A·T*$_{O2\uparrow}$(w) (Figs. 1, 2). These two last processes take place through a common $TS^{A+\cdot T-}_{A\cdot T(WC)\leftrightarrow A+\cdot T-(w)}$ transition state and dynamically unstable intermediate – H-bonded $A^+\cdot T^-$(w) tight ion pair with wobble architecture: exactly at this $A^+\cdot T^-$(w) point it occurs bifurcation of the tautomerisation into two different (in topological and energetic aspects) pathways, which final products are A*·$T_{\uparrow}$(w) and A·T*$_{O2\uparrow}$(w) mispairs. Herewith, the $TS^{A+\cdot T-}_{A\cdot T(WC)\leftrightarrow A+\cdot T-(w)}$ is in fact the transition state for both A·T(WC)↔A*·$T_{\uparrow}$(w) and A·T(WC)↔A·T*$_{O2\uparrow}$(w) tautomerisation reactions (Figs. 1, 2). It was established at the MP2/aug-cc-pVDZ//B3LYP/6-311++G(d,p) level of theory that the $A^+\cdot T^-$(w) intermediate is absent on the potential energy surface, since the $A^+\cdot T^-$(w)↔A·T*$_{O2\uparrow}$(w) and $A^+\cdot T^-$(w)↔A*·$T_{\uparrow}$(w) tautomeric conversions are barrier-less in the continuum with ε=1/ε=4 (see Table S4).

Comparison of the energy barriers of the mutagenic tautomerisation of the isolated A and T bases (45.7 and 39.2 kcal·mol^{-1}, respectively)[52] with similar values in the composition of the pair (17.4 and 20.8 kcal·mol^{-1}, respectively (Table S4)) clearly evidences that complementary H-bonded bases catalyze the tautomeric transformation of each other into the mutagenic tautomeric forms.

On the one hand, population of the tautomerised states of the A·T(WC) DNA base pair – A*·$T_{\uparrow}$(w), A·T*$_{O2\uparrow}$(w) and A·T*$_{\downarrow}$(w) – consists $5.4\cdot10^{-8}/1.5\cdot10^{-7}$, $9.9\cdot10^{-9}/3.3\cdot10^{-8}$ and $2.5\cdot10^{-10}/6.1\cdot10^{-9}$, respectively, obtained at ε=1/ε=4 (the continuum with a low dielectric constant is typical for hydrophobic interfaces of specific protein-nucleic acid interactions[25,54]) (Fig. 1), that fits into the range of the frequencies of the spontaneous point mutations observed experimentally (10^{-11}-10^{-9})[55-57].

On the other hand, wobble A*·T$_{\uparrow}$(w), A·T*$_{O2\uparrow}$(w) and A·T*$_{\downarrow}$(w) H-bonded mispairs satisfy all necessary conditions in order to be successfully dissociated into the monomers by the replication machinery[19-23]. Indeed, they incorporate well into the structure of the DNA double helix[58,59], interaction energy between bases in them does not exceed a similar value for the G·C(WC) DNA base pair (ΔE_{int}=-29.28 and ΔG_{int}=-15.97 kcal·mol^{-1})[21] and they have a fairly long lifetimes (τ=1.6·10^{-8}, 2.9·10^{-9} and 3.0·10^{-8} s, respectively) (Table 1). This indicates that uncovered by us for the first time internally inherent to the A·T(WC) DNA base pair property to exist in a rapid tautomeric equilibrium with the A*·T$_{\uparrow}$(w), A·T*$_{O2\uparrow}$(w) and A·T*$_{\downarrow}$(w) mispairs (time $\tau_{99.9\%}$ necessary to reach 99.9% of the equilibrium concentration between the reactant and the product of the tautomerisation reactions is equal to 1.1·10^{-7}/6.0·10^{-9}, 2.0·10^{-8}/1.4·10^{-9} and 2.1·10^{-7}/9.2·10^{-9} s (ε=1/ε=4), respectively (Table S4)) has direct relevance to the occurrence of the spontaneous replication errors in DNA.

Low population of the tautomerised states of the classical A·T(WC) DNA base pair complicates its direct registration by the experimental physico-chemical methods. Meanwhile, developed in this work conceptions about the spontaneous mutagenic tautomerisation of the A·T(WC) base pair allow us to explain quite naturally microstructural mechanism of the mutagenic pressure exerted by 2-aminopurine (2AP) on the DNA. It is well known[60-64] that this strong mutagen[65] causes A·T→G·C spontaneous point replication error, incorporating into the DNA. According to our data, the tautomerisation reaction of the 2AP·T(WC)[66] base pair into the 2AP·T*$_{\downarrow}$(w) mispair containing T* mutagenic tautomer and back (Figs. 1, 2) causes the increasing of the population of the latest in comparison with the A·T*$_{\downarrow}$(w) mispair in 1887/219 times under normal conditions in the continuum with ε=1/ε=4 (accordingly, probability of the G·T*(WC) mismatch[67] formation in the next cycle of the DNA replication increases in the appropriate number of times) (Figs. 1, 2, Table 1 and Tables S1-S4). One of the basic reasons for this is that the weak C2H···N3 H-bond (1.41 kcal·mol^{-1}) in the A·T*$_{\downarrow}$(w) mispair is replaced by a much stronger classical N2H···N3 H-bond (5.27 kcal·mol^{-1}) in the 2AP·T*$_{\downarrow}$(w) mismatch (Figs. 1, 2, Table 1 and Table S1). In our opinion, the outlined above considerations are important confirmation of the adequacy of our model of the tautomeric non-stability of the A·T(WC) DNA base pair.

Finally, we would like to note that in addition to the pathways analyzed in details above, we have also detected two others A·T*$_{\uparrow}$(w)↔A·T(WC)↔A*·T$_{\downarrow}$(w) tautomerisation routes, controlled by the transition states that are H-bonded A^{-}·T^{+} tight ion pairs (Fig. 4 and Figs. S2, S3 and Tables S1-S4). In view of their extreme slowness ($\tau_{99.9\%}$=6.6·10^{10} and 9.1·10^{9} s, respectively (Table S4), that by orders exceeds time of the DNA replication in the cell[2]), these tautomerisation processes do not present any biological interest and can be attractive only from a theoretical point of view.

Mutagenic tautomeric wobbling of the canonical G·C Watson-Crick DNA base pair. We have managed to establish pioneeringly that classical G·C(WC) DNA base pair stay in the rapid (in comparison with the rate of the DNA replication in the cell[2]) tautomeric equilibrium with the four wobble (w) G·C*$_{\uparrow}$(w), G*·C$_{\downarrow}$(w), G·C*$_{\downarrow}$(w) and G*·C$_{\uparrow}$(w) pairs involving mutagenic tautomers of the DNA bases (Figs. 5, 7, Table 2

and Figs. S4, Tables S6-S10). Our QM computational experiment demonstrate that forced dissociation of these mispairs into monomers by the DNA-polymerase machinery would generate with necessity in the genome mutagenic tautomers of the DNA bases, that are able to cause spontaneous point mutations – replication errors.

It should be noted that all four G·C(WC)↔G·C*$_{\uparrow}$(w), G·C(WC)↔G*·C$_{\downarrow}$(w), G·C(WC)↔G·C*$_{\downarrow}$(w) and G·C(WC)↔G*·C$_{\uparrow}$(w) pathways of the mutagenic tautomerisation occur non-dissociatively (Fig. 8 and Fig. S9, Table S14) *via* the highly stable and highly polar transition states – tight $G^{+}\cdot C^{-}$ and $G^{-}\cdot C^{+}$ ion pairs – formed by the protonated and deprotonated bases (Figs. 5, 7). Bases in the aforementioned complexes, which are involved in these tautomerisation processes, are tightly bound in view of the negative values of the Gibbs free energy of interaction ($\Delta G_{int}<0$) (Table S9), which extremely high values (>100 kcal·mol^{-1}) are achieved at the TSs. This means that impact of the water molecules and stacking interactions[68-70] on the course of the considered processes is negligible.

This non-dissociative quantum mechanism of the sequential DPT in the nucleobase pairs is accompanied by their substantial rearrangement and sequential change of their intermolecular specific interactions (from 10 to 15 patterns), in particular AH···B, NH···HN H-bonds, attractive N···O/N van der Waals contacts and loosened A-H-B covalent bridges along the IRC of tautomerisation (Fig. 8 and Table S14).

This approach[12-16] has stand the test of time for others mispairs and was confirmed experimentally for the G·T(w)[9,10] and G·P(WC)[71,72] base mispairs.

It becomes clear from the comparison of the energy barriers of the mutagenic tautomerisation of the G and C bases in the isolated state[52], with the corresponding values in the composition of the G·C(WC) DNA base pair, respectively (Table 2 and Table S10), that complementary H-bonded bases catalyze the tautomeric transformation of each other.

The G·C(WC)↔G·C*$_{\uparrow}$(w) and G·C(WC)↔G*·C$_{\downarrow}$(w) tautomerisation processes occuring without involvement of the intermediates are significantly simpler than two others G·C(WC)↔G·C*$_{\downarrow}$(w) and G·C(WC)↔G*·C$_{\uparrow}$(w) (Figs. 5, 7, Table 2 and Table S10). These two latest processes proceed *via* the dynamically unstable intermediate – the G*·C*(L) Löwdin's (L) base mispair with Watson-Crick geometry (Figs. 5, 7, Table 2 and Table S10).

It should be emphasized that all four wobble G·C*$_{\uparrow}$(w), G*·C$_{\downarrow}$(w), G·C*$_{\downarrow}$(w) and G*·C$_{\uparrow}$(w) pairs do not tautomerise through the DPT along the neighboring intermolecular H-bonds, because the products of these reactions are dynamically unstable and short-lived structures, for which low-frequency intermolecular vibrations can't develop during their lifetimes (Figs. S5, S6 and Table S12). It also draws attention to the significant non-planarity of the $TS^{G^{-}\cdot C^{+}}_{G^{*}\cdot C^{*}(L)\leftrightarrow G\cdot C^{*}\downarrow(w)}$ (∠C6N1(G^{+})N3C4(C^{-})=±38.0º) and $TS^{G^{-}\cdot C^{+}}_{G^{*}\cdot C^{*}(L)\leftrightarrow G^{*}\cdot C\uparrow(w)}$ (∠C6N1(G^{-})N3C4(C^{+})=±40.2º) transition states and also of the wobble G·C*$_{\downarrow}$(w) (∠N1C2(G)N3C4(C*)=±42.3º) and G*·C$_{\uparrow}$(w) (∠C6N1(G*)N1C6(C)= ±31.0º) mispairs (Figs. S7, S8 and Table

S13). However, our simulation results show that it is not essential obstacle for the incorporation of these complexes into the structure of the DNA double helix.

Detailed physico-chemical analysis revealed that the wobble G·C*$_{\uparrow}$(w), G*·C$_{\downarrow}$(w), G·C*$_{\downarrow}$(w) and G*·C$_{\uparrow}$(w) mismatches can be successfully dissociated into the monomers by the replication machinery, since they fulfill all conditions, which are necessary to accomplish this purpose[19,21,23]. Thus, these tautomerised mispairs can be embedded well[58,59] into the structure of the DNA double helix owing to the fact that their interaction energy between the bases is lower than a similar value for the canonical G·C(WC) DNA base pair[21]; they have rather long lifetimes (21.5, 4.7, $4.9\cdot10^{-6}$ and $1.2\cdot10^{-8}$ s, respectively) and, in addition, the population of these tautomerised states of the canonical G·C(WC) DNA base pair – G·C*$_{\uparrow}$(w), G*·C$_{\downarrow}$(w), G·C*$_{\downarrow}$(w) and G*·C$_{\uparrow}$(w) – consist $1.4\cdot10^{-9}$, $5.3\cdot10^{-10}$, $4.5\cdot10^{-11}$ and $4.2\cdot10^{-11}$, respectively, in the continuum with ε=4[54,73], spanning the experimentally observed range of the frequencies of the spontaneous point mutations (10^{-11}-10^{-9})[55-57]. All these facts gathered together indicate, that the identified here for the first time intrinsic property of the G·C(WC) DNA base pair to stay in rapid tautomeric equilibrium with the wobble G·C*$_{\uparrow}$(w), G*·C$_{\downarrow}$(w), G·C*$_{\downarrow}$(w) and G*·C$_{\uparrow}$(w) pairs, is closely related to the occurrence of spontaneous point mutations – DNA replication errors.

Low population of the tautomerised states of the classical G·C(WC) base pair complicates their registration by the experimental physico-chemical methods. Meanwhile, by NMR spectroscopy it was fixed the conversion of the G·P(WC) base pair including 6H,8H-3,4-dihydropyrimido[4,5-c][1,2]oxazin-7-one (P) mutagen[6] into the wobble configuration and *vice versa* in the DNA composition[71,72]. In this case according to the presented above model, the population of the tautomerised G·P*$_{\uparrow}$(w) and G·P*$_{\downarrow}$(w) states increases by orders of magnitude in comparison with similar G·C*$_{\uparrow}$(w) and G·C*$_{\downarrow}$(w) states, attaining values $4.5\cdot10^{-3}$ and $1.4\cdot10^{-4}$, respectively (ε=4), which allows them to be recorded experimentally (Figs. 5, 6, Table 2 and Tables S8, S10, S11). Contrary, the population of two others tautomerised G*·P$_{\downarrow}$(w) and G*·P$_{\uparrow}$(w) states of the G·P(WC) base pair increases only in 0.2 and 74.2 times in comparison to analogical states of the G·C(WC) base pair in the continuum with ε=4 (Figs. 5 and 6) and that's why they remain inaccessible for experimental observation. All this testifies to the adequacy of the presented above representations of the mutagenic tautomerisation of the G·C(WC) base pair.

Obtained in this paper results significantly deepen and elaborate classical Watson-Crick tautomeric hypothesis[3] that links the nature of spontaneous point mutations occuring during DNA replication with ability of the nucleotide bases to transfer from the canonical into the mutagenic tautomeric form. Developed representations significantly extend possibility for the rational design of the promutagens by using nucleobases with targeted action that are interesting from the point of view of both anticancer and antiviral therapy.

Rather promising task for the future is clarification of the role of the local protein environment of the replisome[74] on the discussed above tautomeric equilibrium and also possible evolutionary value of the mutagenic tautomeric mobility of the canonical Watson-Crick DNA base pairs.

Acknowledgements We gratefully appreciate technical support and computational facilities from 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. This work was supported by the grant of the President of Ukraine to support scientific research of young scientists for 2015 year from the State Fund for Fundamental Research of Ukraine (project № GP/F61/028), by the Grant of the NAS of Ukraine for young scientists for 2015-2016 years and by the Scholarship of the President of Ukraine for young scientists for the years 2014-2016 given to O.O.B.

References.

1. Freese, E.B. On the molecular explanation of spontaneous and induced mutations. *Brookhaven Symp. Biol.*, 1959, **12**, 63-75.
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. Watson, J. D., & Crick, F. H. C. The Structure of DNA. *Cold Spring Harb. Symp. Quant. Biol.*, 1953, **18**, 123-131.
4. Danilov, V.I., Anisimov, V.M., Kurita, N., & Hovorun, D. MP2 and DFT studies of the DNA rare base pairs: the molecular mechanism of the spontaneous substitution mutations conditioned by tautomerism of bases. *Chem. Phys. Lett.*, 2005, **412**, 285-293.
5. Fonseca Guerra, C., Bickelhaupt, F.M., Saha, S., & Wang, F. Adenine tautomers: relative stabilities, ionization energies, and mismatch with cytosine. *J. Phys. Chem. A*, 2006, **110**, 4012-4020.
6. Harris, V. H. *et al.* The effect of tautomeric constant on the specificity of nucleotide incorporation during DNA replication: support for the rare tautomer hypothesis of substitution mutagenesis. *J. Mol. Biol.*, 2003, **326**, 1389‑1401.
7. 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.
8. Wang, W., Hellinga, H. W., & Beese, L. S. Structural evidence for the rare tautomer hypothesis of spontaneous mutagenesis. *Proc. Natl. Acad. Sci. USA,* 2011, **108**, 17644−17648.
9. Kimsey, I.J., Petzold, K., Sathyamoorthy, B., Stein, Z.W., & Al-Hashimi, H.M. Visualizing transient Watson-Crick-like mispairs in DNA and RNA duplexes. *Nature,* 2015, **519**, 315-320.
10. Kimsey, I. J., Zhou, H., Alvey, H., & Al-Hashimi, H. M. 120 Role of dynamic base pair polymorphism in the central dogma of molecular biology. *J. Biomol. Struct. Dynam.,* 2015, **33**(sup1), 75-76.
11. Brovarets', O.O., & Hovorun, D.M. Physicochemical mechanism of the wobble DNA base pairs Gua·Thy and Ade·Cyt transition into the mismatched base pairs Gua*·Thy and Ade·Cyt* formed by the mutagenic tautomers. *Ukr. Bioorg. Acta,* 2009, **8**, 12-18.
12. Brovarets', O. O., & Hovorun, D. M. Tautomeric transition between wobble A·C DNA base mispair and Watson-Crick-like A·C* mismatch: miscrostructural mechanism and biological significance. *Phys. Chem. Chem. Phys.,* 2015, **17**, 15103-15110.
13. Brovarets', O. O., & Hovorun, D. M. How many tautomerisation pathways connect Watson-Crick-like G*·T DNA base mispair and wobble mismatches? *J. Biomol. Struct. & Dynam.*, 2015, **33**, 2297-2315.
14. Brovarets', O. O., & Hovorun, D. M. Novel physico-chemical mechanism of the mutagenic tautomerisation of the Watson-Crick-like A·G and C·T DNA base mispairs: a quantum-chemical picture. *RSC Adv.,* 2015, **5**, 66318-66333.
15. Brovarets', O. O., & Hovorun, D. M. A novel conception for spontaneous transversions caused by homo-pyrimidine DNA mismatches: a QM/QTAIM highlight. *Phys. Chem. Chem. Phys.,* 2015, **17**, 21381-21388.
16. Brovarets', O. O., & Hovorun, D. M. Wobble↔Watson-Crick tautomeric transitions in the homo-purine DNA mismatches: a key to the intimate mechanisms of the spontaneous transversions. *J. Biomol. Struct. & Dynam.*, 2015, DOI: 10.1080/07391102.2015.1077737.

17. Jacquemin, D., Zúñiga, J., Requena, A., & Céron-Carrasco, J. P. Assessing the importance of proton transfer reactions in DNA. *Acc. Chem. Res.*, 2014, **47***,* 2467-2474.
18. Löwdin, P.-O. Proton tunneling in DNA and its biological implications. *Rev. Mod. Phys.,* 1963, **35**, 724-732.
19. 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.
20. Brovarets', O.O., & Hovorun, D.M. Proton tunneling in the A·T Watson-Crick DNA base pair: myth or reality? *J. Biomol. Struct. & Dynam.,* 2015, DOI: 10.1080/07391102.2015.1092886.
21. 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.,* 2014, **32**, 1474-1499.
22. Padermshoke, A., Katsumoto, Y., Masaki, R., & Aida, M. Thermally induced double proton transfer in GG and wobble GT base pairs: a possible origin of the mutagenic guanine. *Chem. Phys. Lett.*, 2008, **457**, 232–236.
23. Brovarets', O.O., Zhurakivsky, R.O., & Hovorun, D.M. DPT tautomerisation of the wobble guanine·thymine DNA base mispair is not mutagenic: QM and QTAIM arguments. *J. Biomol. Struct. & Dynam.,* 2015, **33**, 674-689.
24. Strazewski, P., & Tamm, C. Replication experiments with nucleotide base analogues. *Ang. Chemie Int. Ed.,* 1990, **29**, 36–57.
25. Brovarets' O.O., Yurenko Ye.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.
26. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., & Cheeseman, J.R., … Pople, J.A. (2010). *GAUSSIAN 09* (Revision B.01). Wallingford CT: Gaussian Inc.
27. Tirado-Rives, J., & Jorgensen, W.L. Performance of B3LYP Density Functional Methods for a large set of organic molecules. *J. Chem. Theory Comput.* **4***,* 297–306 (2008).
28. Parr, R.G., & Yang, W. *Density-functional theory of atoms and molecules*. Oxford: Oxford University Press (1989).
29. 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.* **37**, 785-789 (1988).
30. 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.* **31**, 1297–1311 (2010).
31. 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.* **578**, 126-132 (2013).
32. 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.* **34**, 2577-2590 (2013).
33. Brovarets', O.O., Zhurakivsky, R.O., & Hovorun, D.M. Is the DPT tautomerisation 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.* **35**, 451-466 (2014).
34. Brovarets', O.O., & Hovorun, D.M. How the long G·G* Watson-Crick DNA base mispair comprising keto and enol tautomers of the guanine tautomerises? The results of the QM/QTAIM investigation. *Phys. Chem. Chem. Phys.* **6**, 15886-15899 (2014).
35. 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* **114**, 1454-1461 (2012).

36. 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.* **17**, 49–56 (1996).
37. Atkins, P.W. *Physical chemistry*. Oxford: Oxford University Press (1998).
38. Hratchian, H.P., & Schlegel, H.B. (2005). 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* (pp. 195-249). Amsterdam: Elsevier.
39. Brovarets', O.O., & Hovorun, D.M. DPT tautomerisation of the G·A_{syn} and A*·G^*_{syn} DNA mismatches: A QM/QTAIM combined atomistic investigation. *Phys. Chem. Chem. Phys.*, **16**, 9074-9085 (2014).
40. 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.* **16**, 3715-3725 (2014).
41. Brovarets', O.O., & Hovorun, D.M. Does the G·G^*_{syn} DNA mismatch containing canonical and rare tautomers of the guanine tautomerise through the DPT? A QM/QTAIM microstructural study. *Mol. Phys.* **112**, 3033-3046 (2014).
42. Frisch, M.J., Head-Gordon, M., & Pople, J.A. Semi-direct algorithms for the MP2 energy and gradient. *Chem. Phys. Lett.* **166***,* 281-289 (1990).
43. Hariharan, P.C., & Pople, J.A. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta* **28**, 213−222 (1973).
44. 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.* **72**, 650−654 (1980).
45. 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.* **96**, 6796−6806 (1992).
46. Bader, R.F.W. *Atoms in molecules: A quantum theory*. Oxford: Oxford University Press, 1990.
47. Matta, C.F. Modeling biophysical and biological properties from the characteristics of the molecular electron density, electron localization and delocalization matrices, and the electrostatic potential. *J. Comput. Chem.,* 2014, **35***,* 1165-1198.
48. 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.* **32***,* 993-1022 (2014).
49. Brovarets', O. O., Yurenko, Y. P., & Hovorun, D. M. The significant role of the intermolecular CH···O/N hydrogen bonds in governing the biologically important pairs of the DNA and RNA modified bases: a comprehensive theoretical investigation. *J. Biomol. Struct. & Dynam.* **33**, 1624-1652 (2015).
50. Zoete V., Meuwly M. Double proton transfer in the isolated and DNA-embedded guanine-cytosine base pair. *J. Chem. Phys.* **121**, 4377-4388 (2004).
51. Kimsey, I.J., Petzold, K., Sathyamoorthy, B., Stein, Z.W., & Al-Hashimi, H.M. Visualizing transient Watson-Crick-like mispairs in DNA and RNA duplexes. *Nature* **519**, 315-320 (2015).
52. Brovarets', O.O., & Hovorun, D.M. How stable are the mutagenic tautomers of DNA bases? *Biopol. Cell,* 2010, **26***,* 72–76.
53. Lin, Y., Wang, H., Wu, Y., Gao, S., & Schaefer III, H.F. Proton-transfer in hydrogenated guanine–cytosine trimer neutral species, cations, and anions embedded in B-form DNA. *Phys. Chem. Chem. Phys.,* 2014, **16**, 6717-6725.
54. Petrushka, J., Sowers, L.C., & Goodman, M. Comparison of nucleotide interactions in water, proteins, and vacuum: model for DNA polymerase fidelity. *Proc. Natl. Acad. Sci. USA,* 1986, **83**, 1559-1562.
55. Drake, J.W., Charlesworth, B., Charlesworth, D., & Crow, J.F. Rates of spontaneous mutation. *Genetics,* 1998, **148,** 1667–1686.

56. Lynch, M. Rate, molecular spectrum, and consequences of human mutation. *Proc. Natl. Acad. Sci. USA,* 2010, **107**, 961-968.
57. Fijalkowska, I.J., Schaaper, R.M., & Jonczyk, P. DNA replication fidelity in *Escherichia coli*: a multi-DNA polymerase affair. *FEMS Microbiol. Rev.,* 2012, **36**, 1105–1121.
58. Brovarets', O.O., & Hovorun, D.M. Molecular logic of the spontaneous point mutagenesis: variation on a theme … *Ukr. Bioorg. Acta*, 2014, **12**, 48-55.
59. Westhof, E. Isostericity and tautomerism of base pairs in nucleic acids. *FEBS Lett.,* 2014, **588,** 2464-2469.
60. Osborn, M., Person, S., Phillips, S., & Funk, F. A determination of mutagen specificity in bacteria using nonsense mutants of bacteriophage T4. *J. Mol. Biol.,* 1967, **26**, 437-447.
61. Freese, E.B. The mutagenic effect of hydroxyaminopurine derivatives on phage T4. *Mutat. Res.,* 1968, **5**, 299-301.
62. Goodman, M. F., Hopkins, R., & Gore, W. C. 2-Aminopurine-induced mutagenesis in T4 bacteriophage: a model relating mutation frequency to 2-aminopurine incorporation in DNA. *Proc. Natl. Acad. Sci. USA*., 1977, **74**, 4806-4810.
63. Ronen, A. 2-Aminopurine. *Mutat. Res.,* 1979, **75**, 1-47.
64. Watanabe, S. M., & Goodman, M. F. On the molecular basis of transition mutations: frequencies of forming 2-aminopurine·cytosine and adenine·cytosine base mispairs *in vitro*. *Proc. Natl. Acad. Sci. USA*., 1981, **78**, 2864-2868.
65. Champe, S. P., Benzer, S. Reversal of mutant phenotypes by 5-fluorouracil: an approach to nucleotide sequences in messenger-RNA. *Proc. Natl. Acad. Sci. USA*., 1962, **48**, 522-532.
66. Freese, E. The specific mutagenic effect of base analogue on phage T4. *J. Mol. Biol.,* 1959, **1**, 87-105.
67. Brovarets', O. O., & Hovorun, D. M. The nature of the transition mismatches with Watson-Crick architecture: the G*·T or G·T* DNA base mispair or both? A QM/QTAIM perspective for the biological problem. *J. Biomol. Struct. Dyn.,* 2015, **33***,* 925-945.
68. Lin, Y., Wang, H., Wu, Y., Gao, S., & Schaefer III, H.F. Proton-transfer in hydrogenated guanine–cytosine trimer neutral species, cations, and anions embedded in B-form DNA. *Phys. Chem. Chem. Phys.,* 2014, **16**, 6717-6725.
69. Matta, C.F., Castillo, N., & Boyd, R.J. Extended weak bonding interactions in DNA: π-stacking (base-base), base-backbone, and backbone-backbone interactions. *J. Phys. Chem. B,* 2006, **110**, 563-578.
70. Cerón-Carrasco, J. P., Zúñiga, J., Requena, A., Perpète, E. A., Michaux C., & Jacquemin, D. Combined effect of stacking and solvation on the spontaneous mutation in DNA. *Phys. Chem. Chem. Phys.,* 2011, **13**, 14584-14589.
71. Nedderman, A.N., Stone, M.J., Williams, D.H., Lin, P.K.T., & Brown D.M. Molecular basis for methoxyamine-initiated mutagenesis: ^{1}H nuclear magnetic resonance studies of oligonucleotide duplexes containing base-modified cytosine residues. *J. Mol. Biol.,* 1993, **230**, 1068-1076.
72. Nedderman, A.N., Stone, M.J., Lin, P.K.T., Brown, D.M., & Williams, D.H. Base pairing of cytosine analogues with adenine and guanine in oligonucleotide duplexes: evidence for exchange between Watson-Crick and wobble base pairs using ^{1}H NMR spectroscopy. *J. Chem. Soc. Chem. Commun.,* 1991, 1357–1359.
73. Mertz, E. L., & Krishtalik, L. I. Low dielectric response in enzyme active site. *Proc. Natl Acad. Sci. USA*, 2000, **97**, 2081–2086.
74. Baker, T.A., & Bell, S.P. Polymerases and the replisome: machines within machines. *Cell,* 1998, **92**, 295-305.

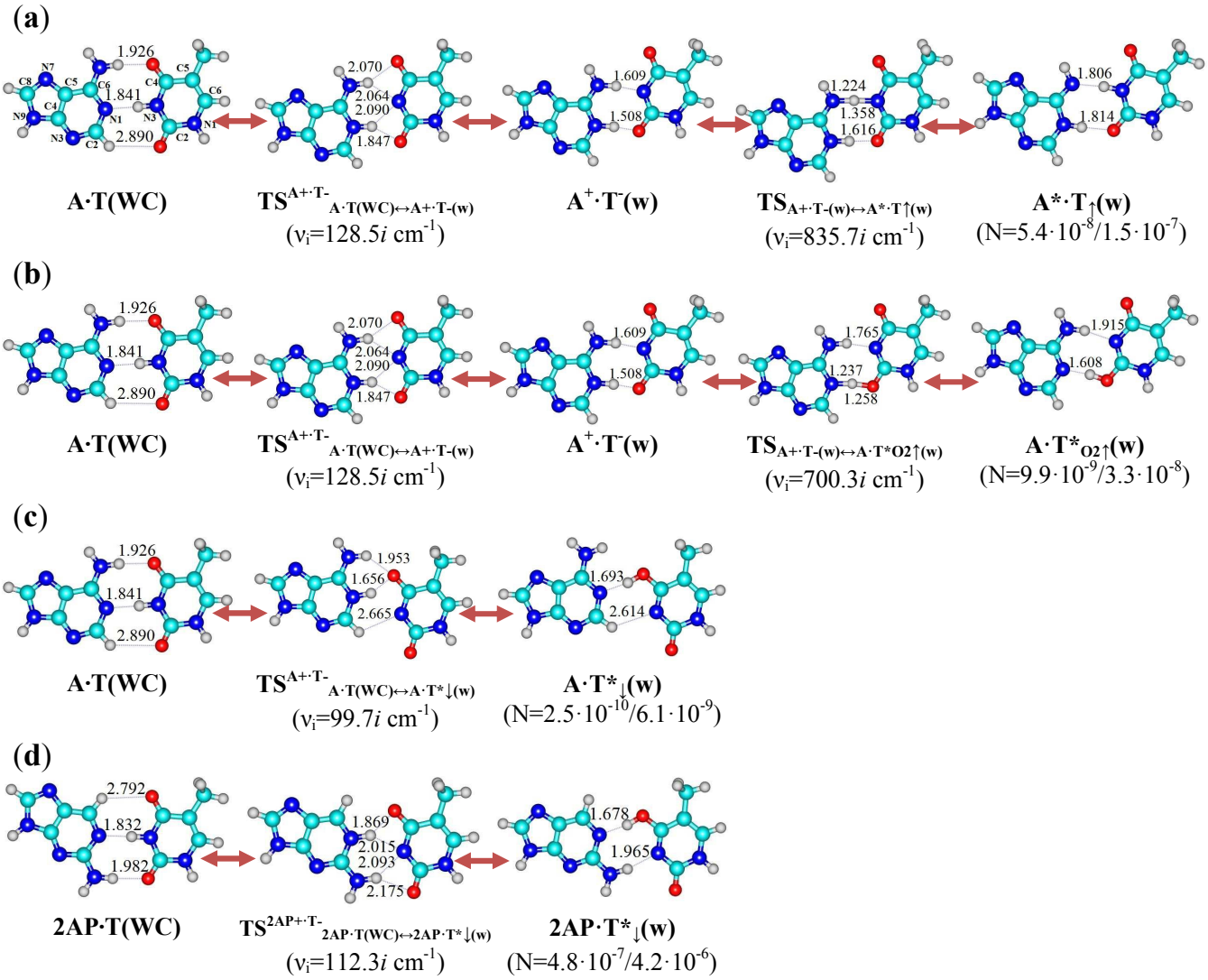
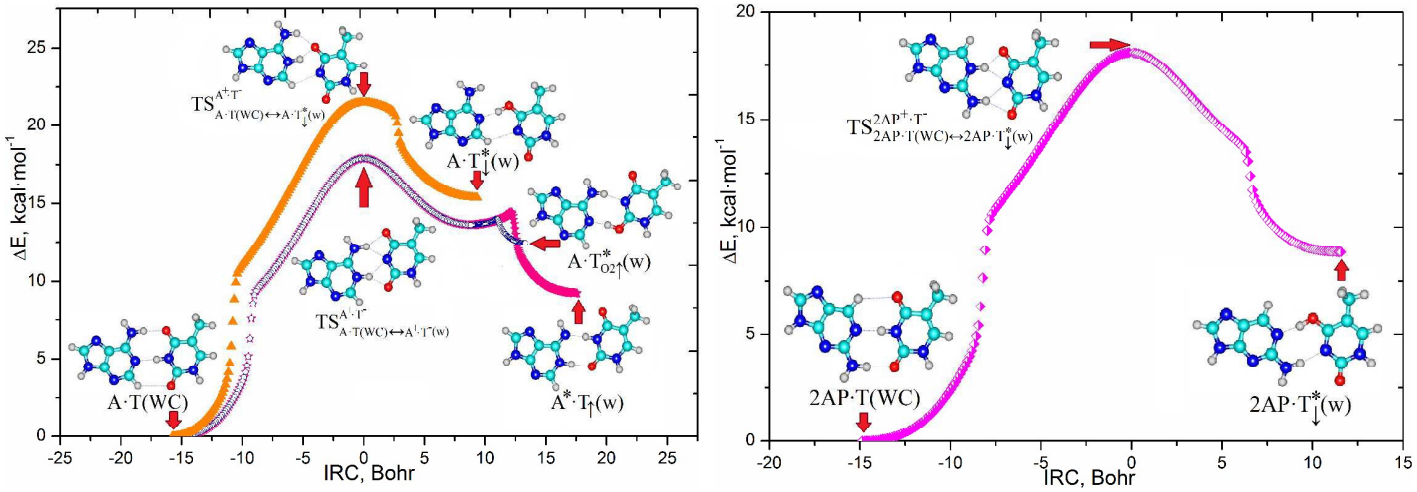


Figure 1. Stationary structures optimized at the B3LYP/6-311++G(d,p) level of theory (ε=1) on the reaction pathways of the (**a**) A·T(WC)↔A^+·T^-(w)↔A*·$T_{\uparrow}$(w), (**b**) A·T(WC)↔A^+·T^-(w)↔A·T*$_{O2\uparrow}$(w), (**c**) A·T(WC)↔A·T*$_{\downarrow}$(w) and (**d**) 2AP·T(WC)↔2AP·T*$_{\downarrow}$(w) tautomerisations. All transition states of these processes represent itself tight A^+·T^-/$2AP^+$·T^- ion pairs, geometry of which is no longer Watson-Crick, but is not yet wobble. Populations N of the wobble mispairs containing rare tautomers are presented below them in brackets (MP2/aug-cc-pVDZ//B3LYP/6-311++G(d,p) level of theory in the continuum with ε=1/ε=4 at T=298.15 K). Dotted lines indicate AH···B H-bonds and continuos – loosened A-H-B covalent bridges (their lengths in the continuum with ε=1 are presented in angstroms; for more detailed physico-chemical characteristics of the H-bonds see Table S1); carbon atoms are in light-blue, nitrogen – in dark-blue, hydrogen – in grey and oxygen – in red; ν_i – imaginary frequency; numeration of atoms depicted on the A·T(WC) base pair is standard.



(**a**) (**b**)

Figure 2. Energetic profiles of the mutagenic tautomeric wobbling of the A·T(WC) and 2AP·T(WC) DNA base pairs due to their tautomeric transformations *via* the (**a**) A·T(WC)↔$A^{+}\cdot T^{-}$(w)↔A*·$T_{\uparrow}$(w), A·T(WC)↔$A^{+}\cdot T^{-}$(w)↔A·T*$_{O2\uparrow}$(w), A·T(WC)↔A·T*$_{\downarrow}$(w) and (**b**) 2AP·T(WC)↔2AP·T*$_{\downarrow}$(w) tautomerisation pathways depicted together with key stationary structures on the potential energy hypersurface (see also Figure 1). These tautomeric conversions occur through the initial nipping off the migrating proton from the T base and joining it to the A complementary base, then shifting and significant rebuilding of the bases relative to each other within the base pair into the major or minor groove sides of the DNA helix and further addition of the mobile or other acidic proton to the neighboring nitrogen or oxygen atoms of the T base. The A·T(WC)↔A*·$T_{\uparrow}$(w) and A·T(WC)↔A·T*$_{O2\uparrow}$(w) tautomerisation reactions proceed *via* the dynamically unstable intermediate – the $A^{+}\cdot T^{-}$ (w) tight ion pair with wobble architecture, exactly at which bifurcation occurs into two different pathways. All transition states of these processes represent itself tight $A^{+}\cdot T^{-}$/$2AP^{+}\cdot T^{-}$ ion pairs, geometry of which is no longer Watson-Crick, but is not yet wobble. These data are obtained by following IRC at the B3LYP/6-311++G(d,p) level of theory in the continuum with ε=1. For the detailed designations of the intermolecular bonds, atoms, their numeration and complex populations see Fig. 1 caption.

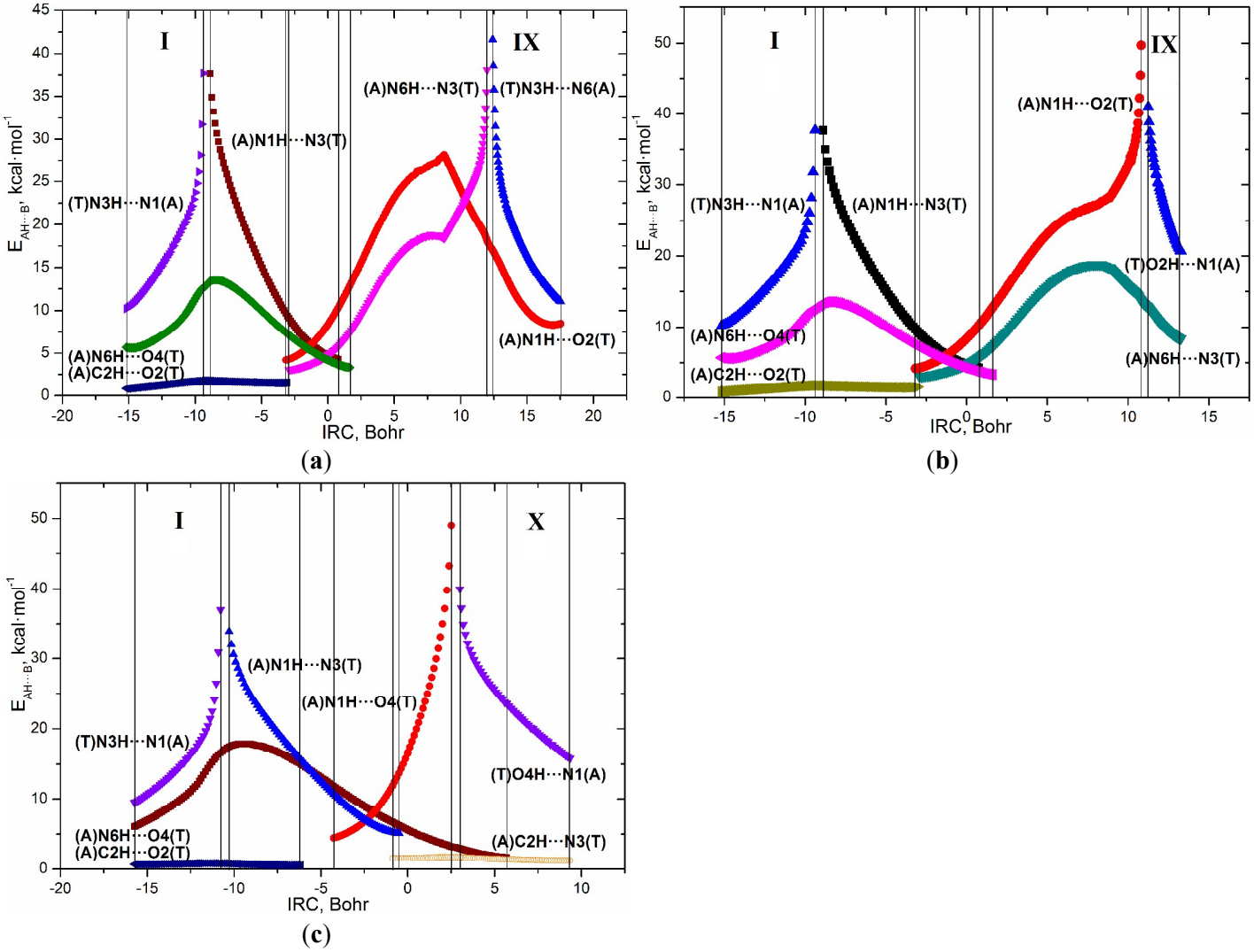


Figure 3. Exchange of the patterns of the intermolecular AH···B H-bonds along the IRC of the biologically important tautomerisations of the A·T(WC) DNA base pair *via* the sequential DPT accompanied with the structural rearrangement of the bases: (**a**) A·T(WC)↔A*·T$_{\uparrow}$(w), (**b**) A·T(WC)↔A·T*$_{O2\uparrow}$(w) and (**c**) A·T(WC)↔A·T*$_{\downarrow}$(w) (B3LYP/6-311++G(d,p) level of theory (ε=1)). Here and below perpendicular lines denote partition of all range into different regions. Detailed numerical characteristics of the unique patterns of the specific intermolecular interactions, which sequential change accompanies non-dissociative processes of tautomerisation, are presented in Table S5.

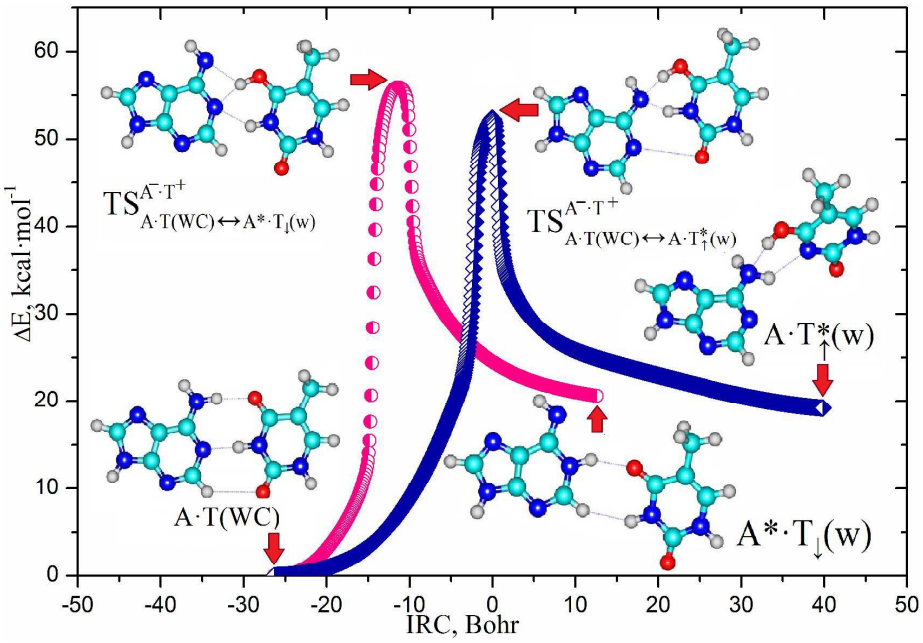


Figure 4. Energetic profiles combined with the stationary points of the high-energy (**a**) A·T(WC)↔A·T*$_{\uparrow}$(w) and (**b**) A·T(WC)↔A*·T$_{\downarrow}$(w) reaction pathways of the mutagenic tautomeric variability of the A·T(WC) DNA base pair proceeding *via* the initial nipping off the migrating proton from the A base and joining it to the T complementary base accompanied with structural rearrangements of the bases relative to each other within the base pair into the major or minor groove sides of the DNA helix and terminal transfer of the mobile or other acidic proton to the N1/N6 nitrogen atoms of the A base (see also Figs. S2, S3 and Tables S1-S4). These data are obtained by following IRC at the B3LYP/6-311++G(d,p) level of theory in the continuum with ε=1. For the detailed designations of the intermolecular bonds, atoms and their numeration see Fig. 1 caption.

Table 1. Energetic and kinetic characteristics of the biologically important A·T(WC)↔A*·$T_{\uparrow}$(w), A·T(WC)↔A·$T^*_{O2\uparrow}$(w), A·T(WC)↔A·$T^*_{\downarrow}$(w) and 2AP·T(WC)↔2AP·$T^*_{\downarrow}$(w) tautomerisations obtained at the MP2/aug-cc-pVDZ//B3LYP/6-311++G(d,p) level of theory in the continuum with ε=1/ε=4 (see also Tables S1-S4).

Tautomerisation	ε[a]	ΔG[b]	ΔE[c]	$\Delta\Delta G_{TS}$[d]	$\Delta\Delta E_{TS}$[e]	ΔΔG[f]	ΔΔE[g]	$\tau_{99.9\%}$[h]	τ[i]	N[j]
A·T(WC)↔A*·$T_{\uparrow}$(w)	1	9.90	9.59	16.72	16.02	6.82	6.43	$1.10\cdot10^{-7}$	$1.59\cdot10^{-8}$	$5.4\cdot10^{-8}$
	4	9.32	8.78	14.41	13.56	5.09	4.78	$9.59\cdot10^{-9}$	$8.62\cdot10^{-10}$	$1.5\cdot10^{-7}$
A·T(WC)↔A·$T^*_{O2\uparrow}$(w)	1	10.91	11.25	16.72	16.02	5.81	4.77	$2.02\cdot10^{-8}$	$2.92\cdot10^{-9}$	$9.9\cdot10^{-9}$
	4	10.19	10.48	14.41	13.56	4.22	3.07	$1.36\cdot10^{-9}$	$1.97\cdot10^{-10}$	$3.3\cdot10^{-8}$
A·T(WC)↔A·$T^*_{\downarrow}$(w)	1	13.08	14.84	20.28	20.41	7.20	5.57	$2.09\cdot10^{-7}$	$3.03\cdot10^{-8}$	$2.5\cdot10^{-10}$
	4	10.69	12.47	16.03	16.42	5.34	3.94	$9.20\cdot10^{-9}$	$1.33\cdot10^{-9}$	$6.1\cdot10^{-9}$
2AP·T(WC)↔2AP·$T^*_{\downarrow}$(w)	1	8.62	8.33	18.66	17.53	10.04	9.21	$2.53\cdot10^{-5}$	$3.66\cdot10^{-6}$	$4.8\cdot10^{-7}$
	4	7.33	7.79	15.48	14.72	8.15	6.93	$1.04\cdot10^{-6}$	$1.50\cdot10^{-7}$	$4.2\cdot10^{-6}$

[a]The dielectric constant; [b]The Gibbs free energy of the product relatively the reactant of the tautomerisation reaction (T=298.15 K), kcal·mol^{-1}; [c]The electronic energy of the product relatively the reactant of the tautomerisation reaction, kcal·mol^{-1}; [d]The Gibbs free energy barrier for the forward reaction of tautomerisation, kcal·mol^{-1}; [e]The electronic energy barrier for the forward reaction of tautomerisation, kcal·mol^{-1}; [f]The Gibbs free energy barrier for the reverse reaction of tautomerisation, kcal·mol^{-1}; [g]The electronic energy barrier for the reverse reaction of tautomerisation, kcal·mol^{-1}; [h]The time necessary to reach 99.9% of the equilibrium concentration between the reactant and the product of the tautomerisation reaction, s; [i]The lifetime of the product of the tautomerisation reaction, s; [j]Populations of the wobble mispairs containing mutagenic tautomers. See also summary Table S2 for the Gibbs and electronic energies of the mispairs and TSs relatively the global minimum – the A·T(WC) DNA base pair.

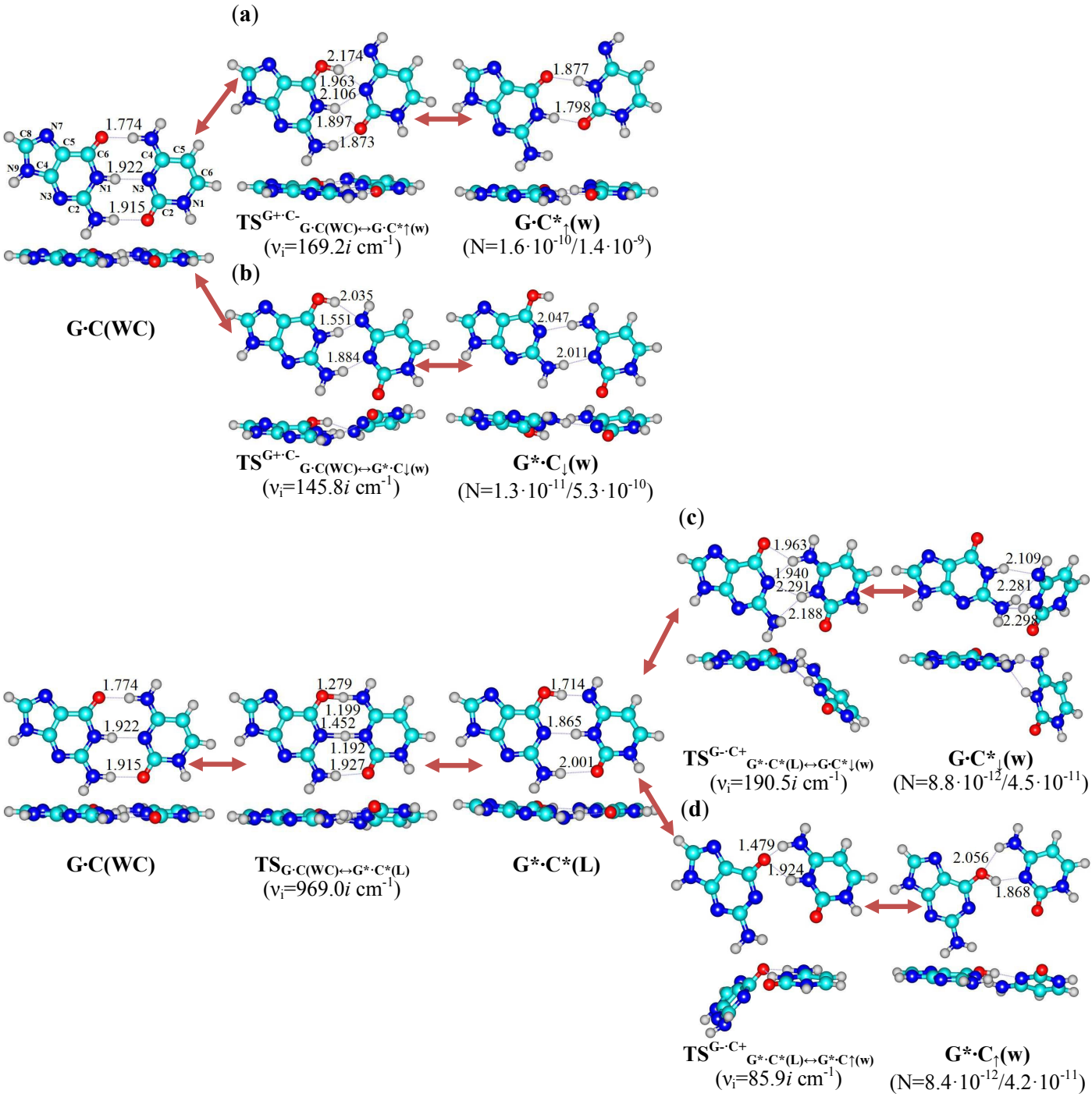


Figure 5. Stationary structures determined on the potential energy hypersurface of the canonical G·C(WC) DNA base pair and pathways of their tautomeric interconversions: calculations at the B3LYP/6-311++G(d,p) level of theory (ε=1). The (**a**) G·C(WC)↔G·C*↑(w), (**b**) G·C(WC)↔G*·C↓(w), (**c**) G·C(WC)↔G*·C*(L)↔G·C*↓(w) and (**d**) G·C(WC)↔G*·C*(L)↔G*·C↑(w) tautomerisations occur through the initial nipping off the migrating proton from the G or C base, then shifting and significant rebuilding of the bases relative to each other within the base pair into the major or minor groove sides of the DNA helix and further addition of the mobile or other acidic proton to the neighboring nitrogen or oxygen atoms of the complementary base. All transition states of these processes represent itself tight $G^+ \cdot C^-$ or $G^- \cdot C^+$ ion pairs, geometry of which is no longer Watson-Crick, but is not yet wobble. Perpendicular projections of all complexes are given below them. For the detailed designations of the intermolecular bonds, atoms, their numeration and complex populations see Fig. 1 caption.

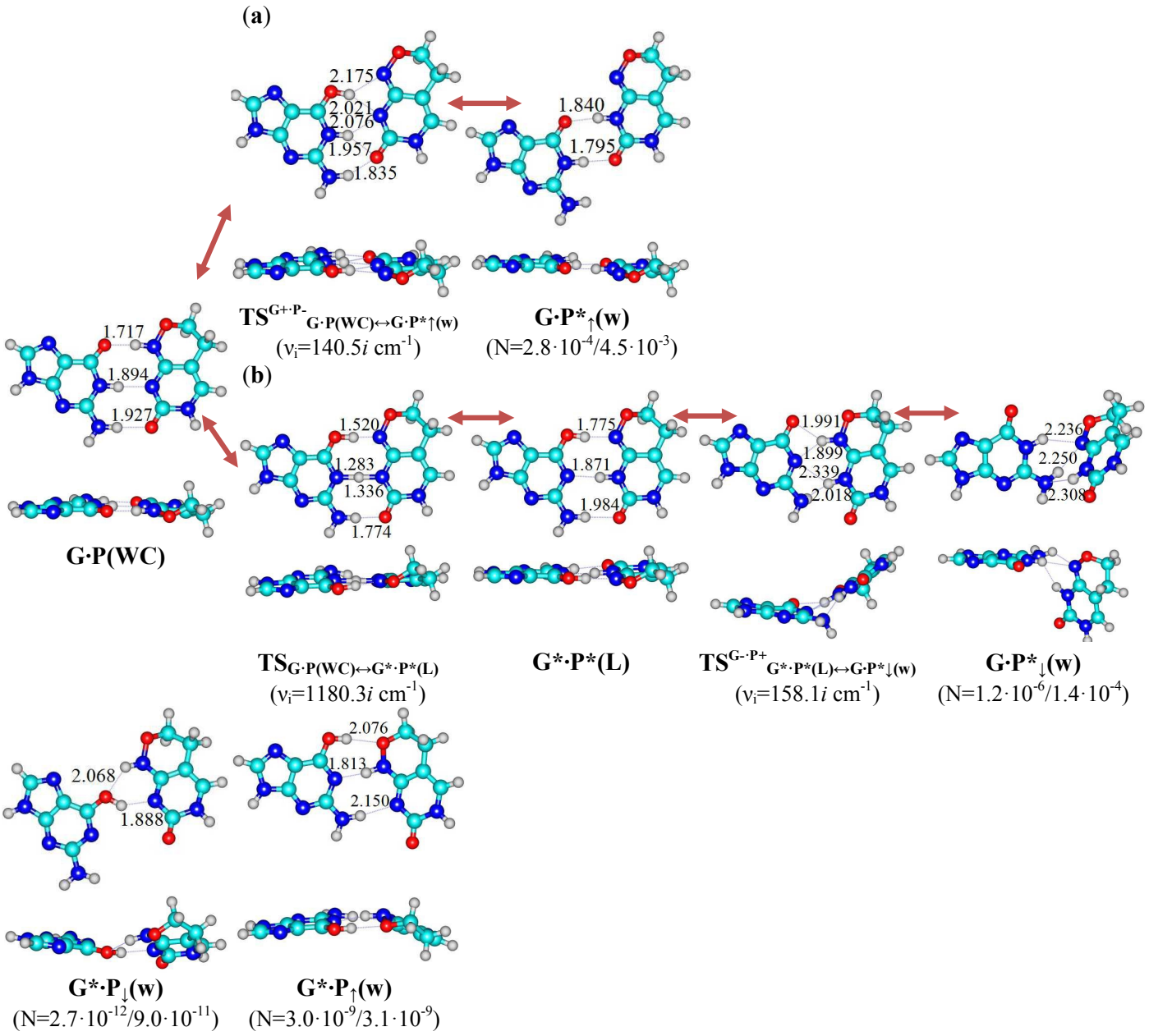


Figure 6. Stationary structures determined on the potential energy hypersurface of the G·P(WC) base pair involving P cytosine analogue and their tautomerisation pathways into the mispairs containing mutagenic tautomer of P: calculations at the B3LYP/6-311++G(d,p) level of theory (ε=1). The (**a**) G·P(WC)↔G·P*$_\uparrow$(w) and (**b**) G·P(WC)↔G·P*$_\downarrow$(w) tautomerisations occur through the initial nipping off the migrating proton from the G or P base, then shifting and significant rebuilding of the bases relative to each other within the base pair into the major or minor groove sides of the DNA helix and further addition of the mobile or other acidic proton to the neighboring nitrogen or oxygen atoms of the complementary base. All transition states of these processes represent itself tight $G^+\cdot P^-$ or $G^-\cdot P^+$ ion pairs, geometry of which is no longer Watson-Crick, but is not yet wobble. Perpendicular projections of all complexes are given below them. For the designation of the intermolecular bonds, atoms, their numeration and complex populations see Fig. 1 caption.

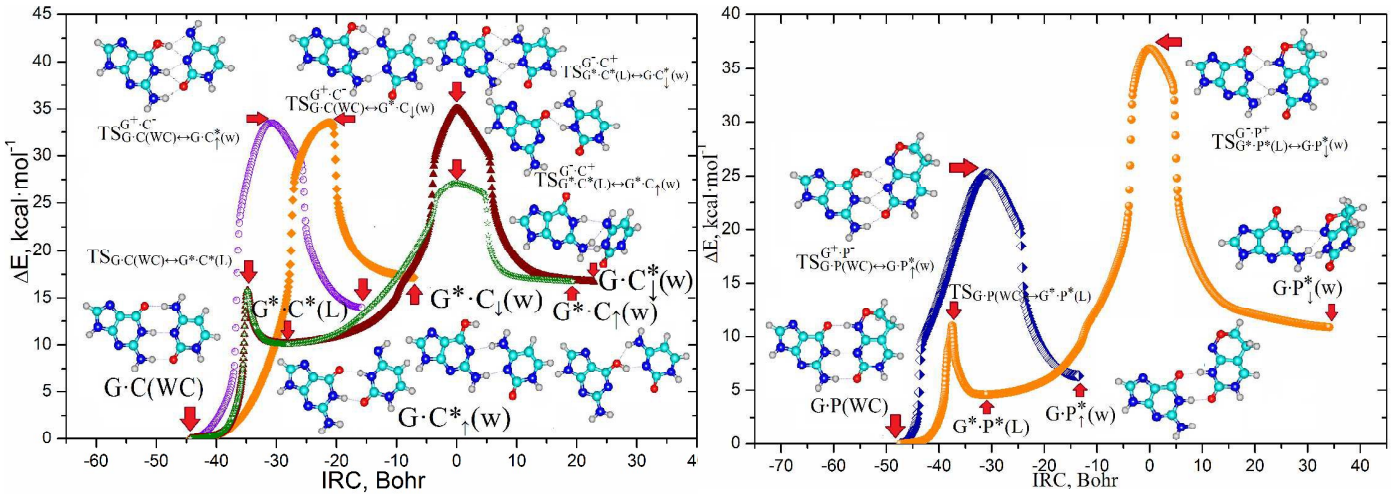


Figure 7. Diversity of the reaction pathways of the (**a**) G·C(WC)↔G·C*↑(w), G·C(WC)↔G*·C↓(w), G·C(WC)↔G·C*↓(w), G·C(WC)↔G*·C↑(w) and (**b**) G·P(WC)↔G·P*↑(w), G·P(WC)↔G·P*↓(w) mutagenic wobbling tautomerisations depicted together with key stationary structures on the potential energy hypersurface (see also Figs. 5 and 6). The G·C(WC)↔G·C*↓(w), G·C(WC)↔G*·C↑(w) and G·P(WC)↔G·P*↓(w) processes proceed *via* the short-lived and dynamically stable intermediates – the G*·C*(L) and G*·P*(L) Löwdin's base mispairs with Watson-Crick geometry. Presented data are obtained by following IRC at the B3LYP/6-311++G(d,p) level of theory in the continuum with ε=1. Various routes of the tautomerisation reactions are depicted in different colors. Peaks curves correspond to the transition states of the tautomerisation processes, while minima – to the base pairs with Watson-Crick or wobble architectures. For more detailed physico-chemical characteristics of these tautomeric transformations see Table 2 and Tables S6-S14.

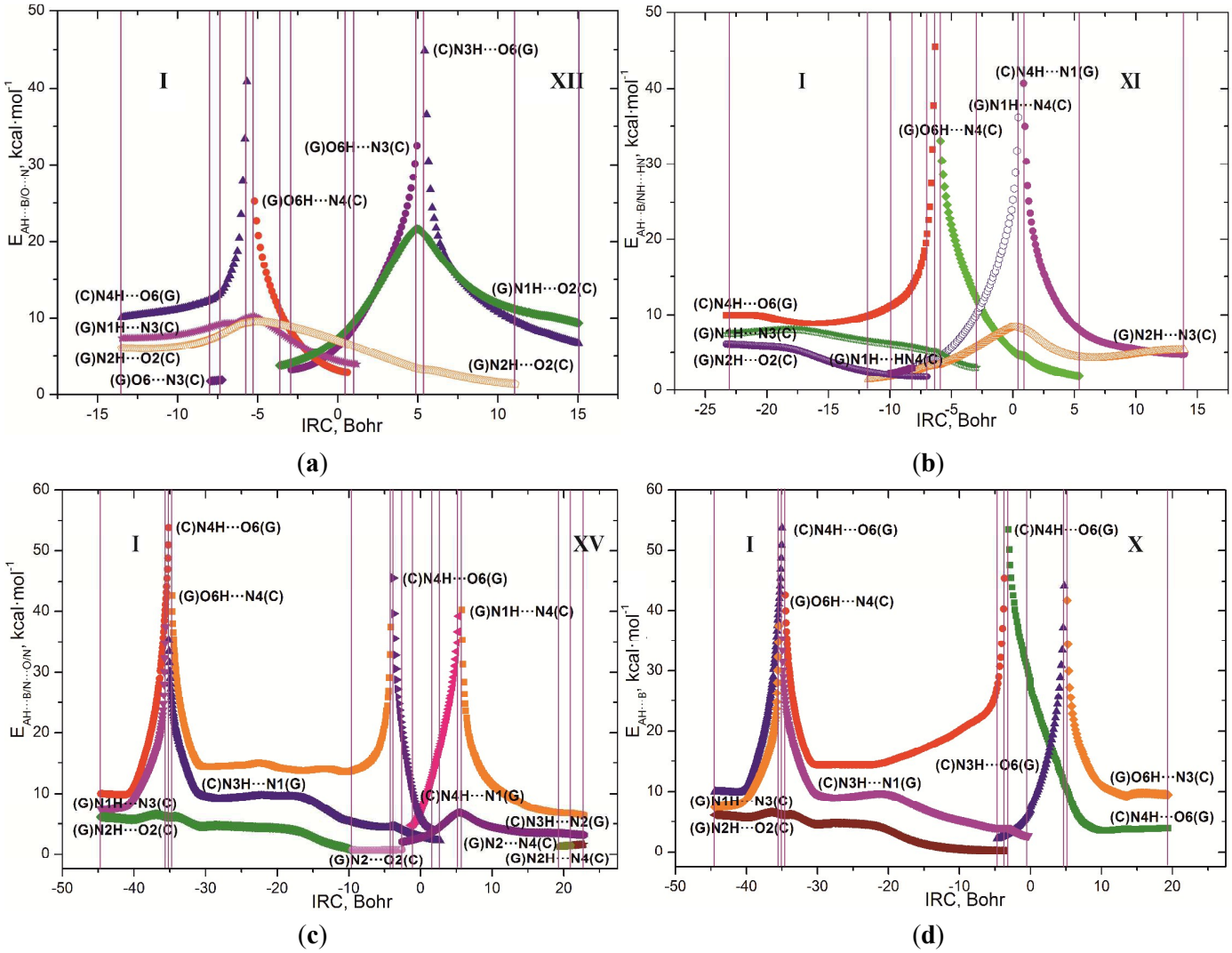


Figure 8. Exchange of the patterns of the intermolecular AH···B, NH···HN H-bonds and attractive N···O/N van der Waals contacts along the IRC of the biologically important tautomerisations of the G·C(WC) DNA base pair *via* the sequential DPT accompanied with structural rearrangements of the bases relative to each other: (**a**) G·C(WC)↔G·C*$_{\uparrow}$(w), (**b**) G·C(WC)↔G*·C$_{\downarrow}$(w), (**c**) G·C(WC)↔G*·C*(L)↔G·C*$_{\downarrow}$(w) and (**d**) G·C(WC)↔G*·C*(L)↔G*·C$_{\uparrow}$(w) (B3LYP/6-311++G(d,p) level of theory (ε=1)). Detailed numerical characteristics of the unique patterns of the specific intermolecular interactions, which sequential changing accompanies non-dissociative processes of tautomerisation, are presented in Table S14.

Table 2. Comprehensive energetic and kinetic parameters of the biologically important G·C(WC)↔G·C*$_{\uparrow}$(w), G·C(WC)↔G*·C$_{\downarrow}$(w), G·C(WC)↔G·C*$_{\downarrow}$(w) and G·C(WC)↔G*·C$_{\uparrow}$(w) wobbling tautomerisations producing mutagenic tautomers of the G and C bases, and also of the G·P(WC)↔G·P*$_{\uparrow}$(w) and G·P(WC)↔G·P*$_{\downarrow}$(w) tautomerisation pathways (P – cytosine analogue). Presented data are obtained at the MP2/aug-cc-pVDZ level of theory for the geometry optimized at the B3LYP/6-311++G(d,p) level of theory in the continuum with ε=1/4 (see also Tables S6-S12).

Tautomerisation	ε	ΔG	ΔE	$\Delta\Delta G_{TS}$	$\Delta\Delta E_{TS}$	$\tau_{99.9\%}$	N
G·C(WC)↔G·C*$_{\uparrow}$(w)	1	13.35	14.10	30.47	30.74	3.95	$1.6\cdot10^{-10}$
	4	12.08	12.55	31.35	31.73	148.66	$1.4\cdot10^{-9}$
G·C(WC)↔G*·C$_{\downarrow}$(w)	1	15.10	16.49	31.08	31.53	0.58	$1.3\cdot10^{-11}$
	4	14.14	15.40	32.50	32.40	32.45	$5.3\cdot10^{-10}$
G·C(WC)↔G·C*$_{\downarrow}$(w)	1	15.08	15.96	30.88	31.41	0.42	$8.8\cdot10^{-12}$
	4	14.11	15.48	24.31	24.59	$3.35\cdot10^{-5}$	$4.5\cdot10^{-11}$
G·C(WC)↔G*·C$_{\uparrow}$(w)	1	14.85	17.22	24.87	25.64	$2.50\cdot10^{-5}$	$8.4\cdot10^{-12}$
	4	12.64	14.59	19.30	19.06	$8.45\cdot10^{-8}$	$4.2\cdot10^{-11}$
G·P(WC)↔G·P*$_{\uparrow}$(w)	1	4.85	5.97	21.77	22.08	2.80	$2.8\cdot10^{-4}$
	4	3.20	4.01	21.99	22.30	65.37	$4.5\cdot10^{-3}$
G·P(WC)↔G·P*$_{\downarrow}$(w)	1	8.08	9.42	31.44	31.83	$1.48\cdot10^{5}$	$1.2\cdot10^{-6}$
	4	5.28	6.89	24.43	24.67	123.32	$1.4\cdot10^{-4}$

For the detailed designation see Table 1. See also summary Table S7 for the Gibbs and electronic energies of the mispairs and TSs relatively the global minimum – the G·C(WC) DNA base pair.

Graphical Abstract

New structural hypostasis of the A·T and G·C Watson-Crick DNA base pairs caused by their mutagenic tautomerisation in a wobble manner: A QM/QTAIM prediction

Ol'ha O. Brovarets'[a,b] & Dmytro M. Hovorun[a,b,✉]

[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. E-mail: dhovorun@imbg.org.ua

Herein, we firstly present biologically important property discovered for the classical A·T(WC) and G·C(WC) Watson-Crick (WC) DNA base pairs to pass non-dissociatively into three A*·T$_{\uparrow}$(w), A·T*$_{O2\uparrow}$(w) and A·T*$_{\downarrow}$(w) or four G·C*$_{\uparrow}$(w), G*·C$_{\downarrow}$(w), G·C*$_{\downarrow}$(w) and G*·C$_{\uparrow}$(w) different wobble (w) mismatches containing mutagenic tautomers of the bases (denoted by asterisks) and *vice versa*.