PCCP

Accepted Manuscript

This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the **Information for Authors**.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Tautomeric transition between wobble A·C DNA base mispair and Watson-Crick-like A·C* mismatch: miscrostructural mechanism and biological significance

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

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
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. Here, we use MP2/DFT quantum-chemical methods combined with Quantum Theory of Atoms in Molecules to study the tautomeric transition between wobble A·C(w) mismatch and Watson-Crick-like A·C*(WC) base mispair, proceeding non-dissociatively *via* the sequential proton transfer between the bases through the planar, highly stable and zwitterionic $TS^{A+·C-}_{A·C(w)\leftrightarrow A·C*(WC)}$ transition state joined by the participation of the $(A)N6^+H\cdots N4^-(C)$, $(A)N1^+H\cdots N4^-(C)$ and $(A)C2^+H\cdots N3^-(C)$ H-bonds. Notably, the A·C(w)↔A·C*(WC) tautomerisation reaction is accompanied by the 10 unique patterns of the specific intermolecular interactions that consistently replace each other. Our data suggest that biologically significant A·C(w)→A·C*(WC) tautomerisation is kinetically controlled pathway for the formation of the enzymatically competent Watson-Crick-like A·C*(WC) DNA base mispair in the essentially hydrophobic recognition pocket of the high-fidelity DNA-polymerase responsible for the occurrence of spontaneous point AC/CA incorporation errors during DNA biosynthesis.

Keywords: DNA biosynthesis · Spontaneous point mutation · AC/CA incorporation error · DPT · Mutagenic tautomer · B3LYP and MP2 · QTAIM

Introduction.

The ability of the DNA bases to tautomerise *via* the double proton transfer (DPT) is vitally important for the origin of the spontaneous point mutations. For a long time irregular A·C*(WC) and A*·C(WC) DNA base mispairs (mutagenic tautomers of the bases are marked with an asterisk [1-3]), which geometrical sizes are close to the dimensions of the classical A·T(WC) and G·C(WC) Watson-Crick (WC) base pairs of nucleobases [4,5], have attracted careful attention of the researchers, who develop the theory of spontaneous point mutagenesis [6-13]. Particular interest to these mispairs has been initiated by the classical Watson and Crick's work [6].

It is nowadays established that A·C*(WC) and A*·C(WC) DNA mismatches, each of which is stabilized by three intermolecular H-bonds [12-15], mutually transform into one another *via* the asynchronous concerted DPT along two neighboring antiparallel H-bonds [13]. At the same time, the A·C*(WC) : A*·C(WC) tautomeric equilibrium consists 98.98 % : 1.02 % under normal conditions [13]. Less stable A*·C(WC) base mispair is short-lived structure ($\tau = 5.8 \cdot 10^{-10}$ s) [13]. It is believed that both of these mispairs can play a pivotal role in the mutagenesis, in particular in the formation of the spontaneous transitions – AC/CA replication errors [12,13]. At the same time, it is assigned to the short-lived A*·C(WC) mispair the role of the supplier of the A·C*(WC) base mispair [13], which lifetime is considerably greater than the time spent by the DNA-polymerase machinery for the forced dissociation of the pairs of the nucleotide bases into the monomers [12]. In these mispairs the base in the mutagenic tautomeric form belongs to the template strand.

It was shown by X-ray diffraction analysis that the mismatched A·C DNA base pair acquires enzymatically competent Watson-Crick-like geometry ((A)N6···N4(C)=3.1 Å) in the recognition pocket of the DNA-polymerase β in its closed conformation [17]. By comparing these experimental data [17] with theoretical results [13], it could be assumed that enzymatically competent conformation of the aforementioned pair corresponds to the A·C*(WC) base pair.

At the same time, the physico-chemical mechanism of the A·C*(WC) base mispair formation in the recognition pocket of the high-fidelity replicative DNA-polymerase

causing transformation into its closed conformation and eventually - the mechanism of the spontaneous point AC/CA incorporation errors during DNA biosynthesis - remain unclarified. Researchers are inclined to think that mutagenic tautomerisation of the nucleobase of the incoming nucleotide is implemented by the endogenous water [17,18] (critical analysis of these approaches is described in the papers [4,5,12,13]).

In this work, we have demonstrated for the first time that isolated A·C*(WC) base mispair with Watson-Crick-like geometry [8,9,13] is in the tautomeric equilibrium with wobble (w) A·C(w) DNA mismatch [14,20]. It is proved that this tautomeric reaction proceeds *via* the transition state like an electroneutral and highly stable $A^+ \cdot C^-$ ion pair (signs "+" and "-"denote the protonation and deprotonation of the bases, respectively) - $TS^{A^+ \cdot C^-}_{A \cdot C(w) \leftrightarrow A \cdot C^*(WC)}$, stabilized by the participation of three intermolecular $(A)N6^+H \cdots N4^-(C)$, $(A)N1^+H \cdots N4^-(C)$ and $(A)C2^+H \cdots N3^-(C)$ H-bonds. It was established that the A·C(w)↔A·C*(WC) tautomerization process, or, figuratively speeking, dynamic polymorphism, which is accompanied by a significant change in the geometry of the mispair, occurs by the non-dissociative mechanism without the disruption of the mismatch. The A·C DNA base mismatch that tautomerizes is connected by 10 unique patterns of the specific intermolecular interactions that successively change each other along the A·C(w)↔A·C*(WC) tautomerisation pathway. We consider these results as principally important for the understanding of the elementary, kinetically controlled mechanism of the origin of the spontaneous point AC/CA incorporation errors during DNA biosynthesis within the framework of the classical Watson-Crick tautomeric hypothesis [6].

Computational Methods.

All calculations have been performed using Gaussian'09 package [21]. Geometries and harmonic vibrational frequencies of the A·C(w) and A·C*(WC) base mispairs and transition state $(TS^{A+·C-}_{A·C(w)\leftrightarrow A·C*(WC)})$ of their tautomerisation *via* the DPT were obtained using B3LYP DFT [22,23] functional combined with Pople's 6-311++G(d,p) basis set, that have been previously used for similar systems and verified to give accurate geometrical structures, normal mode frequencies, barrier heights and characteristics of intra- or intermolecular H-bonds [24-27]. A scaling factor that is equal to 0.9668 has been used in the present work to correct harmonic frequencies of all studied base pairs [28-30]. We have confirmed the minima and TS, located by means of Synchronous Transit-guided Quasi-Newton method [31], on the potential energy landscape by the absence or presence, respectively, of the imaginary frequency in the vibrational spectra of the pair.

In order to consider electronic correlation effects as accurately as possible, we followed geometry optimizations with single point energy calculations using MP2 functional [32] and a wide variety of basis sets, in particular, Pople's basis sets of valence triple-ζ quality [33,34], as well as Dunning's cc-type basis sets [35], augmented with polarization and/or diffuse functions: 6-311++G(2df,pd), 6-311++G(3df,2pd), cc-pVTZ and cc-pVQZ.

Reaction pathway was established by following intrinsic reaction coordinate (IRC) in the forward and reverse directions from $TS^{A+·C-}_{A·C(w)\leftrightarrow A·C*(WC)}$ using Hessian-based predictor-corrector integration algorithm [36] with tight convergence criteria. These calculations eventually ensure that the proper reaction pathway, connecting the expected reactants and products on each side of the $TS^{A+·C-}_{A·C(w)\leftrightarrow A·C*(WC)}$, has been found. We have investigated the evolution of the energetic, geometric, polar and electron-topological characteristics of the H-bonds and base pairs along the pathway of the A·C(w)$\leftrightarrow$A·C*(WC) tautomerisation reaction establishing them at the each point of the IRC [37-39].

Electronic interaction energies E_{int} have been calculated at the MP2/6-311++G(2df,pd) level of theory as the difference between the total energy of the base

mispair and the energies of the isolated monomers. Gibbs free energy of interaction has been obtained using similar equation. In each case the interaction energy was corrected for the basis set superposition error (BSSE) [40,41] through the counterpoise procedure [42,43].

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

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

where E_{el} – electronic energy, while E_{corr} – thermal correction. We applied the standard TS theory [44] to estimate the activation barriers of the tautomerisation reaction.

The time $\tau_{99.9\%}$ necessary to reach 99.9% of the equilibrium concentration of the A·C(w) and A·C*(WC) base mispairs in the system of reversible first-order forward (k_f) and reverse (k_r) reactions can be estimated by formula [44]:

$$\tau_{99.9\%} = \frac{ln10^3}{k_f + k_r}. \tag{2}$$

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

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

we applied standard TS theory [44], in which quantum tunneling effect are accounted by Wigner's tunneling correction [45], that has been successfully used for the DPT reactions [4,5,46-49]:

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

where k_B – Boltzmann's constant, h – Planck's constant, $\Delta\Delta G_{f,r}$ – Gibbs free energy of activation for the tautomerisation reaction in the forward (f) and reverse (r) directions, ν_i – magnitude of the imaginary frequency associated with the vibrational mode at the $TS^{A+·C-}_{A·C(w)\leftrightarrow A·C*(WC)}$.

Bader's quantum theory of Atoms in Molecules was applied to analyse the electron density distribution [50]. The topology of the electron density was analysed using program package AIMAll [51] with all default options. The presence of a bond critical point (BCP), namely the so-called (3,-1) BCP, and a bond path between hydrogen donor and acceptor, as well as the positive value of the Laplacian at this BCP ($\Delta\rho > 0$), were

considered as criteria for the H-bond formation [14,15]. Wave functions were obtained at the level of theory used for geometry optimisation.

The energies of the weak CH···O/N H-bonds [14,15] were calculated by the empirical Espinosa-Molins-Lecomte (EML) formula [52,53] based on the electron density distribution at the (3,-1) BCPs of the H-bonds:

$E_{CH\cdots O/N}=0.5\cdot V(r),$ (5)

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

The energies of the AH···B conventional H-bonds were evaluated by the empirical Iogansen's formula [54]:

$$E_{AH\cdots B} = 0.33\cdot\sqrt{\Delta v - 40},$$ (6)

where Δv – magnitude of the redshift (relative to the free molecule) of the stretching mode of the AH H-bonded group involved in the AH···B H-bond. The partial deuteration was applied to minimize the effect of vibrational resonances [13,14].

The atomic numbering scheme for the A and C bases is conventional [55].

Results and Their Discussion.

We have shown for the first time that isolated biologically important A·C*(WC) and A·C(w) base mispairs stay in the tautomeric equilibrium A·C*(WC):A·C(w)=99.97 % : 0.03 % under normal conditions. The $TS^{A+\cdot C-}_{A\cdot C(w)\leftrightarrow A\cdot C*(WC)}$ transition state of the A·C(w)↔A·C*(WC) tautomeric reaction represents itself electroneutral, highly-polar $A^+\cdot C^-$ ion pair, stabilized by the participation of the $(A)N6^+H\cdots N4^-(C)$ (2.27), $(A)N1^+H\cdots N4^-(C)$ (10.35) and $(A)C2^+H\cdots N3^-(C)$ (2.62 kcal·mol^{-1}) H-bonds (Scheme 1, Figs. 1-3, Tables 1-3). Geometrical architecture of the $TS^{A+\cdot C-}_{A\cdot C(w)\leftrightarrow A\cdot C*(WC)}$ transition state is closer to the Watson-Crick geometry (Fig. 2): its glycosidic sizes are R(H$_1$-H$_9$)=9.822 Å, α_1(N1(C)H$_1$H$_9$)=51.8°, α_2(N9(A)H$_9$H$_1$)=66.4° and similarly to the canonical A·T and G·C Watson-Crick DNA base pairs it contains three intermolecular H-bonds [4,5]. This fact is completely consistent with Hammond's postulate [56], since the A·C(w)→A·C*(WC) tautomerisation is an exothermic process with ΔG=-4.87 kcal·mol^{-1} under normal conditions (Table 2). The A·C(w)↔A·C*(WC) tautomerisation reaction is accompanied by the substantial rearrangement of the base

mispair changing from wobble to Watson-Crick geometry and *vice versa*. This process occurs *via* the non-dissociative mechanism without destruction of the pair that tautomerises (Fig. 2). At this, 10 unique patterns of the intermolecular interactions including H-bonds that consistently switch to each other during the tautomerisation process have been registered by us. At the same time, base pair that tautomerises remains in the $A^+ \cdot C^-$ zwitterionic state [57], stabilized by rather strong electrostatic interactions, in the wide range of the IRC (-2.01÷9.01 Bohr) (Fig. S3 in ESI). In particular, the stabilization energy of the $TS^{A^+ \cdot C^-}_{A \cdot C(w) \leftrightarrow A \cdot C^*(WC)}$ transition state of the $A \cdot C(w) \leftrightarrow A \cdot C^*(WC)$ tautomerisation makes at this 122.5 kcal·mol^{-1}, while the contribution of the H-bonds constitutes 12.4 % (Table 3).

The $A \cdot C(w) \rightarrow A \cdot C^*(WC)$ tautomerization process *via* the sequential DPT occurs at the participation of the A base as the molecule-mediator at the interaction. Initially, the A base detaches the imino proton localized at the N4 nitrogen atom of the C base, adding it to its N1 nitrogen atom (at this proton migrates along the intermolecular (C)N4H···N1(A) H-bond, forming highly stable $A^+ \cdot C^-$ ion pair). After this, ionized A^+ and C^- bases, which constitute the $A^+ \cdot C^-$ ion pair, shift relatively each other and further proton, mentioned above, goes back to another atom of the C base, namely – N3, moving along the $(A)N1^+H···N3^-(C)$ H-bond within the $A^+ \cdot C^-$ ion pair. Finally, after this the $A \cdot C^*(WC)$ pair with Watson-Crick geometry is formed (Scheme 1). Interestingly, in this case DPT takes place, which is a concerted two-phase process for the one and the same proton (Figs. S3a, S3b in ESI).

Talking in other words, A base catalyzes, accelerates the mutagenic tautomerisation of the C base, transforming the wobble $A \cdot C(w)$ base mispair into the $A \cdot C^*(WC)$ base mispair with Watson-Crick geometry (Scheme 1, Figs. 1-3, Tables 1-3). Thus, the Gibbs free energy barrier of the mutagenic tautomerisation of the C base within the $A \cdot C$ pair (20.12 kcal·mol^{-1} (Table 2)) is significantly reduced comparably with the same barrier of the intramolecular mutagenic tautomerisation of the isolated C base (38.59 kcal·mol^{-1} under normal conditions [58]).

Detailed microstructural information about the course of the A·C(w)↔A·C*(WC) tautomerisation *via* the sequential DPT is presented in Scheme1, Figs. 1-3, S1-S3 in ESI and Tables 1-3.

Based on these data we can arrive to the following important conclusions:

- both extrema of the first derivative of the electron energy with respect to the IRC – dE/dIRC (Fig. S1 in ESI), so-called reaction force [59], coincide with the key points of the A·C(w)↔A·C*(WC) tautomerisation reaction, where values of the $\Delta\rho_{(C)N4H\cdots N1(A)}$ and $\Delta\rho_{(C)N3H\cdots N1(A)}$ become equal to 0 (Fig. S3b in ESI);

- investigated tautomerisation reaction is dipole-active process (Fig. S2 in ESI) – dipole moment of the system acquires its maximum value (7.77 D) at the $TS^{A+\cdot C-}_{A\cdot C(w)\leftrightarrow A\cdot C*(WC)}$ transition state;

- in the wide range of the IRC (Figs. 3 and S3-S4 in ESI, Table 4) tautomerisation process is assisted by the (A)C2H···N3(C) and (A)C2H···O2(C) H-bonds, for which maximum values of energies reach 2.89 kcal·mol^{-1} at the IRC=-1.42 Bohr and 1.56 kcal·mol^{-1} at the IRC=10.73 Bohr, respectively.

Obtained results, regarding the mechanism of the A·C(w)→A·C*(WC) tautomeric interconversion between A·C(w) and A·C*(WC) biologically important mispairs as their intrinsic property, open up entirely new perspectives for the interpretation of the nature of the spontaneous point AC/CA incorporation errors during DNA biosynthesis at the atomic level. There are all grounds to associate with the A·C(w)→A·C*(WC) transition the mechanism of the acquisition by the wobble A·C(w) DNA base mispair of the enzymatically competent Watson-Crick conformation in the recognition pocket of the high-fidelity replicative DNA-polymerase in its closed conformation without direct involvement of water molecules as endogenous agent of tautomerisation [60].

The A·C(w)→A·C*(WC) tautomerisation is much slower process ($\tau_{99.9\%}$=4.94·10^2 s (Table 2)) than the rate of the DNA replication (DNA polymerase spends Δt_{pol}≈8.33·10^{-4} s [16] for the incorporation of one incoming nucleotide into the

DNA double helix that is synthesized). So, the P(A·C(w)→A·C*(WC)) probability of the acquisition by the A·C(w) base pair, that initially forms in the recognition pocket of the DNA-polymerase, of the enzymatically competent A·C*(WC) Watson-Crick conformation for this period of time is much less than 1. In other words, the process of the incorrect A·C*(WC) base pair incorporation into the DNA structure during its biosynthesis is under kinetic control, which means that this process is thermodynamically non-equilibrium.

The P(A·C(w)→A·C*(WC)) probability can be easily calculated by the formula of the physico-chemical kinetics given in the works [61,62]. P(A·C(w)→A·C*(WC)) has been calculated to be $1.2 \cdot 10^{-5}$ at $k_f = 1.40 \cdot 10^{-2}$ s^{-1}, $k_r = 3.76 \cdot 10^{-6}$ s^{-1} and $\Delta t_{pol} \approx 8.33 \cdot 10^{-4}$ s [16]. This valuation is most likely underestimated, because the affinity to the recognition pocket of the high-fidelity replicative DNA-polymerase is higher for the A·C*(WC) base mispair [11], than for the A·C(w) base mispair. These facts with necessity should lead to the decreasing of the value of the energy barrier of the A·C(w)→A·C*(WC) tautomerisation in the essentially hydrophobic [63,64] recognition pocket in comparison with the isolated state. Nevertheless, despite the simplicity of our model representations, these data satisfactorily agree with the experimental data on the frequency of the spontaneous transitions, committed by the replicative DNA-polymerase with the inactivated function of proof-reading [65-68]. Obtained data represent convincing evidence of the adequacy of the proposed by us microstructural model of the initiation of the spontaneous point mutations – AC/CA incorporation errors – during DNA biosynthesis.

Concluding Remarks and Perspective.

In this work for the first time using quantum-chemical methods we have reported the tautomeric transition between biologically important A·C(w) and A·C*(WC) base mispairs with wobble and Watson-Crick geometries, respectively, that is controlled by the planar, highly polar and zwitterionic [57] $TS^{A+\cdot C-}_{A \cdot C(w) \leftrightarrow A \cdot C*(WC)}$ transition state stabilized by the participation of the (A)N6$^+$H···N4$^-$(C) (2.27), (A)N1$^+$H···N4$^-$(C) (10.35) and (A)C2$^+$H···N3$^-$(C) (2.62 kcal·mol^{-1}) H-bonds in addition to the strong

electrostatic interactions, mainly contributing toward the H-bonding interactions in this ion pair. Authors associate the biological significance of the A·C(w)→A·C*(WC) DPT tautomerisation, as intrinsic property of these DNA base mispairs, with the mechanism of the formation of the enzymatically competent Watson-Crick-like A·C*(WC) base mispair in the essentially hydrophobic recognition pocket of the high-fidelity replicative DNA-polymerase, responsible for the occurrence of the spontaneous point AC/CA incorporation errors during DNA biosynthesis. It was found out that kinetically controlled A·C(w)→A·C*(WC) tautomerization pathway proceeds *via* the non-dissociative mechanism without the disruption of the pair into the monomers and is accompanied by 10 unique patterns of the intermolecular interactions, including simultaneously maximum 4 H-bonds, that successively change each other.

The simplest estimation of the probability of this process ($1.2 \cdot 10^{-5}$) is satisfactorily agreed with experimental data, indicating the adequacy of the model concepts presented in this work.

In the perspective, it seems to be very interesting to investigate the influence of the microenvironment of the recognition pocket of the high-fidelity replicative DNA-polymerase on the considered process of the tautomerisation. Taking into account high polarity of the transition state and its Watson-Crick-like structure, it is conceivable that the latter would be stabilized, accelerating the process of the A·C(w)→A·C*(WC) mutagenic tautomerization in comparison with the free state.

References.

1. M. K. Shukla and J. Leszczynski (2013). Tautomerism in nucleic acid bases and base pairs: a brief overview. *Wiley Interdisciplinary Reviews: Computational Molecular Science, 3,* 637-649.

2. D. Kosenkov, Y. Kholod, L. Gorb, O. Shishkin, D. M. Hovorun, M. Mons and J. Leszczynski (2009). *Ab initio* kinetic simulation of gas-phase experiments: tautomerization of cytosine and guanine. *Journal of Physical Chemistry B*, *113*, 6140-6150.

3. O. O. Brovarets' and D. M. Hovorun (2015). 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. *Journal of Biomolecular Structure & Dynamics, 33,* 925-945.

4. O. O. Brovarets' and D. M. Hovorun (2014). 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. *Journal of Biomolecular Structure and Dynamics, 32*, 127-154.

5. O. O. Brovarets' and D. M. Hovorun (2014). Why the tautomerization of the G·C Watson–Crick base pair *via* the DPT does not cause point mutations during DNA replication? QM and QTAIM comprehensive analysis. *Journal of Biomolecular Structure & Dynamics, 32*, 1474-1499.

6. J. D. Watson and F. H. C. Crick (1953). The Structure of DNA. *Cold Spring Harbor Symposia on Quantitative Biology, 18,* 123–31.

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

8. V. I. Danilov, V. M. Anisimov, N. Kurita and D. Hovorun (2005). MP2 and DFT studies of the DNA rare base pairs: the molecular mechanism of the spontaneous substitution mutations conditioned by tautomerism of bases. *Chemical Physics Letters, 412*, 285-293.

9. C. Fonseca Guerra, F. M. Bickelhaupt, S. Saha and F. Wang, (2006). Adenine tautomers: relative stabilities, ionization energies, and mismatch with cytosine. *Journal of Physical Chemistry A*, *110*, 4012-4020.

10. O. O. Brovarets' and D. M. Hovorun (2009). 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, *Ukrainica Bioorganica Acta*, *8*, 12-18.

11. W. Wang, H. W. Hellinga and L. S. Beese (2011). Structural evidence for the rare tautomer hypothesis of spontaneous mutagenesis. *Proceedings of the National Academy of Sciences of the United States of America, 124*, 17644–17648.

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

13. O. O. Brovarets' and D. M. Hovorun (2015). The physicochemical essence of the purine·pyrimidine transition mismatches with Watson-Crick geometry in DNA: A·C* *versa* A*·C. A QM and QTAIM atomistic understanding. *Journal of Biomolecular Structure & Dynamics, 33,* 28-55.

14. O. O. Brovarets', Yurenko Y. P. and D. M. Hovorun (2014). Intermolecular CH   O/N H-bonds in the biologically important pairs of natural nucleobases: A thorough quantum-chemical study. *Journal of Biomolecular Structure & Dynamics, 32,* 993-1022.

15. O. O. Brovarets', Yurenko Y.P. and D. M. Hovorun (2014). 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. *Journal of Biomolecular Structure & Dynamics,* DOI: 10.1080/07391102.2014.968623.

16. S. Kirmizialtin, V. Nguyen, K. A. Johnson and R. Elber (2012). How conformational dynamics of DNA polymerase select correct substrates: experiments and simulations. *Structure, 20*, 618–627.

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

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

19. M.-Ch. Koag, K. Nam and S. Lee (2014). The spontaneous replication error and the mismatch discrimination mechanisms of human DNA polymerase β. *Nucleic Acids Research, 42*, 11233-11245.

20. G. Rossetti, P. D. Dans, I. Gomez-Pinto, I. Ivani, G. Gonzalez and M. Orozco (2015). The structural impact of DNA mismatches. *Nucleic Acids Research*, DOI:10.1093/nar/gkv254.

21. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. Montgomery, J. A. T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A.Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson,W. Chen, M. W. Wong, C.Gonzalez and J. A. Pople, *Gaussian 09, Revision B.01 ed*, Gaussian Inc., Wallingford, CT, 2010.

22. R. G. Parr and W. Yang (1989). *Density-functional theory of atoms and molecules*. Oxford: Oxford University Press.

23. C. Lee, W. Yang and R. G. Parr (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B: Condensed Matter and Materials Physics*, *37*, 785-789.

24. M. Lozynski, D. Rusinska-Roszak and H.-G. Mack (1998). Hydrogen bonding and density functional calculations: The B3LYP approach as the shortest way to MP2 results. *Journal of Physical Chemistry A*, *102*, 2899–2903.

25. C. F. Matta (2010). 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. *Journal of Computational Chemistry, 31*, 1297–1311.

26. O. O. Brovarets' and D. M. Hovorun (2014). DPT tautomerisation of the G·A$_{syn}$ and A*·G*$_{syn}$ DNA mismatches: A QM/QTAIM combined atomistic investigation. *Physical Chemistry Chemical Physics*, *16*, 9074-9085.

27. O. O. Brovarets' and D. M. Hovorun (2010). Stability of mutagenic tautomers of uracil and its halogen derivatives: The results of quantum-mechanical investigation. *Biopolymers and Cell, 26*, 295–298.

28. S. P. Samijlenko, Y. P. Yurenko, A. V. Stepanyugin and D. M. Hovorun (2012). Tautomeric equilibrium of uracil and thymine in model protein−nucleic acid contacts. Spectroscopic and quantum chemical approach. *Journal of Physical Chemistry B, 114*, 1454-1461.

29. O. O. Brovarets', R. O. Zhurakivsky and D. M. Hovorun (2013). 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. *Chemical Physics Letters*, *578*, 126-132.

30. O. O. Brovarets', R. O. Zhurakivsky and D. M. Hovorun (2014). A QM/QTAIM microstructural analysis of the tautomerisation *via* the DPT of the hypoxanthine·adenine nucleobase pair. *Molecular Physics, 112*, 2005-2016.

31. C. Peng, P. Y. Ayala, H. B. Schlegel and M. J. Frisch (1996). Using redundant internal coordinates to optimize equilibrium geometries and transition states. *Journal of Computational Chemistry, 17*, 49–56.

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

33. P. C. Hariharan and J. A. Pople (1973). The influence of polarization functions on molecular orbital hydrogenation energies. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta), 28*, 213−222.

34. R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople (1980). Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *Journal of Chemical Physics, 72*, 650−654.

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

36. H. P. Hratchian and H. B. Schlegel (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.

37. O. O. Brovarets', R. O. Zhurakivsky and D. M. Hovorun (2013). The physico-chemical "anatomy" of the tautomerisation through the DPT of the biologically important pairs of hypoxanthine with DNA bases: QM and QTAIM perspectives. *Journal of Molecular Modeling*, *19*, 4119-4137.

38. O. O. Brovarets' and D. M. Hovorun (2014). Does the G•G*$_{syn}$ DNA mismatch containing canonical and rare tautomers of the guanine tautomerise through the DPT? A QM/QTAIM microstructural study. *Molecular Physics, 112*, 3033-3046.

39. O. O. Brovarets' and D. M. Hovorun (2014). 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. *Physical Chemistry Chemical Physics, 6*, 15886-15899.

40. S. F. Boys and F. Bernardi (1970). The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Molecular Physics*, *19,* 553–566.

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

42. J. A. Sordo, S. Chin and T. L. Sordo (1988). On the counterpoise correction for the basis set superposition error in large systems. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta), 74*, 101–110.

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

44. P. W. Atkins (1998). *Physical chemistry*. Oxford: Oxford University Press.

45. E. Wigner (1932). Über das Überschreiten von Potentialschwellen bei chemischen Reaktionen [Crossing of potential thresholds in chemical reactions]. *Zeitschrift für Physikalische Chemie, B19*, 203−216.

46. O. O. Brovarets' and D. M. Hovorun (2013). Atomistic nature of the DPT tautomerisation of the biologically important C•C* DNA base mispair containing amino and imino tautomers of the cytosine: A QM and QTAIM approach. *Physical Chemistry Chemical Physics, 15*, 20091-20104.

47. O. O. Brovarets' and D. M. Hovorun (2013). Atomistic understanding of the C·T mismatched DNA base pair tautomerization via the DPT: QM and QTAIM computational approaches. *Journal of Computational Chemistry, 34*, 2577-2590.

48. O. O. Brovarets', R. O. Zhurakivsky and D. M. Hovorun (2013). DPT tautomerization of the long A·A* Watson-Crick base pair formed by the amino and imino tautomers of adenine: combined QM and QTAIM investigation. *Journal of Molecular Modeling, 19*, 4223-4237.

49. O. O. Brovarets', R. O. Zhurakivsky and D. M. Hovorun (2014). 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. *Physical Chemistry Chemical Physics, 16*, 3715-3725.

50. R. F. W. Bader (1990). *Atoms in molecules: A quantum theory.* Oxford: Oxford University Press.

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

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

53. I. Mata, I. Alkorta, E. Espinosa and E. Molins (2011). Relationships between interaction energy, intermolecular distance and electron density properties in hydrogen bonded complexes under external electric fields. *Chemical Physics Letters, 507*, 185–189.

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

55. W. Saenger (1984). *Principles of nucleic acid structure.* New York: Springer.

56. G. S. Hammond (1955). A correlation of reaction rates. *Journal of the American Chemical Society, 77*, 334–338.

57. J. P. Richard and T. L. Amyes (2004). On the importance of being zwitterionic: enzymatic catalysis of decarboxylation and deprotonation of cationic carbon. *Bioorganic Chemistry, 32,* 354-366.

58. O. O. Brovarets' and D. M. Hovorun (2010). How stable are the mutagenic tautomers of DNA bases? *Biopolymers and Cell, 26,* 72–76.

59. P. Politzer, J. S. Murray and P. Jaque (2013). Perspectives on the reaction force constant. *Journal of Molecular Modeling, 19*, 4111-4118.

60. O. O. Brovarets' and D. M. Hovorun (2015). How many tautomerisation pathways connect Watson-Crick-like G*•T DNA base mispair and wobble mismatches? *Journal of Biomolecular Structure & Dynamics,* in press.

61. W. J. Moore (1983). *Basic Physical Chemistry*. Prentice-Hall: Englewood Cliffs, NJ.

62. Y. Podolyan, L. Gorb and J. Leszczynski (2003). *Ab initio* study of the prototropic tautomerism of cytosine and guanine and their contribution to spontaneous point mutations. *International Journal of Molecular Sciences, 4,* 410-421.

63. Y. Yang, Y. Qin, Q. Ding, M. Bakhtina, L. Wang, M.-D. Tsai and D. Zhong (2014). Ultrafast water dynamics at the interface of the polymerase–DNA binding complex. *Biochemistry, 53*, 5405-5413.

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

65. M.-M. Huang, N. Arnheim and M. Goodman (1992). Extension of base mispairs by Taq DNA polymerase: implications for single nucleotide discrimination in PCR. *Nucleic Acids Research*, *20*, 4567-4573.

66. I. J. Fijalkowska, R. M. Schaaper and P. Jonczyk (2012). DNA replication fidelity in *Escherichia coli*: a multi-DNA polymerase affair. *FEMS Microbiology Reviews, 36*, 1105–1121.

67. T.A. Kunkel and K. Bebenek (2000). DNA replication fidelity. *Annual Review of Biochemistry*, *69,* 497–529.

68. T. A. Kunkel and P. S. Alexander (1986). The base substitution fidelity of eucaryotic DNA polymerases. Mispairing frequencies, site preferences, insertion preferences, and base substitution by dislocation. *Journal of Bioogical Chemistry, 261*, 160-166.

Table 1. Electron-topological, structural, vibrational and energetic characteristics of the intermolecular H-bonds revealed in the base mispairs and TSs of their mutual transformation *via* the sequential DPT and polar characteristics of the latters obtained at the B3LYP/6-311++G(d,p) level of theory

Complex	AH···B H-bond	ρ[a]	$\Delta\rho$[b]	$100·\varepsilon$[c]	$d_{A···B}$[d]	$d_{H···B}$[e]	Δd_{AH}[f]	$\angle$AH···B[g]	$\Delta\nu$[h]	E_{HB}[i]	μ[j]
A·C(w)	(C)N4H···N1(A)	0.025	0.074	7.72	3.052	2.030	0.016	175.6	285.2	5.17	5.20
	(A)C2H···N3(C)	0.010	0.030	5.53	3.452	2.513	0.0005	144.2	−5.8	1.60*	
A·C*(WC)	(A)N6H···N4(C)	0.029	0.082	7.63	2.983	1.959	0.021	173.8	371.7	6.01	3.10
	(C)N3H···N1(A)	0.040	0.093	6.59	2.895	1.852	0.031	178.9	551.6	7.46	
	(A)C2H···O2(C)	0.005	0.017	1.48	3.628	2.798	0.0001	133.1	−5.9	0.96*	
TS$^{A+·C-}_{A·C(w)\leftrightarrow A·C*(WC)}$	(A)N6$^+$H···N4$^-$(C)	0.014	0.045	6.68	3.151	2.317	0.006	138.8	87.4	2.27	7.77
	(A)N1$^+$H···N4$^-$(C)	0.055	0.091	4.92	2.749	1.688	0.064	166.8	1023.3	10.35	
	(A)C2$^+$H···N3$^-$(C)	0.015	0.048	8.72	3.068	2.375	0.003	120.1	24.0	2.62*	

[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 or EML (marked with an asterisk) formulas, kcal·mol^{-1}
[j]The dipole moment of the complex, D

Table 2. Energetic and kinetic characteristics of the A·C(w)↔A·C*(WC) tautomerisation *via* the sequential DPT 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]
MP2/6-311++G(2df,pd)	4.66	6.56	20.12	18.98	24.77	25.55	$6.24·10^2$
MP2/6-311++G(3df,2pd)	4.82	6.72	19.70	18.57	24.52	25.29	$3.09·10^2$
MP2/cc-pVTZ	5.28	7.18	20.45	19.32	25.72	26.50	$1.09·10^3$
MP2/cc-pVQZ	4.87	6.77	19.98	18.85	24.84	25.62	$4.94·10^2$

[a]The relative Gibbs free energy of the A·C(w) base pair ($\Delta G_{A·C*(WC)}$=0; T=298.15 K), kcal·mol^{-1}
[b]The relative electronic energy of the A·C(w) base pair ($\Delta E_{A·C*(WC)}$=0), kcal·mol^{-1}
[c]The Gibbs free energy of activation for the forward reaction of the A·C(w)→A·C*(WC) tautomerisation, kcal·mol^{-1}
[d]The activation electronic energy for the forward reaction of the A·C(w)→A·C*(WC) tautomerisation, kcal·mol^{-1}
[e]The Gibbs free energy of activation for the reverse reaction of the A·C(w)→A·C*(WC) tautomerisation, kcal·mol^{-1}
[f]The activation electronic energy for the reverse reaction of the A·C(w)→A·C*(WC) tautomerisation, kcal·mol^{-1}
[g]The time necessary to reach 99.9% of the equilibrium concentration between the reactant A·C(w) and the product A·C*(WC) of the A·C(w)↔A·C*(WC) tautomerisation reaction, s

Table 3. Interbase interaction energies (in kcal·mol^{-1}) for the investigated base mispairs and TS of their mutual transformation *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory

Complex	$-\Delta E_{int}$[a]	ΣE_{HB}[b]	$\Sigma E_{HB}/|\Delta E_{int}|$, %	$-\Delta G_{int}$[c]
A·C(w)	7.47	6.77	90.6	0.84
A·C*(WC)	15.73	14.44	91.8	2.27
TS$^{A+·C-}_{A·C(w)\leftrightarrow A·C*(WC)}$	122.54	15.24	12.4	110.00

[a]The BSSE-corrected electronic interaction energy
[b]The total energy of the intermolecular H-bonds (see Table 1)
[c]The BSSE-corrected Gibbs free energy of interaction (T=298.15 K)

Table 4. Patterns of the intermolecular interactions including AH···B H-bonds and loosened A-H-B covalent bridges that sequentially replace each other along the IRC of the A·C(w)↔A·C*(WC) tautomerisation *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory (see Figs. 3 and S3 in ESI)

Patterns	IRC range, Bohr	Intermolecular interactions, forming patterns
I	[-14.42÷-2.01)	(C)N4H···N1(A), (A)C2H···N3(C)
II	[-2.01÷-1.66)	(C)N4-H-N1(A), (A)C2H···N3(C)
III	[-1.66÷-1.54)	(A)N6H···N4(C), (C)N4-H-N1(A), (A)C2H···N3(C)
IV	[-1.54÷0.36)	(A)N6H···N4(C), (A)N1H···N4(C), (A)C2H···N3(C)
V	[0.36÷3.44)	(A)N6H···N4(C), (A)N1H···N4(C), (A)N1H···N3(C), (A)C2H···N3(C)
VI	[3.44÷3.67)	(A)N6H···N4(C), (A)N1H···N4(C), (A)N1H···N3(C)
VII	[3.67÷4.98)	(A)N6H···N4(C), (A)N1H···N4(C), (A)N1H···N3(C), (A)C2H···O2(C)
VIII	[4.98÷9.01)	(A)N6H···N4(C), (A)N1H···N3(C), (A)C2H···O2(C)
IX	[9.01÷9.60)	(A)N6H···N4(C), (A)N1-H-N3(C), (A)C2H···O2(C)
X	[9.60÷15.61]	(A)N6H···N4(C), (C)N3H···N1(A), (A)C2H···O2(C)
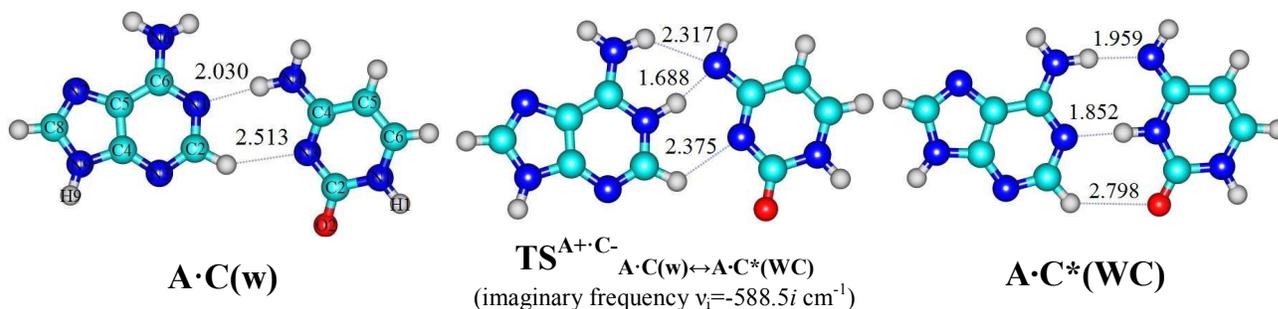
Scheme 1. Geometrical structures of the reactant (A·C(w)), transition state (TS$^{A+\cdot C-}_{A·C(w)\leftrightarrow A·C^*(WC)}$) and product (A·C*(WC)) of the A·C(w)↔A·C*(WC) tautomerisation reaction *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory. The dotted lines indicate the AH···B H-bonds, their H···B lengths are presented above them in angstroms. Carbon atoms are in light blue, nitrogen – in dark blue, hydrogen – in grey and oxygen – in red.
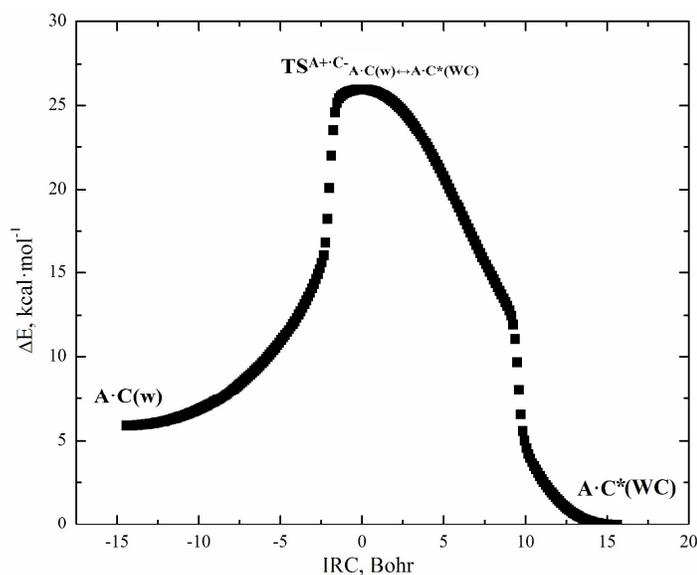

Fig. 1. Profile of the relative electronic energy ΔE of the A·C(w) base mispair along the IRC of the A·C(w)↔A·C*(WC) tautomerisation *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory.
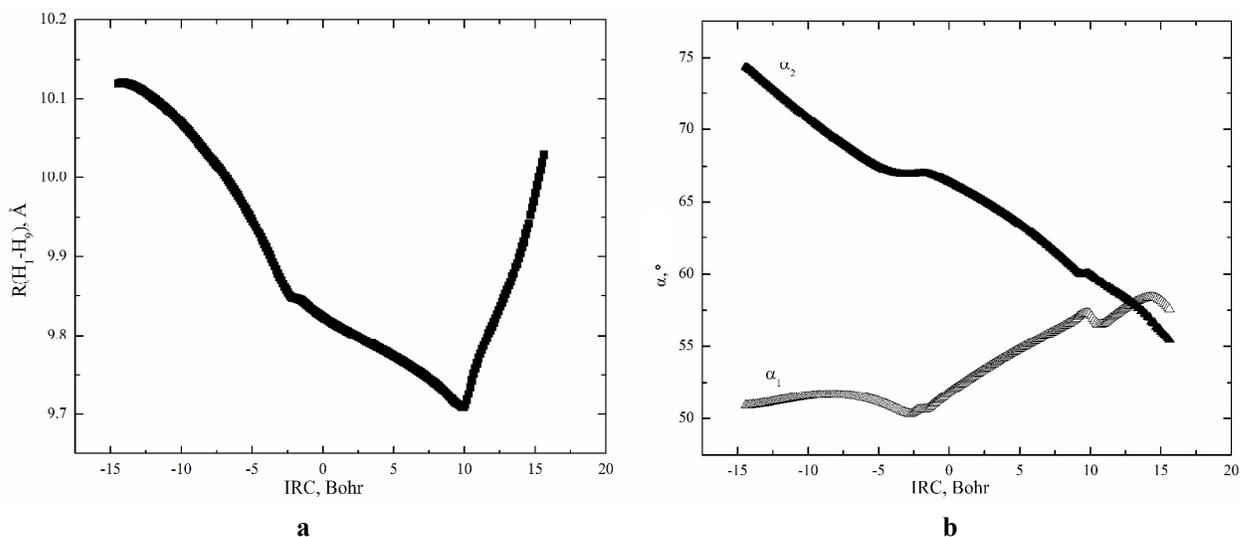
Fig. 2. Profiles of: (**a**) the distance R(H$_1$-H$_9$) between the H$_1$ and H$_9$ glycosidic hydrogens and (**b**) the α$_1$ (∠N1H$_1$(C)H$_9$(A)) and α$_2$ (∠N9H$_9$(A)H$_1$(C)) glycosidic angles along the IRC of the A·C(w)↔A·C*(WC) tautomerisation *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory (see Scheme 1).
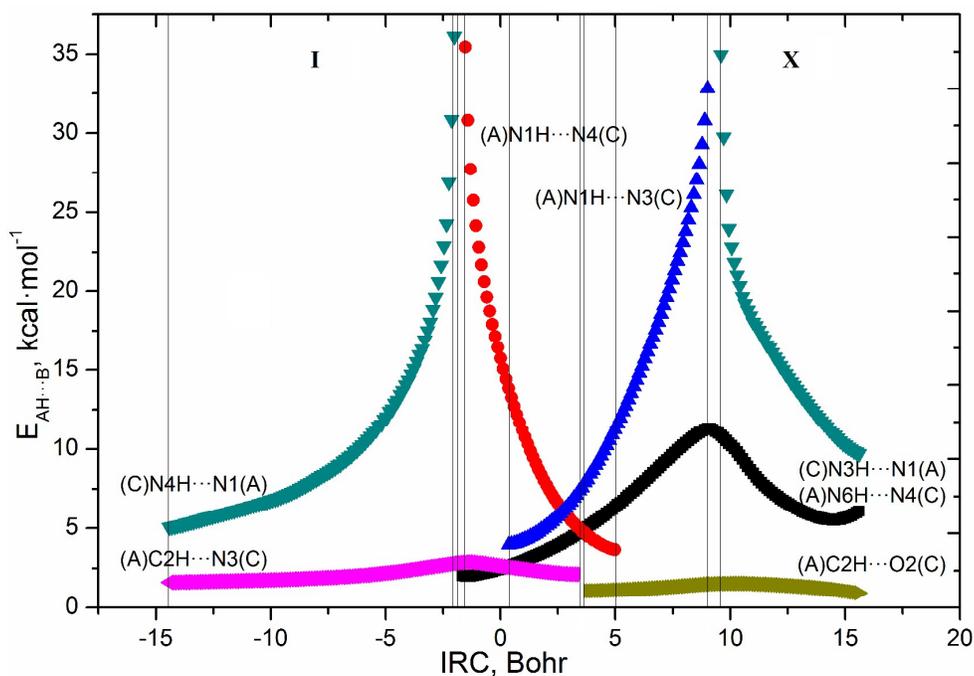

Fig. 3. Profiles of the energy of the intermolecular H-bonds E$_{AH\cdots B}$ estimated by the EML formula at the (3,-1) BCPs along the IRC of the A·C(w)↔A·C*(WC) tautomerisation *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory (see Table 4).