

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](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

A novel conception for spontaneous transversions caused by homo-pyrimidine DNA mismatches: A QM/QTAIM highlight

OI'ha O. Brovarets^{a,b} & Dmytro M. Hovorun^{a,b,✉}

^aDepartment of Molecular and Quantum Biophysics, Institute of Molecular Biology and Genetics, National Academy of Sciences of Ukraine, 150 Akademika Zabolotnoho Str., 03680 Kyiv, Ukraine

^bDepartment 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. We have firstly shown that the T·T(w) and C·C(w) DNA mismatches with wobble (w) geometry stay in the slow tautomeric equilibrium with short T·T*(WC) and C·C*(WC) Watson-Crick (WC) mispairs. These non-dissociative tautomeric rearrangements are controlled by the plane-symmetric, highly stable, highly polar and zwitterionic transition states. Obtained results allow us to understand, in what way the T·T(w) and C·C(w) mismatches acquire enzymatically competent T·T*(WC) and C·C*(WC) conformations directly in the hydrophobic recognition pocket of the high-fidelity DNA-polymerase, thereby producing thermodynamically non-equilibrium spontaneous transversions. The simplest numerical estimation of the frequency ratio of the TT to CC spontaneous transversions satisfactorily agrees with experimental data.

Keywords: DNA biosynthesis · Incorporation error · Spontaneous transversion · Wobble conformation · Watson-Crick mispair · Homo-pyrimidine mismatch · Tautomeric transition

† Electronic supplementary information (ESI) available: (i) Computational details; (ii) Physico-chemical parameters of the specific intermolecular contacts; (iii) Energetic and kinetic characteristics of the mispairs tautomerisation; (iv) Interaction energies for the investigated DNA mispairs; (v) Structures of the stationary points of the tautomeric conversions; (vi) Profiles of the glycosidic parameters of the mispairs along the IRC of the tautomerisation; (vii) Ranges of the existence of the obtained patterns of the specific intermolecular interactions along the IRC of the tautomerisation. See DOI: 10.1039/x0xx00000x.

Introduction. An intriguing theme of the origin of the spontaneous point mutations has excited researchers' mind during several decades suggesting DNA mispairs as their source [1-6]. Characterizing current state of this biologically important area of knowledge, it can be certainly stated that one of the stumbling stones in the theory of spontaneous point mutagenesis, which slowly gets on its feet, is understanding of the nature of the spontaneous transversions [7-10] caused by the homo-pyrimidine DNA base mispairs. Thorough analysis of the literature evidences that progress in this biologically important area is constrained by the lack of experimental data, especially of the X-ray crystallographic analysis of the enzymatically competent conformation of the incorrect C·C and T·T nucleobase pairs in the active center of the high-fidelity replicative DNA-polymerase at the site of incorporation [11-13], and, from the other hand, with the absence of new model conceptions and approaches.

The rare tautomer hypothesis of spontaneous point mutagenesis formulated by Watson and Crick [1] has been recently challenged by computational [14,15] and experimental [11,12,16-18] studies demonstrating that displacement of one single proton alters H-bonding scheme of the mispair, making it indistinguishable from the canonical Watson-Crick base pairs, that give impetus to further studies. It is logical to assume within the framework of this hypothesis that exceptionally the C·C*(WC) [19-21] and T·T*(WC) [20-22] base mispairs possessing Watson-Crick-like (WC) geometry and containing rare tautomers [23,24] of the C and T DNA bases (here and below they are marked with an asterisk) can be enzymatically incorporated into the structure of the DNA double helix and so are implicated in the origin of spontaneous transversions. Albeit wobble (w) [25-27] C·C(w) and T·T(w) base mispairs cause mild structural and dynamic distortions in the composition of the DNA double helix [28], we have suggested that the C·C*(WC) [19] and T·T*(WC) [20] mismatches with *cis*-oriented glycosidic bonds [20,21] are their enzymatically competent configurations causing spontaneous transversions.

In this study we have examined in details all possible tautomerisation ways connecting the homo-pyrimidine T·T(w) and C·C(w) mispairs into the C·C*(WC) and T·T*(WC) DNA mismatches and occurring without breakage of the mutual *cis*-orientation of their glycosidic bonds. Formed C·C*(WC) and T·T*(WC) mismatches are able to accommodate into the hydrophobic recognition pocket of the high-fidelity replicative DNA-polymerase and thus to be responsible for the origin of transversion mutations.

All geometric, energetic and vibrational calculations of the considered base mispairs and transition states (TSs) of their conversion have been performed by Gaussian'09 package [29] using B3LYP [30] and MP2 [31] levels of theory combined with a wide variety of basis sets. Bader's quantum theory of Atoms in Molecules was applied to analyse the electron density distribution [32-34]. IRC calculations have been performed to further confirm the proposed mechanisms of conversion [35] and to obtain profiles of the energetic and geometric characteristics of the base pairs and H-bonds in them along the reaction pathway [36-38] (for more details see ESI†).

Thus, we have shown for the first time that plane-symmetric (C_s) T·T(w) mismatch, that in fact is observed in the DNA double helix under physiological conditions [25-27], stays in slow tautomeric equilibrium with two T·T*(WC) [20] and T*_{O2}·T(WC) mispairs representing itself propeller-like and structurally non-rigid complexes [39,40] (Figs. 1, S1 and Tables 1, S1, S4). Notably, the easiness of their acquisition of the plane-symmetric architecture ($\Delta\Delta G_{TS}=1.68$ and 1.54 kcal·mol⁻¹, respectively, T=298.15 K [39]) enables them to be effectively incorporated into the structure of the DNA double helix.

Both T·T(w) \leftrightarrow T·T*(WC) and T·T(w) \leftrightarrow T*_{O2}·T(WC) tautomerisation processes are controlled by the plane-symmetric, highly polar, highly stable and zwitterionic TS^{T+·T-}_{T·T(w) \leftrightarrow T·T*(WC)} and TS^{T-·T+}_{T·T(w) \leftrightarrow T*_{O2}·T(WC)} transition states, respectively (Figs. 1, S1 and Tables S1-S4). Their structures are stabilized by strong electrostatic interactions (~ 100 kcal·mol⁻¹) and four intermolecular H-bonds [41], whose contribution to the interaction energy is not decisive (Tables S1, S3). High stability of these TSs excludes direct participation in these tautomerisation processes of the endogenous water molecules as active participants of these events [42] and allows to consider the influence of the stacking interactions with neighboring Watson-Crick base pairs [5,6,43,44] on the course of these tautomerisation reactions as slight perturbation. It is important that both T·T(w) \leftrightarrow T·T*(WC) and T·T(w) \leftrightarrow T*_{O2}·T(WC) tautomerisation processes occur by the non-dissociative mechanism (Figs. S2, S3) and are controlled by the 12 and 10 unique patterns of the specific intermolecular contacts, respectively, that successively replace each other along the IRC of tautomerisation (Fig. 2 and Tables S1, S5).

Another important finding is that the T·T*_{O2}(WC) mispair is in the rapid tautomeric equilibrium with the T*_{O2}·T(WC) mismatch ($\tau_{99.9\%}=3.88\cdot 10^{-10}$ s), that is implemented by the mechanism of the concerted synchronous double proton transfer (DPT) [45,46] along neighboring H-bonds (Figs. 1b, S1 and Tables 1, S1-S4).

At the same time, the DPT tautomerisation of the T·T(w) DNA mismatch into the T*·T*_{O2}(w) mispair with C_s symmetry actually does not happen, since the terminal T*·T*_{O2}(w) mismatch is short-lived ($\tau=5.61\cdot 10^{-15}$ s), dynamically unstable structure, for which low-frequency intermolecular vibrations can't develop (Figs. 1b, S1 and Tables 1, S1-S4).

By comparison of the energetic and kinetic characteristics for the T·T(w) \leftrightarrow T·T*(WC) and T·T(w) \leftrightarrow T·T*_{O2}(WC) tautomerisation processes, we have established that exactly first of them is responsible for the acquisition by the T·T(w) mispair of the enzymatically competent T·T*(WC) conformation in the hydrophobic recognition pocket of the high-fidelity replicative DNA-polymerase, that is a necessary and sufficient condition for the occurrence of the corresponding spontaneous transversions [11-13].

In the isolated state the T·T(w) mispair has a pseudo-two-fold symmetry. Calculations show that the transition between these pseudo-symmetrical states is provided by the non-dissociative mechanism

of the tautomeric nature (Fig. 3): exactly it, in our opinion, adequately explains existing experimental data [47] that have been left without any microstructural interpretation for a long period of time.

We have also established for the first time (Figs. 4, S4 and Tables 1, S1-S3, S6) that propeller-like and non-rigid C·C*(WC) mismatch with C_1 symmetry [19] slowly tautomerises into the C·C(w) ($\angle C4N3(C)N3C4(C^*)=9.0^\circ$) and C*·C*(w) (C_s) mispairs through the highly polar, highly stable and zwitterionic $TS^{C^-C^+}_{C\cdot C(w)\leftrightarrow C\cdot C^*(WC)}$ ($\angle C4N3(C^-)N3C4(C^+)=4.3^\circ$) and $TS^{C^-C^+}_{C\cdot C^*(WC)\leftrightarrow C^*\cdot C^*(w)}$ (C_s) transition states, respectively. These TSs are stabilized by strong electrostatic interactions (~ 100 kcal·mol⁻¹ (Table S3)) involving intermolecular H-bonds of medium strength that, similarly to the transformations of the T·T(w) mispair, excludes direct participation of the water molecules in these tautomerisation processes and indicates the negligible impact of the π - π stacking interactions [5,43,44] and sequence [6] on the considered tautomerisation processes.

Our results show that the C·C(w) \leftrightarrow C·C*(WC) and C·C*(WC) \leftrightarrow C*·C*(w) reactions occur by the non-dissociative mechanism (Figs. S5, S6) and are accompanied by 11 and 10 unique patterns of the specific intermolecular contacts including H-bonds, attractive van der Waals contacts and loosened covalent bridges, respectively, that consistently change each other along the IRC of tautomerisation (Fig. 5 and Tables S1, S7).

Moreover, we have also tested the C*·C*(w) mispair for the possibility of its tautomerisation *via* the DPT along intermolecular H-bonds according to Löwdin's mechanism [48,49]. Consequently, it was established that the C*·C*(w) \leftrightarrow C*_{O2}·C(w) DPT tautomerisation does not occur, since the C*_{O2}·C(w) mismatch is dynamically unstable structure ($\tau=2.45\cdot 10^{-13}$ s), for which low-frequency intermolecular vibrations can not develop (Figs. 4, S4 and Tables 1, S6).

Biological importance of the C·C(w) \rightarrow C·C*(WC) tautomeric conversion relies in the fact that it is kinetically controlled pathway for the formation directly in the essentially hydrophobic recognition pocket of the high-fidelity DNA-polymerase of the enzymatically competent C·C*(WC) mispair responsible for the occurrence of spontaneous point CC incorporation errors.

Our results raise at least two interesting issues representing a challenge for the experiment. Firstly, – in what proportions coexist in double helical DNA highly polar C·C(w), C·C*(WC), which in the free state is a global minimum, and plane symmetric C*_{O2}·C(w) mispairs? Secondly, since all these three mispairs have pseudo-two-fold symmetry, so the question arises about the mechanisms of their mutual transformation in the composition of the double helical DNA. Calculations indicate that these processes proceed *via* the non-dissociative mechanisms of the tautomeric origin (Fig. 6).

Our data are in a good agreement with other theoretical calculations [28]. Thus, fixation in the DNA duplex of the C·C(w) mispair, stabilized by the N4H \cdots N3 H-bond and N3 \cdots O2 van der Waals contact, and the T·T(w) mismatch, joined by the N3H \cdots O4 and N3H \cdots O2 H-bonds (Table S1), can be attributed to the fact that they are transformed into the C·C*(WC) and T·T*(WC) mismatches, respectively, through the tautomeric rearrangement accompanied with structural rebuilding of the base

mispairs (Figs. 3 and 6). This finding can't be reflected by the MD calculations operating solely with the canonical tautomers [28].

Obtained data shed light on the microstructural mechanisms of the occurrence of the spontaneous transversions caused by the homo-pyrimidine DNA mismatches and recognition of the latest by the repair systems [50-52].

Numerical evaluation of the probability ratio of the acquisition of the enzymatically competent conformation by the T·T(w) mispair relatively to the analogical probability for the C·C(w) mismatch ($2.3 \cdot 10^2$) is consistent with experimental data showing values of the extending efficiencies that are higher by several orders for the T·T, than for the C·C transversion mispairs [53-55], and also significantly lower frequencies of spontaneous transversions in comparison with spontaneous transitions [56,57]. At this we have applied known formulas of physico-chemical kinetics [58,59]: thus, we have suggested that $k_f=1.09 \cdot 10^{-10}/k_r=4.19 \cdot 10^{-4} \text{ s}^{-1}$ (T·T(w) \leftrightarrow T·T*(WC)), $k_f=1.59 \cdot 10^{-6}/k_r=4.71 \cdot 10^{-13} \text{ s}^{-1}$ (C·C(w) \leftrightarrow C·C*(WC)) and high-fidelity DNA-polymerase spends time $\Delta t_{\text{pol}}^{\text{wrong}} \approx 8.3 \cdot 10^{-3} \text{ s}$ [60,61] for the incorporation of one incoming incorrect nucleotide. Notably, we do not compare here absolute values of the mentioned above probabilities with corresponding literature data, since we are convinced that for their correct estimation it should be necessarily taken into account quantum tunneling of protons [62], which significantly accelerates tautomerisation processes.

Our findings also add substantially to the understanding, why the pyrimidine-pyrimidine T·T(w) mismatches are much better deleted from the genome by the repair systems "sharpened" for the wobble architecture than the C·C(w) mispairs (see works [28, 63] and refs. therein): the point is that the T·T(w) \leftrightarrow T·T*(WC) tautomeric equilibrium is much more shifted to the left than the C·C(w) \leftrightarrow C·C*(WC) tautomeric balance.

Acknowledgments. O.O.B. 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) and by the grant of the NAS of Ukraine for the realization of the scientific and research work of the young scientists of the NAS of Ukraine for 2015-2016 years. This work was performed using computational facilities of joint computer cluster of SSI "Institute for Single Crystals" of the National Academy of Sciences of Ukraine and Institute for Scintillation Materials of the National Academy of Sciences of Ukraine incorporated into Ukrainian National Grid. The authors sincerely thank Dr. Ivan S. Voiteshenko (Institute of High Technologies, Taras Shevchenko National University of Kyiv) and Dr. Fernando R. Clemente (Gaussian, Inc.) for their technical support of the work.

References.

1. J. D. Watson and F. H. C. Crick (1953). The structure of DNA. *Cold Spring Harb. Symp. Quant. Biol.*, 18, 123–131.
2. R. C. von Borstel (1994). Origins of spontaneous base substitutions. *Mut. Res.*, 307, 131-140.
3. E. B. Freese (1959). On the molecular explanation of spontaneous and induced mutations. *Brookhaven Symp. Biol.*, 12, 63-73.
4. Y. Lin, H. Wang, Y. Wu, S. Gao and H. F. Schaefer III (2014). Proton-transfer in hydrogenated guanine-cytosine trimer neutral species, cations, and anions embedded in B-form DNA. *Phys. Chem. Chem. Phys.*, 16, 6717-6725.

5. J. P. Cerón-Carrasco and D. Jacquemin (2015). DNA spontaneous mutation and its role in the evolution of GC-content: assessing the impact of the genetic sequence *Phys. Chem. Chem. Phys.*, *17*, 7754-7760.
6. D. Jacquemin, J. Zúñiga, A. Requena and J. P. Céron-Carrasco (2014). Assessing the importance of proton transfer reactions in DNA. *Acc. Chem. Res.*, *47*, 2467-2474.
7. M. D. Topal and J. R. Fresco (1976). Complementary base pairing and the origin of substitution mutation. *Nature*, *263*, 285-289.
8. V. I. Poltev and V. I. Bruskov (1977). Molecular mechanisms of spontaneous transversions and transitions. *Mol. Biol.*, *11*, 661-670.
9. O. O. Brovarets' and D. M. Hovorun (2010). Quantum chemical investigation of the elementary molecular mechanisms of pyrimidine-purine transversions. *Ukr. Biochem. J.*, *82*, 55-60.
10. 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. *J. Comput. Chem.*, *34*, 2577-2590.
11. K. Bebenek, L. C. Pedersen and T. A. Kunkel (2011). Replication infidelity *via* a mismatch with Watson-Crick geometry. *Proc. Natl. Acad. Sci. USA*, *108*, 1862-1867.
12. W. Wang, H. W. Hellinga and L. S. Beese (2011). Structural evidence for the rare tautomer hypothesis of spontaneous mutagenesis. *Proc. Natl. Acad. Sci. USA*, *108*, 17644-17648.
13. S. Xia, J. Wang and W. H. Konigsberg (2013). DNA mismatch synthesis complexes provide insights into base selectivity of a B family DNA polymerase. *J. Am. Chem. Soc.*, *135*, 193-202.
14. O. O. Brovarets' and D. M. Hovorun (2015). How many tautomerisation pathways connect Watson-Crick-like G*·T DNA base mispair and wobble mismatches? *J. Biomol. Struct. Dynam.* DOI: 10.1080/07391102.2015.1046936.
15. O. O. Brovarets' and D. M. Hovorun (2015). Tautomeric transition between wobble A·C DNA base mispair and Watson-Crick-like A·C* mismatch: microstructural mechanism and biological significance. *Phys. Chem. Chem. Phys.*, *17*, 15103-15110.
16. V. H. Harris, C. L. Smith, W. J. Cummins, A. L. Hamilton, H. Adams, M. Dickman, D. P. Hornby and D. M. Williams (2003). 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.*, *326*, 1389-1401.
17. H. Robinson, Y.-G. Gao, C. Bauer, C. Roberts, C. Switzer and A.H.-J. Wang (1998). 2'-Deoxyisoguanosine adopts more than one tautomer to form base pairs with thymidine observed by high-resolution crystal structure analysis. *Biochemistry*, *37*, 10897-10905.
18. I. J. Kimsey, K. Petzold, B. Sathyamoorthy, Z. W. Stein and H. M. Al-Hashimi (2015). Visualizing transient Watson-Crick-like mispairs in DNA and RNA duplexes. *Nature*, *519*, 315-320.
19. 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. *Phys. Chem. Chem. Phys.*, *15*, 20091-20104.
20. J. C. Lee and R. R. Gutell (2004). Diversity of base-pair conformations and their occurrence in rRNA structure and RNA structural motifs. *J. Mol. Biol.*, *344*, 1225-1249.
21. R. R. Gutell and C. R. Woese (1990). Higher-order structural elements in ribosomal RNAs: pseudoknots and the use of non-canonical pairs. *Proc. Natl. Acad. Sci. USA*, *87*, 663-667.
22. O. O. Brovarets', R. O. Zhurakivsky and D. M. Hovorun (2014). Structural, energetic and tautomeric properties of the T·T*/T*·T DNA mismatch involving mutagenic tautomer of thymine: A QM and QTAIM insight. *Chem. Phys. Lett.*, *592*, 247-255.
23. 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. *J. Phys. Chem. B*, *114*, 1454-1461.

24. O. O. Brovarets' and D. M. Hovorun (2010). Stability of mutagenic tautomers of uracil and its halogen derivatives: The results of quantum-mechanical investigation. *Biopolym. Cell*, *26*, 295–298.
25. N. Peyret, P. A. Seneviratne, H. T. Allawi and J. SantaLucia Jr (1999). Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A·A, C·C, G·G, and T·T mismatches. *Biochemistry*, *38*, 3468–3477.
26. A. Tikhomirova, I. V. Beletskaya and T. V. Chalikian (2006). Stability of DNA duplexes containing GG, CC, AA, and TT mismatches. *Biochemistry*, *45*, 10563–10571.
27. M. Kouchakdjian, B. F. L. Li, F. P. Swann and J. D. Patel (1988). Pyrimidine-pyrimidine base-pair mismatches in DNA. A nuclear magnetic resonance study of T·T pairing at neutral pH and C·C pairing at acidic pH in dodecanucleotide duplexes. *J. Mol. Biol.*, *202*, 139–155.
28. G. Rossetti, P. D. Dans, I. Gomez-Pinto, I. Ivani, G. Gonzalez and M. Orozco (2015). The structural impact of DNA mismatches. *Nucl. Acids Res.*, *43*, 4309–4321.
29. 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.
30. J. Tirado-Rives and W. L. Jorgensen (2008). Performance of B3LYP Density Functional Methods for a large set of organic molecules. *J. Chem. Theor. Comput.*, *4*, 297–306.
31. M. J. Frisch, M. Head-Gordon and J. A. Pople (1990). Semi-direct algorithms for the MP2 energy and gradient. *Chem. Phys. Lett.*, *166*, 281–289.
32. R. F. W. Bader (1990). *Atoms in molecules: A quantum theory*. Oxford: Oxford University Press.
33. C. F. Matta (2014). 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.*, *35*, 1165–1198.
34. C. Lecomte, E. Espinosa and C. F. Matta (2015). On atom–atom ‘short contact’ bonding interactions in crystals. *IUCrJ*, *2*, 161–163.
35. C. Peng, P. Y. Ayala, H. B. Schlegel and M. J. Frisch (1996). Using redundant internal coordinates to optimize equilibrium geometries and transition states. *J. Comput. Chem.*, *17*, 49–56.
36. 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. *J. Mol. Model.*, *19*, 4223–4237.
37. 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. *Phys. Chem. Chem. Phys.*, *16*, 9074–9085.
38. 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. *Phys. Chem. Chem. Phys.*, *16*, 3715–3725.
39. O. O. Brovarets' (2012). Structural-dynamical properties of pyrimidine-pyrimidine DNA base pairs with *cis*-oriented glycosidic bonds: a quantum-chemical study. *Ukr. Bioorg. Acta*, *10*, 39–46.
40. T. Y. Nikolaienko, L. A. Bulavin and D. M. Hovorun (2011). How flexible are DNA constituents? The quantum-mechanical study. *J. Biomol. Struct. Dynam.*, *29*, 563–575.
41. O. O. Brovarets', Y. P. Yurenko 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. *J. Biomol. Struct. Dynam.*, *33*, 1624–1652.
42. 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. *Phys. Chem. Chem. Phys.*, *13*, 4311–4317.
43. J. P. Cerón-Carrasco, J. Zúñiga, A. Requena, E. A. Perpète, C. Michaux and D. Jacquemin (2011). Combined effect of stacking and solvation on the spontaneous mutation in DNA. *Phys. Chem. Chem. Phys.*, *13*, 14584–14589.

44. G. Villani (2014). Coupling between hydrogen atoms transfer and stacking interaction in adenine-thymine/guanine-cytosine complexes: A theoretical study *J. Phys. Chem. B*, 118, 5439-5452.
45. 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. *J. Biomol. Struct. Dynam.*, 33, 28-55.
46. 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. *Chem. Phys. Lett.*, 578, 126-132.
47. G. He, C. K. Kwok and S. L. Lam (2011). Preferential base pairing modes of T·T mismatches. *FEBS Lett.*, 585, 3953-3958.
48. P.-O. Löwdin (1963). Proton tunneling in DNA and its biological implications. *Rev. Mod. Phys.*, 35, 724-732.
49. 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. *J. Biomol. Struct. Dynam.*, 32, 1474-1499.
50. E. C. Friedberg (2003). DNA damage and repair. *Nature*, 421, 436-440.
51. M. Sharma, A. V. Predeus, S. Mukherjee and M. Feig (2013). DNA bending propensity in the presence of base mismatches: implications for DNA repair. *J. Phys. Chem. B*, 117, 6194-6205.
52. M. Feig, B. Wang, M. Sharma, Z. Burton, K. Opron, R. Cukier, A. Predeus, N. Kovacs, S. Law and Sh. Mukherjee (2015). 10 How is fidelity maintained in nucleic acids? Two tales in DNA repair and DNA transcription from computer simulations. *J. of Biomol. Struct. Dynam.* 33, Suppl. 1, 6-7.
53. M.-M. Huang, N. Arnheim and M. Goodman (1992). Extension of base mispairs by Taq DNA polymerase: implications for single nucleotide discrimination in PCR. *Nucl. Acids Res.*, 20, 4567-4573.
54. F. W. Perrino and L. A. Loeb (1989). Differential extension of 3' mispairs is a major contribution to the high fidelity of calf thymus DNA polymerase- α . *J. Biol. Chem.*, 264, 2898-2905.
55. T. A. Kunkel and P. S. Alexander (1986). The base substitution fidelity of eucaryotic DNA polymerases. *J. Biol. Chem.*, 261, 160-166.
56. H. Lee, E. Popodi, H. Tang and P. L. Foster (2012). Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. *Proc. Natl. Acad. Sci. USA*, 109, E2774- E2783.
57. M. Lynch (2010). Rate, molecular spectrum, and consequences of human mutation. *Proc. Natl. Acad. Sci. USA*, 107, 961-968.
58. Moore, W. J. (1983). *Basic Physical Chemistry*. Prentice-Hall: Englewood Cliffs, NJ.
59. 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. *Int. J. Mol. Sci.*, 4, 410-421.
60. L. Sun, R. I. Cukier and Y. Bu (2007). Factors determining the deriving force of DNA formation: geometrical differences of base pairs, dehydration of bases and the arginine assisting. *J. Phys. Chem.*, 111, 1802-1808.
61. 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.
62. A. D. Godbeer, J. S. Al-Khalili and P. D. Stevenson (2015). Modelling proton tunneling in the adenine-thymine base pair. *Phys. Chem. Chem. Phys.*, 17, 13034-13044.
63. C. Dohet, R. Wagner and M. Radman (1985) Repair of defined single base-pair mismatches in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, 82, 503-505.

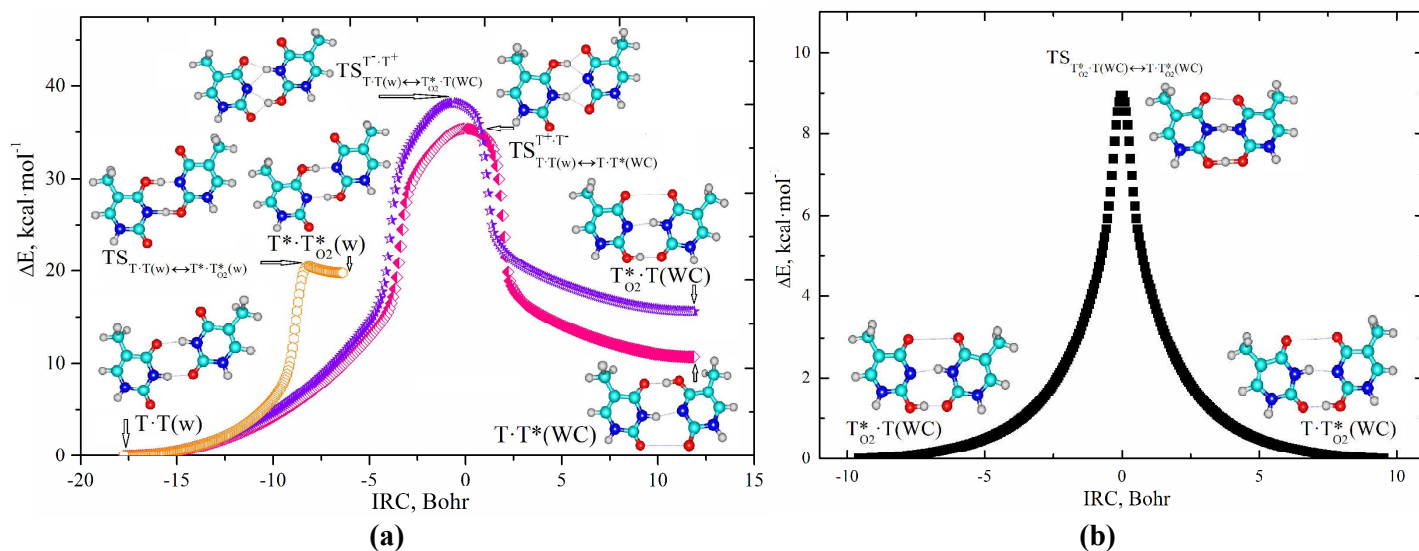


Fig. 1. Stationary structures on the reaction pathways of the (a) $T \cdot T(w) \leftrightarrow T \cdot T^*(WC)/T^* \cdot O_2 \cdot T(WC)/T^* \cdot T^*O_2(w)$ and (b) $T^* \cdot O_2 \cdot T(w) \leftrightarrow T \cdot T^*O_2(WC)$ tautomerisations *via* the sequential DPT accompanied by the structural rearrangement of the base pair obtained at the B3LYP/6-311++G(d,p) level of theory. ΔE - electronic energy of the base mispairs and TSs. Dotted lines indicate $AH \cdots B$ H-bonds and attractive $A \cdots B$ van der Waals contacts, while continuous – covalent bonds; carbon atoms are in light-blue, nitrogen – in dark-blue, hydrogen – in grey and oxygen – in red.

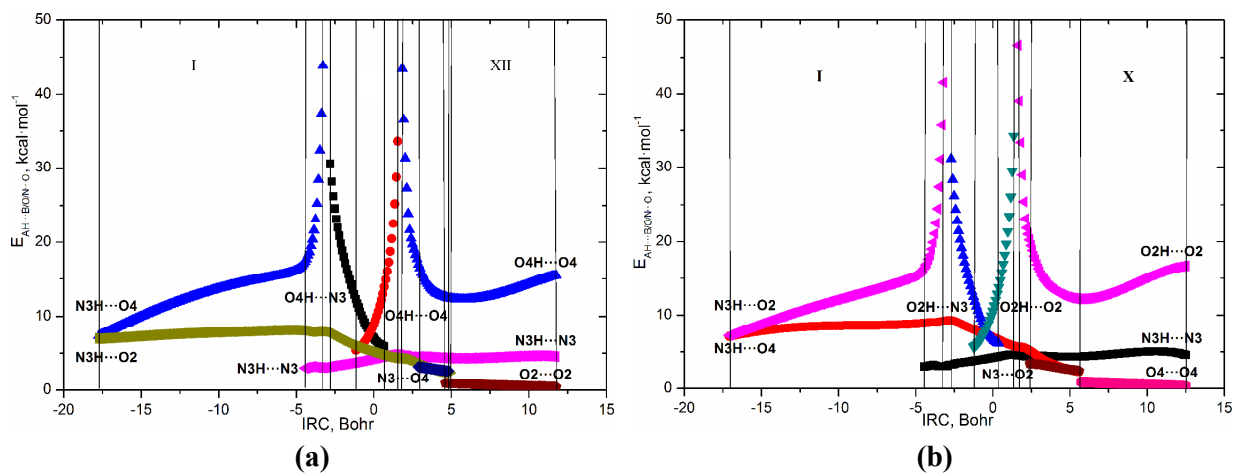


Fig. 2. Exchange of the patterns of the intermolecular $AH\cdots B$ H-bonds and attractive $O/N\cdots O$ van der Waals contacts (their energies $E_{AH\cdots B/O/N\cdots O}$ are estimated by the EML formula at the (3,-1) BCPs) along the IRC of the biologically important (a) $T\cdot T(w)\leftrightarrow T\cdot T^*(WC)$ and (b) $T\cdot T(w)\leftrightarrow T^*O_2\cdot T(WC)$ tautomerisations *via* the sequential DPT accompanied by the structural rearrangement of the base pair obtained at the B3LYP/6-311++G(d,p) level of theory (see Figs. 1, S1 and Tables S1, S5).

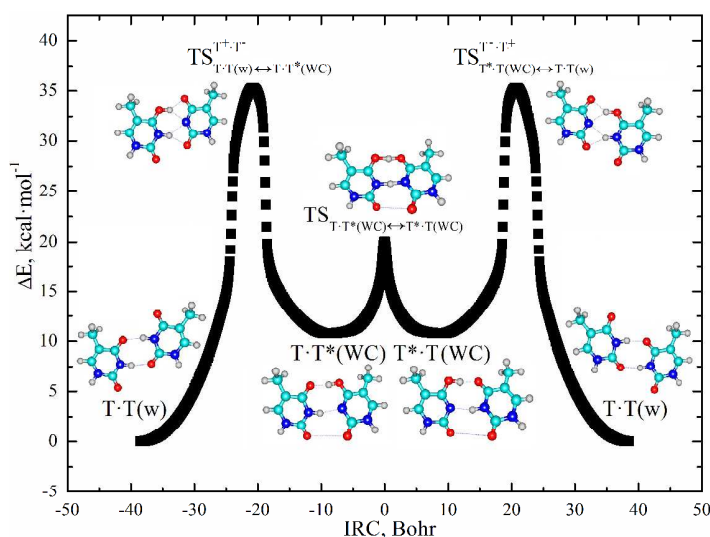


Fig. 3. Interconversion of the pseudo-twofold symmetrical $T\cdot T(w)$ DNA mismatches *via* the tautomeric transition. Presented isoenergetic structures are linked with each other by the pseudo-dyadic axis of symmetry. For the designations see Fig. 1.

Table 1. Energetic and kinetic characteristics of the $T \cdot T(w) \leftrightarrow T \cdot T^*(WC)$, $T \cdot T(w) \leftrightarrow T^*_{O2} \cdot T(WC)$, $T^*_{O2} \cdot T(w) \leftrightarrow T \cdot T^*_{O2}(WC)$, $T \cdot T(w) \leftrightarrow T^* \cdot T^*_{O2}(w)$, $C \cdot C(w) \leftrightarrow C \cdot C^*(WC)$, $C \cdot C^*(WC) \leftrightarrow C^* \cdot C^*(w)$ and $C^* \cdot C^*(w) \leftrightarrow C^*_{O2} \cdot C(w)$ tautomerisations *via* the sequential DPT accompanied by the structural rearrangement of the base pair obtained at the MP2/cc-pVQZ//B3LYP/6-311++G(d,p) level of theory (see also Figs. 1, 4, S1 and S4).

Tautomerisation	ΔG^a	ΔE^b	$\Delta\Delta G_{TS}^c$	$\Delta\Delta E_{TS}^d$	$\Delta\Delta G^e$	$\Delta\Delta E^f$	$\tau_{99.9\%}^g$
$T \cdot T(w) \leftrightarrow T \cdot T^*(WC)$	8.98	8.64	31.06	31.90	22.09	23.26	$1.65 \cdot 10^4$
$T \cdot T(w) \leftrightarrow T^*_{O2} \cdot T(WC)$	12.72	12.81	32.82	33.96	20.11	21.15	$5.82 \cdot 10^2$
$T^*_{O2} \cdot T(w) \leftrightarrow T \cdot T^*_{O2}(WC)$	0.00	0.00	4.40	7.56	4.40	7.56	$3.88 \cdot 10^{-10}$
$T \cdot T(w) \leftrightarrow T^* \cdot T^*_{O2}(w)$	15.61	15.99	13.82	17.03	-1.79	1.03	$3.88 \cdot 10^{-14}$
$C \cdot C(w) \leftrightarrow C \cdot C^*(WC)$	-8.90	-10.73	25.38	24.32	34.28	35.05	$4.35 \cdot 10^6$
$C \cdot C^*(WC) \leftrightarrow C^* \cdot C^*(w)$	1.34	1.50	18.34	19.44	17.00	17.94	2.90
$C^* \cdot C^*(w) \leftrightarrow C^*_{O2} \cdot C(w)$	6.07	6.36	6.59	9.27	0.53	2.91	$1.69 \cdot 10^{-12}$

^aThe Gibbs free energy of the product relatively the reactant of the tautomerisation reaction (T=298.15 K), kcal·mol⁻¹

^bThe electronic energy of the product relatively the reactant of the tautomerisation reaction, kcal·mol⁻¹

^cThe Gibbs free energy barrier for the forward reaction of tautomerisation, kcal·mol⁻¹

^dThe electronic energy barrier for the forward reaction of tautomerisation, kcal·mol⁻¹

^eThe Gibbs free energy barrier for the reverse reaction of tautomerisation, kcal·mol⁻¹

^fThe electronic energy barrier for the reverse reaction of tautomerisation, kcal·mol⁻¹

^gThe time necessary to reach 99.9% of the equilibrium concentration between the reactant and the product of the tautomerisation reaction, s

See also summary Table S2 for the Gibbs and electronic energies of the mispairs and TSs relatively the global minima – the $T \cdot T(w)$ and $C \cdot C^*(WC)$ mispairs.

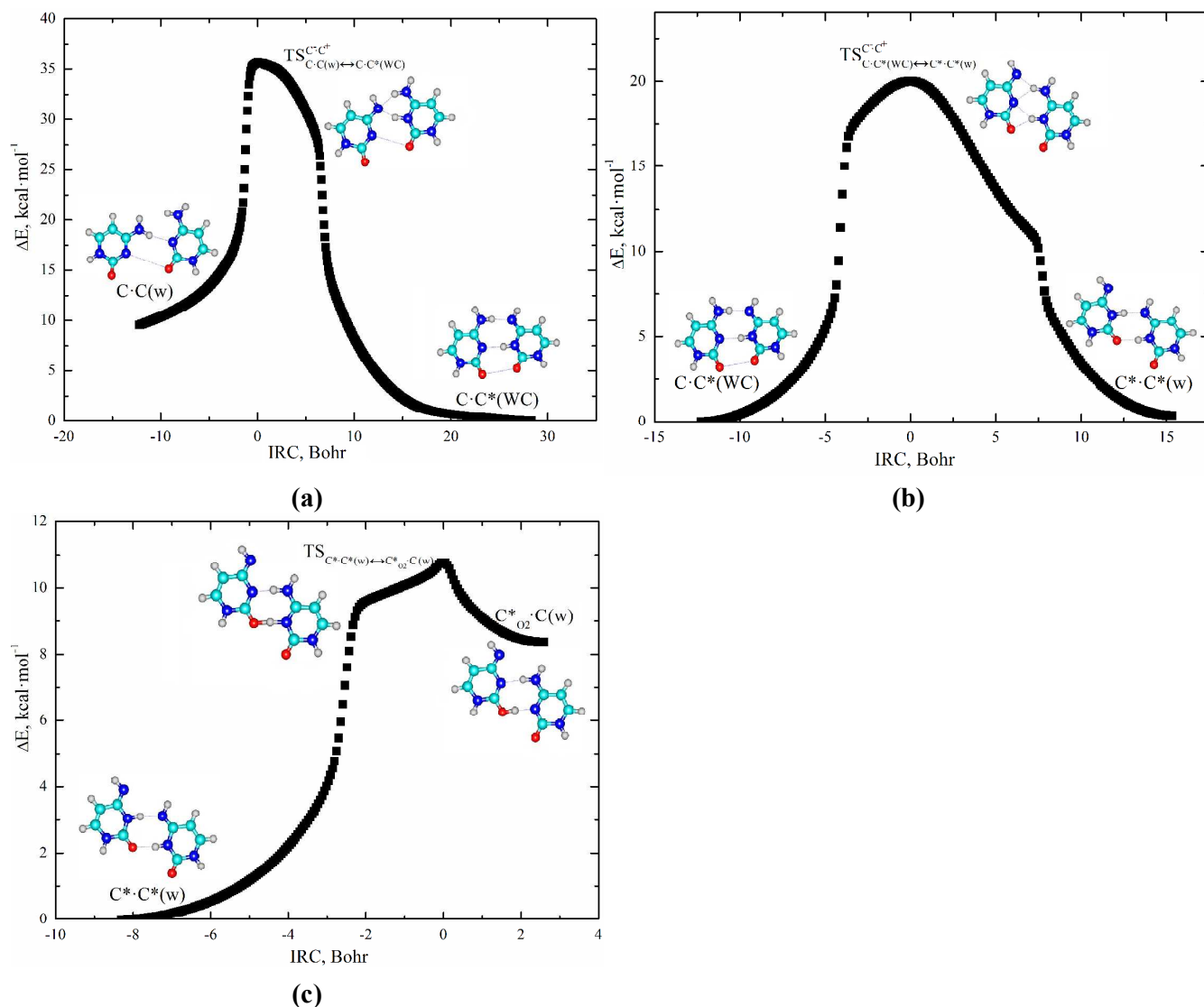


Fig. 4. Stationary structures on the reaction pathways of the (a) $C\cdot C(w) \leftrightarrow C\cdot C^*(WC)$, (b) $C\cdot C^*(WC) \leftrightarrow C^*\cdot C^*(w)$ and (c) $C^*\cdot C^*(w) \leftrightarrow C^*_{O2}\cdot C(w)$ tautomerisations *via* the sequential DPT accompanied by the structural rearrangement of the base pair obtained at the B3LYP/6-311++G(d,p) level of theory. For the designations see Fig. 1.

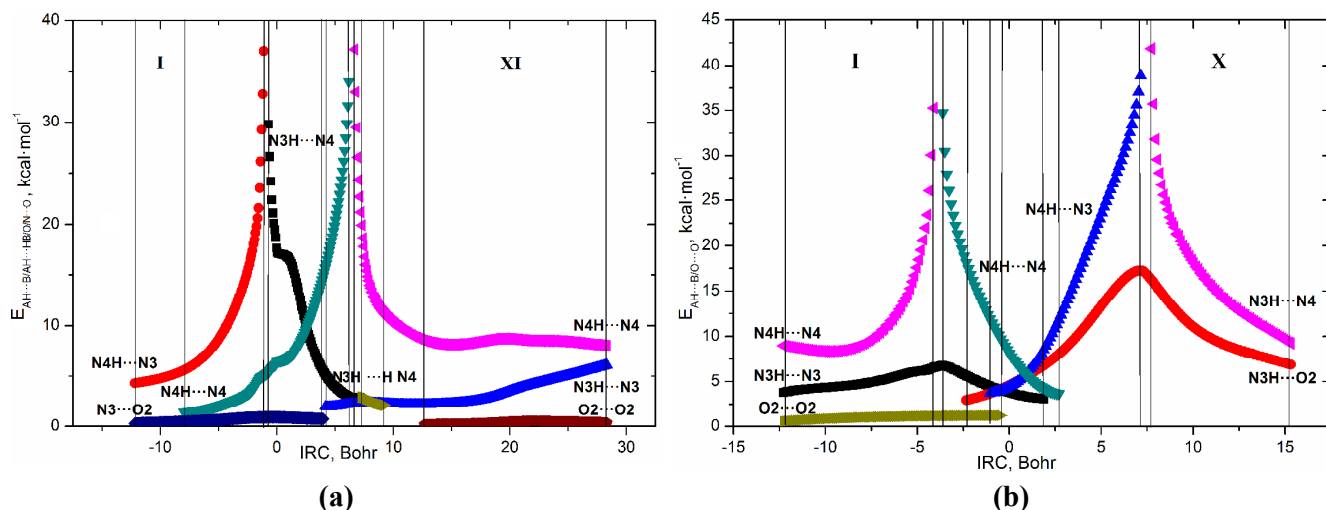


Fig. 5. Exchange of the patterns of the intermolecular $\text{AH}\cdots\text{B}$, $\text{AH}\cdots\text{HB}$ H-bonds and attractive $\text{O}/\text{N}\cdots\text{O}$ van der Waals contacts (their energies $E_{\text{AH}\cdots\text{B}/\text{AH}\cdots\text{HB}/\text{O}/\text{N}\cdots\text{O}}$ are estimated by the EML formula at the (3,-1) BCPs) along the IRC of the biologically important (a) $\text{C}\cdot\text{C}(\text{w}) \leftrightarrow \text{C}\cdot\text{C}^*(\text{WC})$ and (b) $\text{C}\cdot\text{C}^*(\text{WC}) \leftrightarrow \text{C}^*\cdot\text{C}^*(\text{w})$ tautomerisations *via* the sequential DPT accompanied by the structural rearrangement of the base pair obtained at the B3LYP/6-311++G(d,p) level of theory (see Figs. 4, S4 and Tables S1, S7).

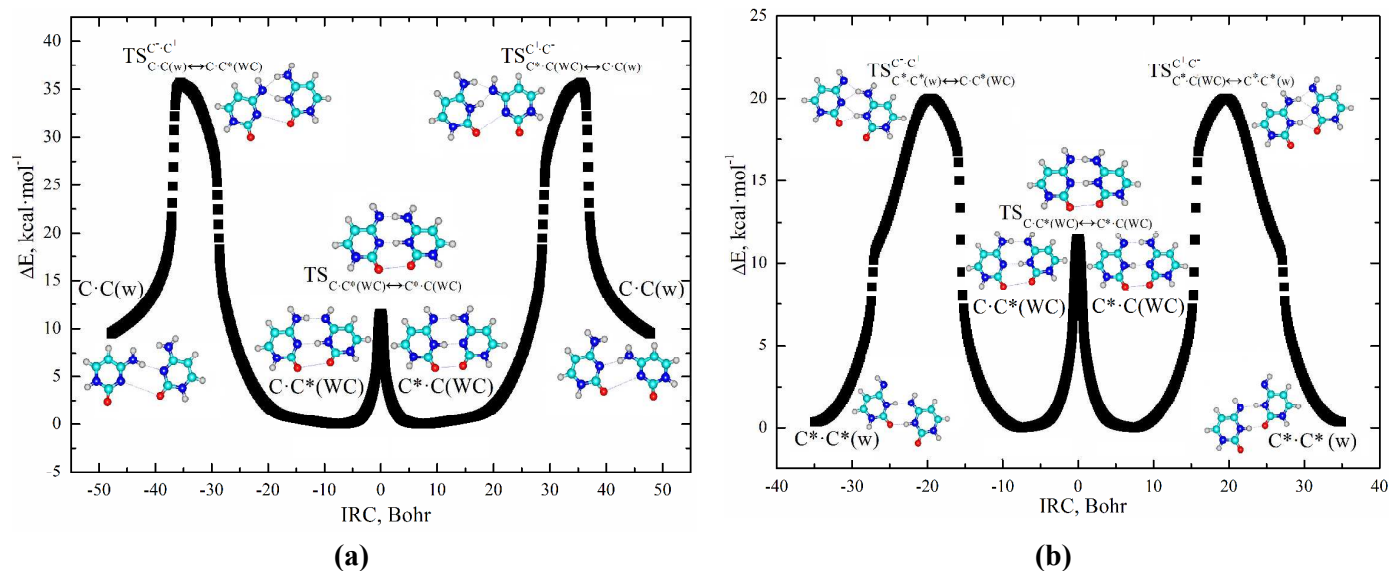


Fig. 6. Interconversion of the pseudo-twofold symmetrical (a) $\text{C}\cdot\text{C}(\text{w})$ and (b) $\text{C}^*\cdot\text{C}^*(\text{w})$ DNA mismatches *via* the tautomeric transition. Presented isoenergetic structures are linked with each other by the pseudo-dyadic axis of symmetry. For the designations see Fig. 1.