J Mol Model (2013) 19:4119–4137
DOI 10.1007/s00894-012-1720-9

ORIGINAL PAPER

The physico-chemical "anatomy" of the tautomerization through the DPT of the biologically important pairs of hypoxanthine with DNA bases: QM and QTAIM perspectives

Ol'ha O. Brovarets' · Roman O. Zhurakivsky · Dmytro M. Hovorun

Received: 16 October 2012 / Accepted: 3 December 2012 / Published online: 5 January 2013

Abstract The biologically important tautomerization of the Hyp·Cyt, Hyp*·Thy and Hyp·Hyp base pairs to the Hyp*·Cyt*, Hyp·Thy* and Hyp*·Hyp* base pairs, respectively, by the double proton transfer (DPT) was comprehensively studied *in vacuo* and in the continuum with a low dielectric constant (ε=4) corresponding to hydrophobic interfaces of protein–nucleic acid interactions by combining theoretical investigations at the B3LYP/6-311++G(d,p) level of QM theory with QTAIM topological analysis. Based on the sweeps of the energetic, electron-topological, geometric and polar parameters, which describe the course of the tautomerization along the intrinsic reaction coordinate (IRC), it was proved that the tautomerization through the DPT is concerted and asynchronous process for the Hyp·Cyt and Hyp*·Thy base pairs, while concerted and synchronous for the Hyp·Hyp homodimer. The continuum with ε=4 does not affect qualitatively the course of the tautomerization reaction for all studied complexes. The nine key points along the IRC of the Hyp·Cyt↔Hyp*·Cyt* and Hyp*·Thy↔Hyp·Thy* tautomerizations and the six key points of the Hyp·Hyp↔Hyp*·Hyp* tautomerization have been identified and fully characterized. These key points could be considered as electron-topological "fingerprints" of concerted asynchronous (for Hyp·Cyt and Hyp*·Thy) or synchronous (for Hyp·Hyp) tautomerization process *via* the DPT. It was found, that in the Hyp*·Cyt*, Hyp·Thy*, Hyp·Hyp and Hyp*·Hyp* base pairs all H-bonds are significantly cooperative and mutually reinforce each other, while the C2H…O2 H-bond in the Hyp·Cyt base pair and the O6H…O4 H-bond in the Hyp*·Thy base pair behave anti-cooperatively, i.e., they become weakened, while two others become strengthened.

Keywords Base pairs with Watson-Crick-like geometry that contains hypoxanthine · Cooperativity of H-bonds · B3LYP level of QM theory · Sweeps of the energetic · Electron-topological · Geometric and polar parameters · Point mutations in DNA · Prototropic tautomerizm · QTAIM analysis · The double proton transfer

Electronic supplementary material The online version of this article (doi:10.1007/s00894-012-1720-9) contains supplementary material, which is available to authorized users.

O. O. Brovarets' · R. O. Zhurakivsky · D. M. Hovorun
Department of Molecular and Quantum Biophysics, Institute of Molecular Biology and Genetics, National Academy of Sciences of Ukraine, 150 Zabolotnoho Str, 03680 Kyiv, Ukraine

O. O. Brovarets' · R. O. Zhurakivsky · D. M. Hovorun
Research and Educational Center "State Key Laboratory of Molecular and Cell Biology", 150 Zabolotnoho Str, 03680 Kyiv, Ukraine

O. O. Brovarets' · D. M. Hovorun (✉)
Department of Molecular Biology, Biotechnology and Biophysics, Institute of High Technologies, Taras Shevchenko National University of Kyiv, 2 Hlushkova Ave, 03127 Kyiv, Ukraine
e-mail: dhovorun@imbg.org.ua

Introduction

Though accidental changes in a genomic DNA sequence, called mutations [1], can be harmful to a cell, in which they occur, changing the functions of proteins and causing aging, illnesses and inherited diseases [1–4] in some circumstances, may be beneficial providing the fuel for phenotypic evolution [5] in others.

Spontaneous mutations generally occur during DNA replication due to a variety of sources such as endogenous chemical lesions generated spontaneously during normal

cell metabolism, errors in normal cellular processes and transposable genetic elements [5]. Spontaneous point mutations may arise due to the substitution of one nucleotide base for another [4].

A common cause of spontaneous point mutations on the molecular level is the deamination of the canonical DNA base to an atypical (modified) base containing a keto-imino group in place of the original amino group [1, 2]. Thus, for example, the oxidative deamination of the adenine (Ade) [6–11] leads to the formation of the hypoxanthine (Hyp) [11, 12] that possesses a structure similar to DNA base guanine (Gua) lacking its exocyclic 2-amino group and so has the H-bonding pattern, which is complementary to cytosine (Cyt) [7, 9, 13–19]. In such a way, Hyp is able to pair with Cyt forming a Hyp·Cyt base pair during DNA replication [7–9, 13–21]. So, the deamination of Ade leads to the Ade·Thy→Gua·Cyt transition [7, 8, 15, 19, 22–24]. However, a review of the current literature indicates a lack of both experimental and theoretical data concerning the molecular mechanisms underlying the transitions induced by the Hyp [15, 23–25].

In our previous work [25] we discovered the mechanism of the prototropic tautomerizm of Hyp, which is known to play a fundamental role in the mutagenesis as the formation of rare tautomers (maked with an asterisk) of Hyp can induce alterations in the normal base pairing, leading to changes in the genetic code [26]. We have investigated the mechanism of the intermolecular tautomerization of the most energetically favorable Hyp·Hyp homodimer (symmetry C_{2h}), stabilized by the two equivalent N1H…O6 H-bonds, *via* the DPT [25]. Moreover, we established the molecular mechanism of the first discovered keto-enol tautomerization of the Hyp·Hyp homodimer (C_{2h}) *via* the zwitterionic near-orthogonal transition state (TS), stabilized by the $N1^{+}H…N1^{-}$ and $O6^{+}H…N1^{-}$ H-bonds, to the Hyp*·Hyp heterodimer (C_s), stabilized by the O6H…O6 and N1H…N1 H-bonds. We first showed that the Hyp*·Thy mispair (C_s), stabilized by the O6H…O4, N3H…N1 and C2H…O2 H-bonds, mimicking Watson-Crick base pairing, converts to the wobble Hyp·Thy base pair (C_s), stabilized by the N3H…O6 and N1H…O2 H-bonds, *via* the high- and low-energy TSs and the Hyp·Thy* intermediate, stabilized by the O4H…O6, N1H…N3 and C2H…O2 H-bonds. The most energetically favorable TS of the Hyp*·Thy→Hyp·Thy tautomerization is the Hyp^{+}·Thy^{-} zwitterionic pair (C_s), stabilized by the $O6^{+}H…O4^{-}$, $O6^{+}H…N3^{-}$, $N1^{+}H…N3^{-}$ and $N1^{+}H…O2^{-}$ H-bonds. We also provided investigation of the Hyp·Cyt→Hyp*·Cyt* tautomerization *via* the DPT both in vacuum and in solution.

Based on the obtained results, it was expressed and substantiated the hypothesis, that the keto tautomer of Hyp is a mutagenic compound, while enol tautomer Hyp* does not possess mutagenic properties [25]. The lifetime of the nonmutagenic tautomer Hyp* exceeds by many orders the time required to complete a round of DNA replication in the cell. An influence of the surrounding environment (ε=4) corresponding to a hydrophobic interfaces of protein–nucleic acid interactions [25, 27–34] on the stability of studied complexes and corresponding TSs was established to be negligible [25].

Since these stationary points for the base pairs and TSs of their interconversion were located, the reaction pathways were established by following the IRC in the forward and reverse directions from each TS to ensure the expected reactants and products on each side of the TS.

However, these data cannot give an exhaustive understanding of the detailed physico-chemical mechanisms of the tautomerization and mutagenic action of Hyp as a product of the Ade deamination in DNA.

In our recent work [27] devoted to elementary physico-chemical mechanisms of the tautomerization of the Ade·Thy Watson-Crick DNA base pair, we first applied a novel approach, based on the sweeps along the IRC of such important physico-chemical characteristics of the DPT as energetic, electron-topological, polar and geometric. We have investigated the evolution of these characteristics of the H-bonds and base pairs along the reaction pathway establishing them at the each point of the IRC.

For the first time, the nine key points along the IRC of the Ade·Thy base pair tautomerization, which could be considered as electron-topological "fingerprints" of the concerted asynchronous process of the Ade·Thy→Ade*·Thy* tautomerization *via* the DPT, have been identified and fully characterized. We first suggested that the nine key points are characteristic for any asynchronous tautomerization process *via* the DPT in the H-bonded complexes. Based on this approach, it becomes possible to clearly distinguish reactant, transition state and product regions [27, 35, 36] of the DPT in the Ade·Thy base pair separated by key points 2 and 8, where the $\Delta\rho$=0 at the BCPs of the corresponding H-bonds. It was established, that the extrema of the first derivative of the electronic energy along the IRC dE/dIRC (the so-called reaction force [35, 36]) is reached exactly at these key points.

Investigation of the energy dependencies of the H-bonds taken at the beginning and at the end of the movement of protons in the Ade·Thy base pair along the IRC, which stabilize the Ade·Thy Watson-Crick and Ade*·Thy* Löwdin's DNA base pairs, respectively, allowed us to numerically characterize such an important characteristic of intermolecular H-bonds as their cooperativity [37, 38]. It was found that all H-bonds are cooperative in the Ade·Thy base pair, reinforcing each other, while the C2H…O2 H-bond behaves anti-cooperatively in the Ade*·Thy* base pair, that is it gets weakened, while two others get strengthened.

The aim of this work is to extend our approach [27] to a wider range of biologically important objects, such as the H-

bonded pairs of Hyp with DNA bases. Realization of this goal is important for understanding the mechanisms of genomic instability associated with point mutations, caused by the tautomerization of nucleotide bases. Biologically important base pairs involving Hyp and DNA bases – namely, Hyp·Cyt, Hyp*·Cyt*, Hyp*·Thy, Hyp·Thy*, Hyp·Hyp, Hyp*·Hyp* – investigated earlier [25], serve as the objects of this study.

The biologically important tautomerization of the Hyp·Cyt, Hyp*·Thy and Hyp·Hyp base pairs to the Hyp*·Cyt*, Hyp·Thy* and Hyp*·Hyp* base pairs, respectively, by the DPT was comprehensively studied *in vacuo* and in the continuum with a low dielectric constant (ε=4) corresponding to a hydrophobic interfaces of protein–nucleic acid interactions [25, 27–34] by combining theoretical investigations at the B3LYP/6-311++G(d,p) level of QM theory with QTAIM topological analysis. Based on the sweeps of the energetic, electron-topological, geometric and polar parameters, which describe the course of the tautomerization along the IRC, it was proved that the tautomerization through the DPT is concerted and asynchronous process for the Hyp·Cyt and Hyp* - Thy base pairs, while concerted and synchronous process for the Hyp·Hyp homodimer. The limiting stage for the Hyp·Cyt→Hyp*·Cyt* and Hyp*·Thy→Hyp·Thy* tautomerizations is the final proton transfer along the N4H…O6 and O6H…O4 H-bonds, respectively, exposed in the major groove of the double-stranded DNA. The continuum with ε=4 does not affect qualitatively the course of the tautomerization reaction for all studied complexes.

The nine and six key points for the Hyp·Cyt, Hyp*·Thy and Hyp·Hyp base pairs, respectively, which could be considered as electron-topological "fingerprints" of their tautomerization *via* the DPT, have been identified and fully characterized along the IRC. These key points have been used to define the reactant, transition state and product regions of the DPT in the investigated base pairs.

We have revealed for the first time that the third C2H…O2 H-bond in the Hyp·Cyt and Hyp*·Thy base pairs assists their tautomerization *via* the DPT.

It was found, considering the energy dependence of each of the H-bonds along the IRC in the investigated base pairs, that in the Hyp*·Cyt*, Hyp·Thy*, Hyp·Hyp and Hyp*·Hyp* base pairs all H-bonds are significantly cooperative [37, 38] and mutually reinforce each other, while in the Hyp·Cyt and Hyp*·Thy base pairs the C2H…O2 and O6H…O4 H-bonds, respectively, behave anti-cooperatively, i.e., they become weakened, while two others become strengthened.

The obtained results enable us to regard the proposed approach [27] not only as a powerful tool for the study of the mechanisms underlying the tautomerization of any H-bonded complex *via* the DPT, but also as a method for researching the cooperativity of the H-bonds that stabilize them [37, 38].

Computational methods

All calculations have been carried out with the Gaussian 09 suite of programs [39]. Geometries and harmonic vibrational frequencies of the base pairs and the TSs of their tautomerization were obtained using density functional theory (DFT) with the B3LYP hybrid functional [40], which includes Becke's three-parameter exchange functional (B3) [41] combined with Lee, Yang and Parr's (LYP) correlation functional [42] in connection with Pople's 6-311++G(d,p) basis set. A scaling factor of 0.9668 has been used in the present work at the B3LYP level of theory to correct the harmonic frequencies of all the studied base pairs. We performed single point energy calculations at the correlated MP2 level of theory [43] with the 6-311++G(2df,pd) basis set for B3LYP/6-311++G(d,p) geometries to consider electronic correlation effects as accurately as possible. MP2/6-311++G(2df,pd)//B3LYP/6-311++G(d,p) level of theory has been successfully applied on similar systems recently studied and has been verified to give accurate normal mode frequencies, barrier heights, characteristics of intra- and intermolecular H-bonds and geometries [4, 25, 27, 28, 44–61]. Moreover, an excellent agreement between computational and experimental NMR, UV and IR spectroscopic data [51, 55, 56, 60, 61] evidences that the applied level of theory for the single-point energy calculations (MP2/6-311++G(2df,pd)), as well as the method employed for geometry optimization (B3LYP/6-311++G(d,p)) are reliable.

In order to take into account the impact of the surrounding effect on the tautomerization process of the investigated complexes, we have repeated the geometry optimizations at the B3LYP/6-311++G(d,p) level of theory [50] using the conductor-like polarizable continuum model (CPCM) [62, 63], choosing the continuum with a dielectric constant of ε=4 to mimic hydrophobic interfaces of protein-nucleic acids interactions [30–34].

The correspondence of the stationary points to minimum on the potential energy landscape or TS has been checked by the absence or the presence, respectively, of one and only one imaginary frequency corresponding to the normal mode that identifies the reaction coordinate. TSs were located by means of synchronous transit-guided quasi-Newton (STQN) method [64, 65].

Since the stationary points were located, the reaction pathway was established by following the IRC in the forward and reverse directions from each TS using the Hessian-based predictor-corrector (HPC) integration algorithm [66–68] 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 have investigated the evolution of the energetic, geometric, polar and electron-topological characteristics of the H-bonds and base pairs along the

reaction pathway establishing them at each point of the IRC in vacum and in the continuum with ε=4.

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

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

where E_{el} - the electronic energy, while E_{corr} – thermal correction. We applied the standard transition state theory [69] to estimate activation barriers of the tautomerization reactions.

Bader's quantum theory "atoms in molecules" (QTAIM) was applied to analyze electron density [70]. The topology of the electron density was analyzed using program package AIMAll [71] with all the default options. The presence of a bond critical point [70], namely the so-called (3,-1) bond critical point (BCP) and a bond path between hydrogen donor and acceptor, as well as the positive value of the Laplacian of the electron density at this BCP ($\Delta\rho \geq 0$), were considered as criteria for H-bond formation. Wave functions were obtained at the level of theory used for geometry optimization.

The energies of the non-conventional CH…O H-bonds and conventional H-bonds (at the investigation of the corresponding sweeps) were evaluated by the empirical Espinosa-Molins-Lecomte (EML) formula [72, 73] based on the electron density distribution at the (3,-1) BCPs of the H-bonds:

$$E_{HB} = 0.5 \cdot V(r), \quad (2)$$

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

The energy of the O6H…O4 H-bond in the $TS_{Hyp^*\cdot Thy \leftrightarrow Hyp\cdot Thy^*}$ were estimated by the Nikolaienko-Bulavin-Hovorun formula [74]:

$$E_{HB} = -2.03 + 225 \cdot \rho, \quad (3)$$

where ρ is the electron density at the (3,-1) BCP of the O6H…O4 H-bond.

The energies of the conventional H-bonds were evaluated by the empirical Iogansen's formula [75]:

$$E_{HB} = 0.33 \cdot \sqrt{\Delta\nu - 40}, \quad (4)$$

where $\Delta\nu$ - the magnitude of the redshift (relative to the free molecule) of the stretching mode of H-bonded AH groups in the AH…B H-bond (A and B – N, O). The partial deuteration was applied to minimize the effect of vibrational resonances [4, 25, 27, 28, 44–52, 55–57, 60].

The atomic numbering scheme for the purine and pyrimidine bases is conventional [76].

Results and discussion

The obtained results are collected in Table 1 and presented in Figs. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15 (see also Figs. S1-S39 in the Supplementary materials). They were analyzed separately for each studied base pair.

The Hyp·Cyt, Hyp*·Cyt*, Hyp*·Thy and Hyp·Thy* Watson-Crick-like base pairs are stabilized by three intermolecular H-bonds each, while the Hyp·Hyp and Hyp*·Hyp* homodimers are joined by two equivalent H-bonds: their basic physico-chemical characteristics are presented in Table 1. It draws attention to the fact that at the Hyp·Cyt→Hyp*·Cyt*, Hyp*·Thy→Hyp·Thy* and Hyp·Hyp→Hyp*·Hyp* tautomerizations the total energy of the intrapair H-bonds increases in 1.27/1.37, 1.12/1.09 and 1.47/1.58 times, respectively (hereinafter in this Article ε=1/ε=4) (Table 1).

In the Hyp·Cyt and Hyp*·Thy base pairs the average N1H…N3 (7.05/7.58) and N3H…N1 (7.23/6.76 kcalmol^{-1}) H-bonds, respectively, are the strongest ones. It should be noted that the DPT in this base pairs starts from the single PT along these H-bonds.

The classical geometric criteria for the identification of H-bonds are satisfied for all conventional AH…B H-bonds, which physico-chemical parameters are presented in Table 1.

The spectroscopic data collected in Table 1 confirm geometric results. The shift in frequency of the stretching mode of the AH donor group (the difference between the frequency of the AH group in monomer and the frequency of this group involved in H-bonding) is mostly positive (shift to the red) for all conventional H-bonds, while negative (shift to the blue) for the CH…O H-bonds.

All studied intrapair H-bonds without exceptions show positive values of the Laplacian of the electron density $\Delta\rho$ at the BCPs of the AH…B H-bonds and the values of the electron density ρ at this BCPs, usually treated as a measure of H-bonding strength [70, 77], are situated within the range 0.002÷0.127/0.002÷0.102 a.u., at that the minimum value of the ρ corresponds to the C2H…O2 H-bond in the Hyp*·Thy base pair both in vacuum and in the continuum with ε=4, while the maximum value of the ρ in vacuum corresponds to the N4H…O6 H-bond in the $TS_{Hyp\cdot Cyt \leftrightarrow Hyp^*\cdot Cyt^*}$ and in the continuum with ε=4 corresponds to the O6H…O4 H-bond in the $TS_{Hyp^*\cdot Thy \leftrightarrow Hyp\cdot Thy^*}$. This clearly shows that generally the N4H…O6 H-bond in the $TS_{Hyp\cdot Cyt \leftrightarrow Hyp^*\cdot Cyt^*}$ and the O6H…O4 H-bond in the $TS_{Hyp^*\cdot Thy \leftrightarrow Hyp\cdot Thy^*}$ are the strongest interactions, while the C2H…O2 H-bond in the Hyp*·Thy base pair is the weakest one.

In the Hyp·Cyt, Hyp*·Cyt*, Hyp*·Thy, Hyp·Thy* base pairs and in the $TS_{Hyp\cdot Cyt \leftrightarrow Hyp^*\cdot Cyt^*}$ and the $TS_{Hyp^*\cdot Thy \leftrightarrow Hyp\cdot Thy^*}$ of their interconversion *via* the DPT we revealed non-conventional or so-called weak [78] C2H…O2 H-bonds with the energy in the range 0.40÷0.97/0.32÷0.98 kcalmol^{-1}, which is acceptable for a weak H-bonds of the CH…O type (Table 1) [4, 25, 28, 57].

Table 1 Electron-topological, structural, vibrational and energetic characteristics of the intermolecular H-bonds in pairs of Hyp with DNA bases and TSs of their tautomerization *via* the DPT obtained at the B3LYP/6-311++G(d,p) level of theory in vacuum and in the continuum with ε=4

Complex	H-bond AH…B	ρ[a]		$\Delta\rho$[b]		$100\cdot\varepsilon$[c]		$d_{A\ldots B}$[d]		$d_{H\ldots B}$[e]		$\angle AH\ldots B$[f]		Δd_{AH}[g]		$\Delta\nu$[h]		E_{HB}[i]	
		ε=1	ε=4	ε=1	ε=4	ε=1	ε=4	ε=1	ε=4	ε=1	ε=4	ε=1	ε=4	ε=1	ε=4	ε=1	ε=4	ε=1	ε=4
Hyp·Cyt	N4H…O6	0.032	0.028	0.109	0.097	3.94	4.76	2.861	2.919	1.833	1.897	175.9	176.0	0.021	0.016	363.6	286.7	5.94	5.18
	N1H…N3	0.039	0.041	0.096	0.095	6.76	6.81	2.876	2.872	1.837	1.828	176.5	178.1	0.028	0.031	496.5	567.2	7.05	7.58
	C2H…O2	0.005	0.005	0.015	0.016	5.14	2.23	3.637	3.603	2.858	2.807	128.8	130.2	-0.0001	0.00006	-10.6	-15.8	0.79*	0.89*
Hyp*·Cyt*	O6H…N4	0.056	0.068	0.100	0.093	5.17	4.99	2.689	2.633	1.676	1.604	173.8	172.7	0.048	0.065	916.4	1240.1	9.77	11.43
	N3H…N1	0.040	0.035	0.094	0.090	6.27	6.41	2.878	2.921	1.839	1.889	173.9	172.1	0.031	0.025	528.6	473.0	7.29	6.87
	C2H…O2	0.003	0.002	0.010	0.008	9.75	18.14	3.861	3.973	3.069	3.207	130.3	128.4	0.001	0.00067	5.5	4.7	0.49*	0.36*
$TS_{Hyp\cdot Cyt\leftrightarrow Hyp^*\cdot Cyt^*}$	N4H…O6	0.127	–	0.027	–	2.29	–	2.498	–	1.317	–	176.2	–	0.172	–	1840.4	–	14.00	–
	N3H…N1	0.077	0.077	0.071	0.071	5.10	6.14	2.683	2.834	1.574	1.788	176.3	172.1	0.095	–	1493.9	–	12.58	11.71*
	C2H…O2	0.004	0.004	0.014	0.014	6.80	24.24	3.672	3.933	2.894	3.187	128.7	126.6	-0.006	–	-86.2	–	0.76*	0.76*
Hyp*·Thy	O6H…O4	0.043	0.047	0.131	0.138	2.48	2.62	2.697	2.660	1.708	1.669	173.2	171.1	0.025	0.029	495.4	599.0	7.04	7.80
	N3H…N1	0.039	0.036	0.093	0.091	6.14	6.17	2.886	2.927	1.851	1.897	172.2	171.2	0.029	0.025	520.3	459.8	7.23	6.76
	C2H…O2	0.002	0.002	0.009	0.008	14.54	19.02	3.936	4.019	3.150	3.257	129.9	128.2	0.001	0.00036	4.8	2.5	0.40*	0.32*
Hyp·Thy*	O4H…O6	0.054	0.052	0.143	0.142	2.17	2.64	2.623	2.635	1.616	1.635	174.0	171.7	0.041	0.036	766.8	714.8	8.90	8.57
	N1H…N3	0.041	0.039	0.098	0.095	6.20	6.20	2.854	2.875	1.824	1.843	170.0	170.2	0.028	0.028	493.6	507.7	7.03	7.14
	C2H…O2	0.003	0.003	0.011	0.010	27.37	21.77	3.793	3.831	3.087	3.123	123.4	123.6	-0.001	-0.00007	-13.7	-18.8	0.49*	0.45*
$TS_{Hyp^*\cdot Thy\leftrightarrow Hyp\cdot Thy^*}$	O6H…O4	0.114	0.102	0.078	0.108	1.84	2.12	2.447	2.460	1.350	1.381	173.6	173.0	–	–	–	–	22.15**	23.62**
	C2H…O2	0.006	0.006	0.017	0.017	2.99	2.06	3.559	3.564	2.763	2.767	130.0	130.1	–	–	–	–	0.97*	0.98*
Hyp·Hyp	N1H…O6	0.039	0.037	0.123	0.117	3.34	3.78	2.798	2.824	1.762	1.790	176.9	177.1	0.025	0.022	430.3	405.6	6.52	6.31
Hyp*·Hyp*	O6H…N1	0.056	0.058	0.099	0.098	4.99	4.96	2.698	2.690	1.686	1.675	175.3	175.4	0.045	0.047	879.0	953.1	9.56	9.97

[a] The electron density at the BCP, a.u

[b] The Laplacian of the electron density at the BCP, a.u

[c] The ellipticity at the BCP

[d] Distance between A (a H-bond donor) and B (a H-bond acceptor) atoms, Å

[e] The distance between H and B atoms, Å

[f] The H-bond angle, degree

[g] The elongation of the H-bond donating group AH upon H-bonding, Å

[h] The redshift (positive value) or the blueshift (negative value) of the stretching vibrational mode of H-bonded AH group, cm^{-1}

[i] H-bond energy, estimated by Iogansen's [60], Espinosa-Molins-Lecomte (marked with an asterisk) [57, 58] and Nikolaienko-Bulavin-Hovorun (marked with a double asterisk) [59] formulae, $kcalmol^{-1}$

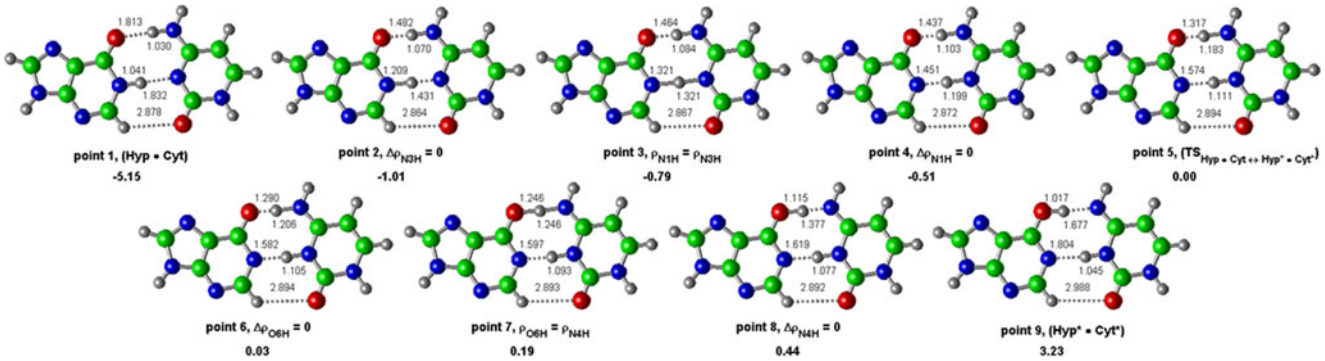

Fig. 1 Geometric structures of the nine key points describing the evolution of the Hyp·Cyt→Hyp*·Cyt* tautomerization *via* the DPT along the IRC (see also Fig. S1 in the Supplementary materials) obtained at the DFT level of theory *in vacuo*. Coordinates of the nine key points are presented in Bohr for each structure. The dotted lines indicate H-bonds, while continuous lines show covalent bonds (their lengths are presented in angstroms). Carbon atoms are in *green*, nitrogen in *blue*, hydrogen in *grey* and oxygen in *red*

We have performed the IRC calculations both in vacuum and in the continuum with ε=4 corresponding to a hydrophobic interfaces of protein–nucleic acid interactions [25–34] to map out a reaction pathway of the Hyp·Cyt↔Hyp*·Cyt*, Hyp*·Thy↔Hyp·Thy* and Hyp·Hyp↔Hyp*·Hyp* transformations and to test whether the optimized TS structures are connected to the products (the Hyp*·Cyt*, Hyp·Thy* and Hyp*·Hyp* base pairs) and the reactants (the Hyp·Cyt, Hyp*·Thy and Hyp·Hyp base pairs) of the tautomerization reactions starting from the corresponding TSs downhill, both in forward and reverse directions (Figs. 2,7 and 12).

The pathways of the Hyp*·Thy↔Hyp·Thy* and Hyp·Hyp↔Hyp*·Hyp* transformations yield any dynamically stable intermediate as in vacuum, so in the continuum with ε=4, while the Hyp·Cyt↔Hyp*·Cyt* transformation occurs without any intermediate in vacuum and proceeds through the zwitterionic $Hyp^{-}\cdot Cyt^{+}$ base pair, which is dynamically unstable [25, 79], in the continuum with ε=4.

Tautomerization of the Hyp·Cyt base pair into the Hyp*·Cyt* mispair

A local minimum on the electronic energy surface corresponding to the zwitterionic $Hyp^{-}\cdot Cyt^{+}$ base pair is stabilized by the N4H…O6 (21.05), N3H…N1 (22.92) and C2H…O2 (0.96 $kcal\,mol^{-1}$) H-bonds and observed at the IRC=-4.58 Bohr in the continuum with ε=4 (Table 1). The zwitterionic $Hyp^{-}\cdot Cyt^{+}$ base pair formally can not be considered as an intermediate, due to the negligible value of the electronic energy of the reverse barrier of the Hyp·Cyt→$Hyp^{-}\cdot Cyt^{+}$ transformation (41.9 cm^{-1} obtained at the B3LYP/6-311++G(d,p) level of theory) or absence of this barrier (-69.5 cm^{-1} obtained at the MP2/6-311++G(2df,pd)//B3LYP/6-311++G(d,p) level of theory) [25].

To characterize the evolution of the Hyp·Cyt→Hyp*·Cyt* tautomerization *via* the DPT along the IRC as comprehensively as possible, we have identified in vacuum and in the continuum with ε=4 the nine key points, three of which are stationary - Hyp·Cyt, Hyp*·Cyt* and $TS_{Hyp\cdot Cyt\rightarrow Hyp^{*}\cdot Cyt^{*}}$ (Figs. 1 and S1). The energetic, geometric, electron-topological and polar characteristics of these nine key points along the IRC are exhaustively presented in Figs. 1–5 and S1-S16.

It should be noted that these nine key points differ in vacuum and in the continuum with ε=4.

Key point 1. The starting structure along the IRC pathway is the Hyp·Cyt base pair with Watson-Crick-like geometry. It is stabilized by the N4H…O6, N1H…N3 and C2H…O2 H-bonds (Table 1, Figs. 1 and S1).

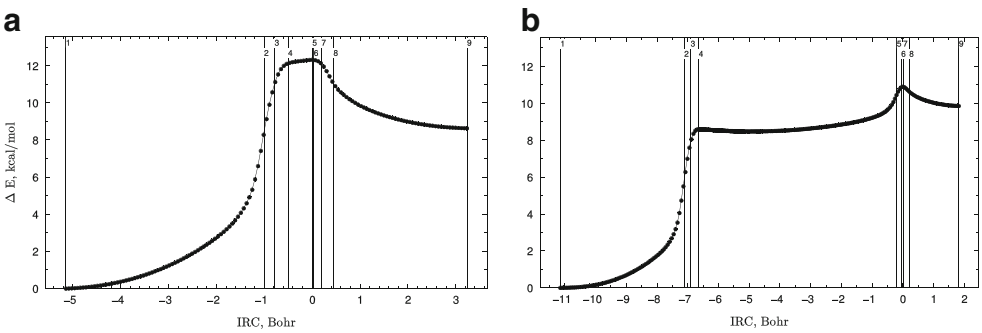

Fig. 2 Energy profiles of the Hyp·Cyt→Hyp*·Cyt* tautomerization along the IRC obtained at the DFT level of theory **a** *in vacuo* and **b** in the continuum with a low dielectric constant (ε=4). Coordinates of the nine key points are listed in Figs. 1 and S1

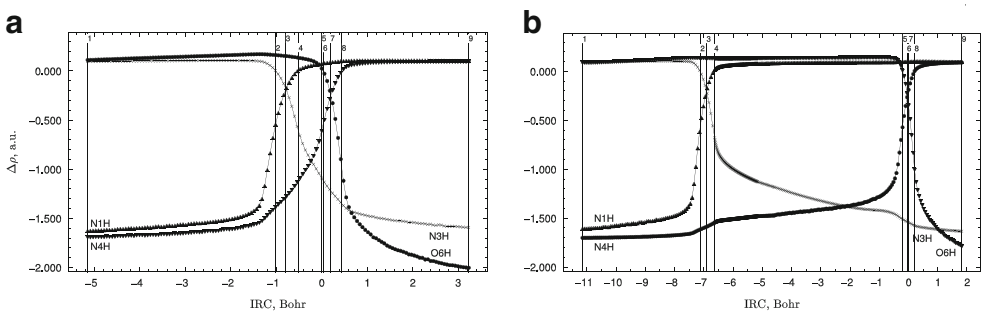

Fig. 3 Profiles of the Laplacian of the electron density $\Delta\rho$ at the BCPs along the IRC of the Hyp·Cyt→Hyp*·Cyt* tautomerization obtained at the DFT level of theory **a** *in vacuo* and **b** in the continuum with a low dielectric constant (ε=4)

Key point 2. The structure of the base pair corresponding to the situation in which the N1-H chemical bond of the Hyp base is significantly weakened and the N3…HH-bond actually becomes the N3-H covalent bond (Fig. 1). A characteristic feature of this structure is a zero value of the $\Delta\rho$ for the BCP of the N3…HH-bond (Fig. 3). The maximum value of the energy of the N3…HH-bond is attained at this key point (Fig. 4).

Key point 3. This structure is characterized by the equivalent loosened N3-H and N1-H covalent bonds that have equal values of the electron density ρ, the Laplacian of the electron density $\Delta\rho$ at the BCPs and the $d_{N3H/N1H}$ distances of the N3-H and N1-H covalent bonds. χ-like dependencies of electron-topological and geometric characteristics are observed for the loosened N3-H-N1 bridge (Figs. 3, 5 and S3).

Key point 4. At this key point the dynamically unstable [79] zwitterionic Hyp^{-}·Cyt^{+} base pair begins to form, i.e., the N1-H covalent bond begins to acquire characteristics of the N1…HH-bond (Fig. 1). A characteristic feature of this structure is a zero value of the $\Delta\rho$ for the BCP of the N1…HH-bond (Fig. 3). The maximum value of the energy of the N1…HH-bond is attained at this key point (Fig. 4).

Key point 5. In vacuum the $TS_{Hyp\cdot Cyt\leftrightarrow Hyp^{*}\cdot Cyt^{*}}$ of the Hyp·Cyt↔Hyp*·Cyt* tautomerization *via* the DPT, stabilized by the N4H…O6, N3H…N1 and C2H…O2 H-bonds, is observed at this key point (Table 1, Fig. 1).

In the continuum with ε=4 the N4-H chemical bond of the Cyt base is significantly weakened and the O6…HH-bond actually becomes the O6-H covalent bond (Table 1, Fig. S1). A characteristic feature of this structure is a zero value of the $\Delta\rho$ for the BCP of the O6…HH-bond (Fig. 3b). The maximum value of the energy of the O6…HH-bond is attained at this key point (Fig. 4b).

Key point 6. In vacuum is observed the structure corresponding to key point 5 in the continuum with ε=4 (Figs. 1 and S1).

In the continuum with ε=4 it is observed the structure that possesses equivalent loosened N4-H and O6-H covalent bonds that have equal values of the electron density ρ, the Laplacian of the electron density $\Delta\rho$ at the BCPs and the $d_{N4H/O6H}$ distances of the N4-H and O6-H covalent bonds. χ-like dependencies of

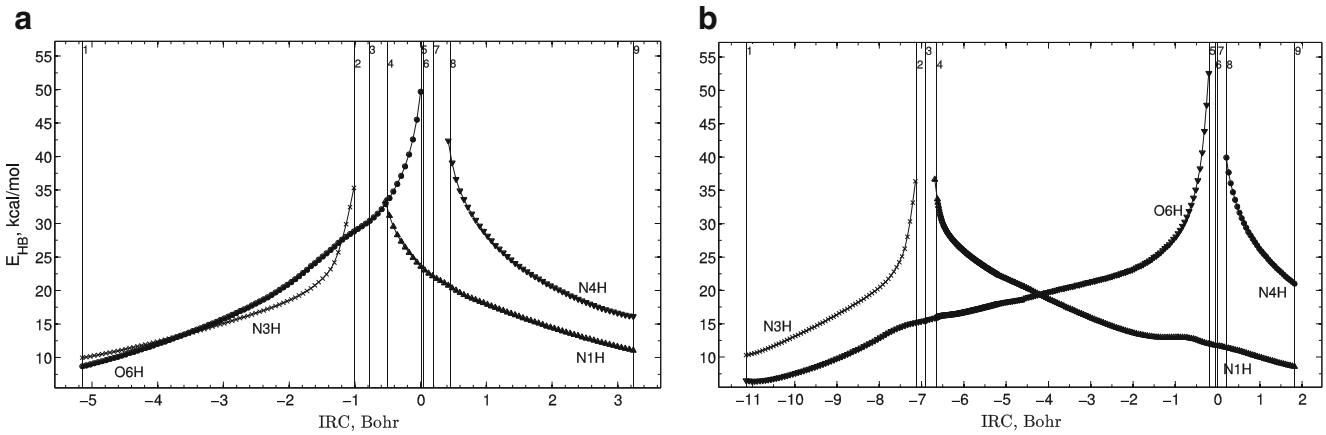

Fig. 4 Profiles of the energy of the H-bonds E_{HB}, estimated by the EML formula [72, 73], along the IRC of the Hyp·Cyt→Hyp*·Cyt* tautomerization obtained at the DFT level of theory **a** *in vacuo* and **b** in the continuum with a low dielectric constant (ε=4)

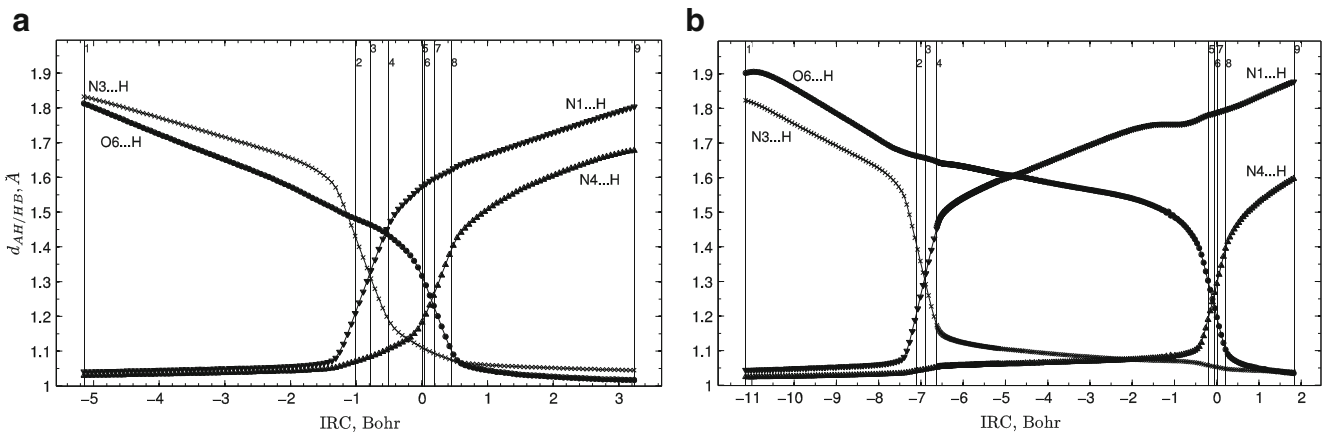

Fig. 5 Profiles of the distance $d_{AH/HB}$ between the hydrogen H and electronegative A or B atoms of the AH…B H-bond along the IRC of the Hyp·Cyt→Hyp*·Cyt* tautomerization obtained at the DFT level of theory **a** *in vacuo* and **b** in the continuum with a low dielectric constant (ε=4)

electron-topological and geometric characteristics are observed for the loosened N4-H-O6 bridge (Figs. 3b, 5b and S3). It should be noted that χ-like dependencies of the ρ, $\Delta\rho$ and the $d_{AH/HB}$ distances of the N3-H-N1 and N4-H-O6 bridges are shifted along the IRC, so the DPT at the tautomerization process occurs with a time gap (Figs. 3b, 5b and S3).

Key point 7. In vacuum is observed the structure corresponding to key point 6 in the continuum with ε=4 (Figs. 1 and S1).

In the continuum with ε=4 the $TS_{Hyp\cdot Cyt\leftrightarrow Hyp^*\cdot Cyt^*}$, stabilized by the N3H…N1 and C2H…O2 H-bonds and by the O6-H-N4 bridge, corresponds to this key point (Table 1, Fig. S1).

Key point 8. Both in vacuum and in the continuum with ε=4 at this key point the mispair containing rare tautomers of the Hyp and Cyt bases begins to form, i.e., the significantly weakened N4-H covalent bond begins to acquire characteristics of the N4…HH-bond (Figs. 1 and S1). A characteristic feature of this structure is a zero value of the $\Delta\rho$ for the BCP of the N4…HH-bond (Fig. 3). The maximum value of the energy of the N4…HH-bond is attained at this key point (Fig. 4).

Key point 9. The final structure in vacuum and in the continuum with ε=4 – the tautomerized Hyp*·Cyt* mispair. It is stabilized by the O6H…N4, N3H…N1 and C2H…O2 H-bonds (Table 1, Figs. 1 and S1).

These nine key points are used to define the reactant, transition state and product regions [27, 35, 36] of the DPT in the Hyp·Cyt base pair (Fig. 2).

Interestingly, the extrema of the first derivative of the electron energy with respect to the IRC dE/dIRC are reached exactly at key points 2 and 8 (Fig. S2).

Nucleotide bases do not lose their chemical individuality at *the reactant region* and acquire such mutual deformation and

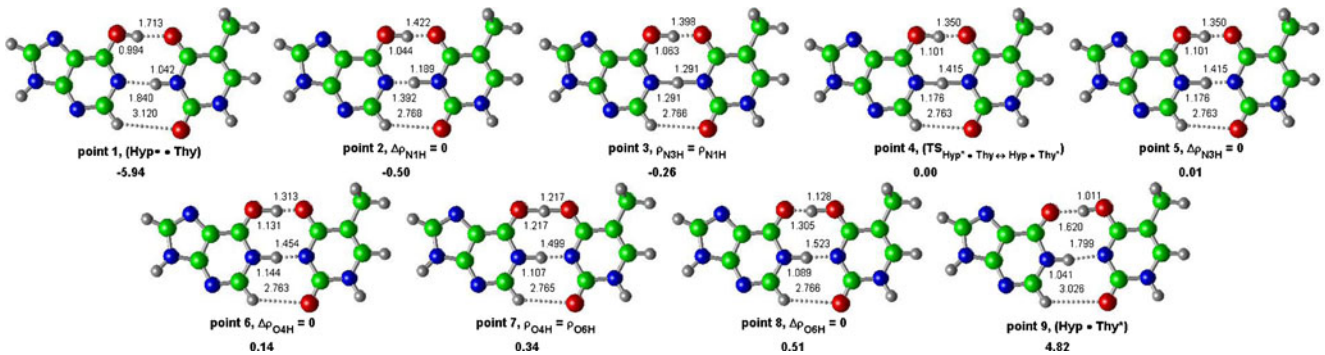

Fig. 6 Geometric structures of the nine key points describing the evolution of the Hyp*·Thy→Hyp·Thy* tautomerization *via* the DPT along the IRC obtained at the DFT level of theory *in vacuo* (see also Fig. S17 in the Supplementary materials). Coordinates of the nine key points are presented in Bohr for each structure. The *dotted lines* indicate H-bonds, while *continuous lines* show covalent bonds (their lengths are presented in angstroms). Carbon atoms are in *green*, nitrogen in *blue*, hydrogen in *grey* and oxygen in *red*

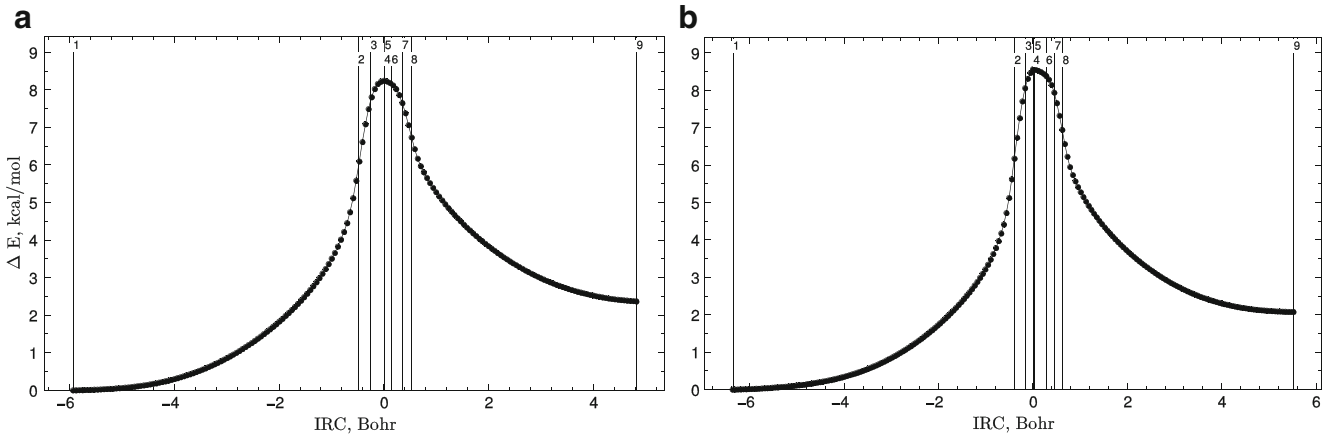

Fig. 7 Energy profiles of the Hyp*·Thy→Hyp·Thy* tautomerization along the IRC obtained at the DFT level of theory **a** *in vacuo* and **b** in the continuum with a low dielectric constant (ε=4). Coordinates of the nine key points are listed in Figs. 6 and S17

orientation, that eventually leads to the chemical reaction at *the transition state region*, namely to the DPT. It quite logically follows from our data (sweeps) on the $\Delta\rho$ that the reactant region starts at key point 1 and ends at key point 2, after passing which bases lose their chemical individuality and the DPT chemical reaction starts. We obtained that the electronic energy necessary to bring the donor and acceptor atoms as close as possible to each other to activate the DPT reaction, that is the energy difference between key points 2 and 1, is 8.28/5.89 kcalmol^{-1}. Actually, *the transition state region*, where the DPT occurs, is located between key points 2 and 8.

The product region, where rare tautomers of nucleotide bases do not lose their chemical individuality and where the relaxation to the final Hyp*·Cyt* base pair takes place, begins at key point 8 and ends at the final key point 9. Comparably small amount of energy 2.51/0.78 kcalmol^{-1}, that is the difference between the energy of key points 8 and 9, releases at the relaxation of the base pair, corresponding to key point 8, under its tautomerization into the reaction product – the Hyp*·Cyt* base pair.

Movement along the IRC of the Hyp·Cyt↔Hyp*·Cyt* tautomerization reaction in vacuum and in the continuum with ε=4 shows that this reaction involves no stable intermediates (it is *concerted*) and the DPT does not occur simultaneously (this event occurs *asynchronously*) including the concerted movement of two protons along the H-bridges that join the Hyp and Cyt bases. Moreover, we established that first the proton localized at the N1 nitrogen atom of the Hyp base migrates along the strongest N1H…N3H-bond [25] to the N3 nitrogen atom of the Cyt base inducing the formation of the unstable Hyp^{-}·Cyt^{+} zwitterion (Fig. 1). This migration further induces the transition of the proton at the N4 nitrogen atom of the Cyt to the O6 oxygen atom of the Hyp, that leads to the formation of the Hyp*·Cyt* mispair. The limiting stage of this phenomenon is the final proton transfer along the N4H…O6 H-bond exposed in the major groove of the double-stranded DNA. In the continuum with a low dielectric constant (ε=4) the mechanism of the DPT reaction

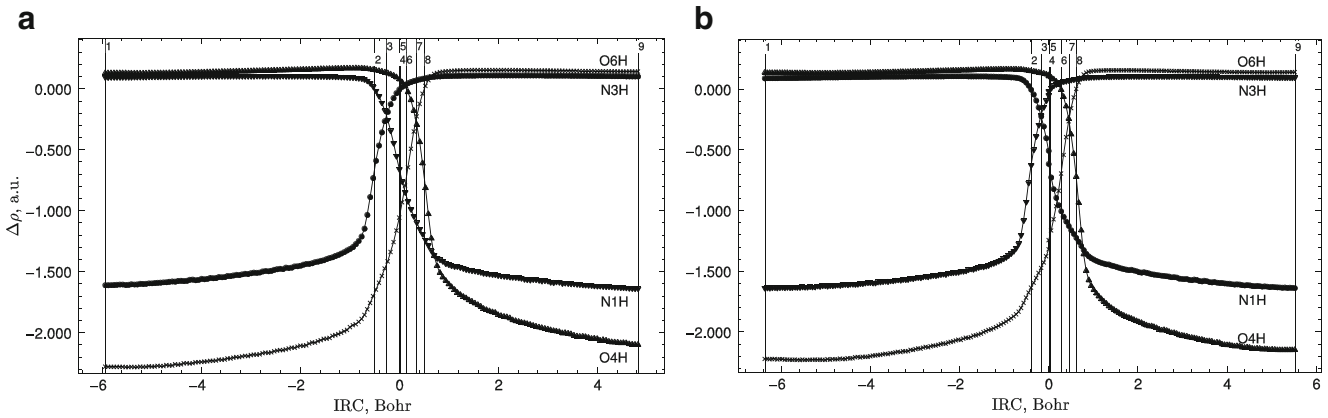

Fig. 8 Profiles of the Laplacian of the electron density $\Delta\rho$ at the BCPs along the IRC of the Hyp*·Thy→Hyp·Thy* tautomerization obtained at the DFT level of theory **a** *in vacuo* and **b** in the continuum with a low dielectric constant (ε=4)

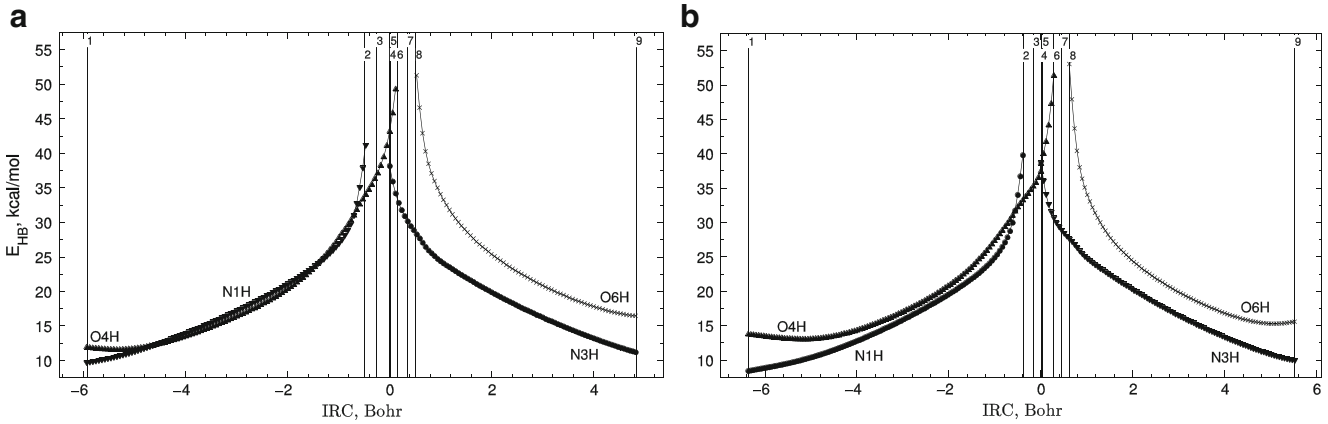

Fig. 9 Profiles of the energy of the H-bonds E_{HB}, estimated by the EML formula [72, 73], along the IRC of the Hyp*·Thy→Hyp·Thy* tautomerization obtained at the DFT level of theory **a** *in vacuo* and **b** in the continuum with a low dielectric constant (ε=4)

remains the same for the Hyp·Cyt base pair differing only by the presence of the local minimum $Hyp^{+}\cdot Cyt^{-}$, which is dynamically unstable. We revealed that the $TS_{Hyp\cdot Cyt\leftrightarrow Hyp^{*}\cdot Cyt^{*}}$ is more similar to product Hyp*·Cyt* of the Hyp·Cyt→Hyp*·Cyt* tautomerization reaction.

Characteristic structural feature of the Hyp·Cyt→Hyp*·Cyt* tautomerization is that it noticeably disturbs the Watson-Crick-like geometry of the Hyp·Cyt base pair (see Figs. S5, S6, S15 and S16), namely the DPT reaction is accompanied by a contraction of the distance between the two bases. Graphs, reflecting the changes in glycosidic parameters and geometric characteristics of the H-bonds, especially $d_{AH/HB}$ distances (Fig. 5), vividly demonstrate the "breathing" of the Hyp·Cyt base pair throughout the tautomerization process. The Hyp·Cyt base pair deforms in the range 52.4÷56.7/49.1÷57.5° and 10.025÷10.178/9.856÷10.359 Å (Figs. S15 and S16) of the glycosidic angles and distances, correspondingly.

It is connected with the change of distances between pairs of electronegative atoms: N3 and N1, N4 and O6, C2 and O2 (Figs. S5 and S11). It should be noted that in vacuum and in the continuum with ε=4 the C2…O2 and N3…N1 distances monotonically decrease reaching the minimum at key point 2, that coincides with the minimum of the R(H-H) glycosidic distance in the continuum with ε=4 and then non-monotonically increase. At the same time the N4…O6 distance non-monotonically decreases reaching its minimum at key point 7 in vacuum and at key point 6 in the continuum with ε=4. The Hyp·Cyt↔Hyp*·Cyt* tautomerization is represented by the non-monotonical changes of α_1 and α_2 glycosidic angles and R(H-H) glycosidic distance (Figs. S15 and S16).

Analysis of the energetic characteristics listed in Fig. 4 allows us to make a definite conclusion about the interdependence (cooperativity) of H-bonds [37, 38], which can be estimated by the $dE^{i}_{HB}/dIRC/dE^{j}_{HB}/dIRC$ formula proposed in our previous work [27], where the ratio of derivatives is

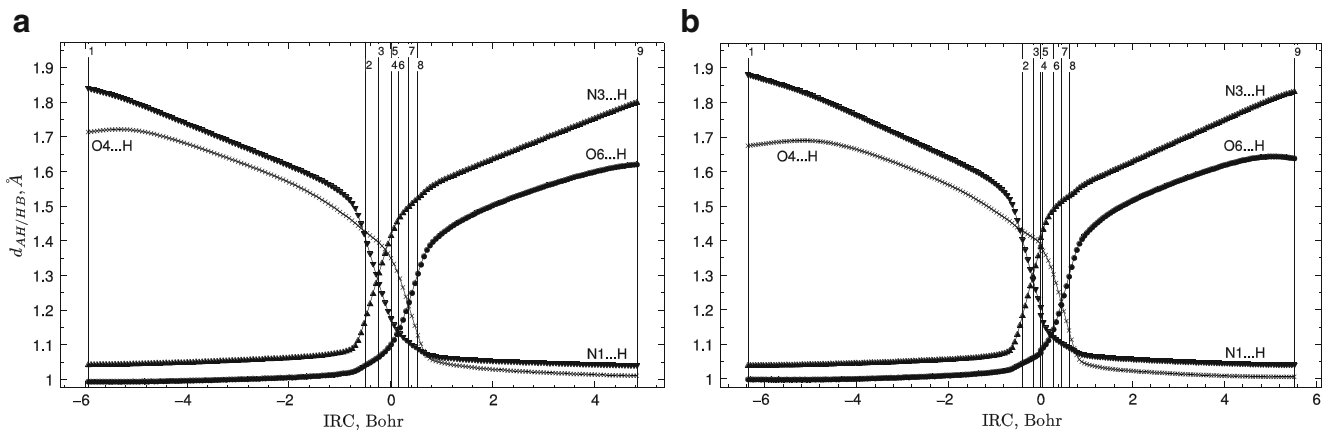

Fig. 10 Profiles of the distance $d_{AH/HB}$ between the hydrogen H and electronegative A or B atoms of the AH…B H-bond along the IRC of the Hyp*·Thy→Hyp·Thy* tautomerization obtained at the DFT level of theory **a** *in vacuo* and **b** in the continuum with a low dielectric constant (ε=4)

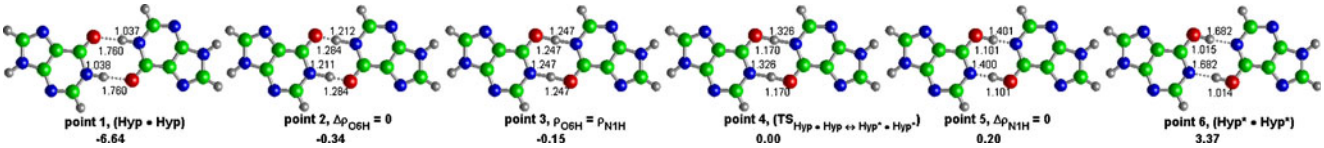

Fig. 11 Geometric structures of the six key points describing the evolution of the Hyp·Hyp→Hyp*·Hyp* tautomerization *via* the DPT along the IRC obtained at the DFT level of theory *in vacuo* (see also Fig. S33 in the Supplementary materials). Coordinates of the six key points are presented in Bohr for each structure. The *dotted lines* indicate H-bonds, while *continuous lines* show covalent bonds (their lengths are presented in angstroms). Carbon atoms are in *green*, nitrogen in *blue*, hydrogen in *grey* and oxygen in *red*

taken at the beginning and at the ending of the movement of protons along the IRC. We obtained the following numerical relations of the H-bonds cooperativity in the Hyp·Cyt and Hyp*·Cyt* base pairs in vacuum using the EML formula [72, 73]: $dE_{N1H\ldots N3}/dE_{N4H\ldots O6}/dE_{C2H\ldots O2}$=1.00/1.45/-0.01 and $dE_{N3H\ldots N1}/dE_{O6H\ldots N4}/dE_{C2H\ldots O2}$=1.00/0.65/0.06, respectively. Thus, unlike the Hyp*·Cyt* base pair, in which all three H-bonds are significantly cooperative and mutually strengthen each other, in the Hyp·Cyt base pair the C2H…O2 H-bond behaves anti-cooperatively, that is it becomes weakened, while two others become strengthened (Fig. 4).

The middle N1H…N3 and the upper N4H…O6 H-bonds in the Hyp·Cyt base pair exist within the 1-2 and 1-6/1-5 (ε=1/ε=4) structures, respectively, coherently becoming stronger during the tautomerization process, the middle N3H…N1 and the upper O6H…N4H-bonds in the Hyp*·Cyt* base pair exist within the 4-9 and 8-9 structures, respectively, coherently becoming weaker during the tautomerization both in vacuum and in the continuum with ε=4 (Fig. 4). The transition from vacuum to the continuum with ε=4 practically has no impact on the energy of H-bonds (Table 1). It should be noted that on the graphs is shown only that energy of H-bonds corresponding to the value $\Delta\rho \geq 0$ (Figs. 3 and 4).

We established that the tautomerization of the Hyp·Cyt base pair is assisted by the third C2H…O2 H-bond, that continuously exists as the H-bond during the entire process of tautomerization, profiles of the C2H…O2 H-bond energy are presented in Figs. S7-S13. The energy value of the C2H…O2 H-bond $E_{C2H\ldots O2}$ maximally increases by 7.1/34.4 % at the IRC=-1.67/-7.59 Bohr in comparison with the starting Hyp·Cyt base pair. $E_{C2H\ldots O2}$ energy decreases to 0.49/0.38 kcalmol^{-1} in the Hyp*·Cyt* mispair. The geometric characteristics of the C2H…O2 H-bond – C2…O2 and H…O2 distances reach their minima along the IRC at key point 2 in the continuum with ε=4 and non-monotonically increase in vacuum (Figs. S11 and S12). The angle of the C2H…O2 H-bond non-monotonically changes along the IRC (Fig. S13).

It draws attention that maximum value of the energy of the C2H…O2 H-bond along the IRC is realized in those key points, where maximum values of the electron density ρ and the Laplacian of the electron density $\Delta\rho$ are implemented (Figs. S7, S8 and S10).

Characteristically, the ellipticity ε for the BCP of the C2H…O2 H-bond reaches its maximum value 0.07/0.26 exactly at the IRC=3.23/-0.52 Bohr (Fig. S9). At the same time the values of the ellipticity ε of the C2H…O2 H-bond possess the same range of values as classical H-bonds (Figs. S4 and S9).

The dipole moment μ shows significant changes in vacuum and noticeably non-monotonic Ω-like dependence in the continuum with ε=4 along the IRC of the Hyp·Cyt→Hyp*·Cyt* tautomerization *via* the DPT (Fig. S14). In the

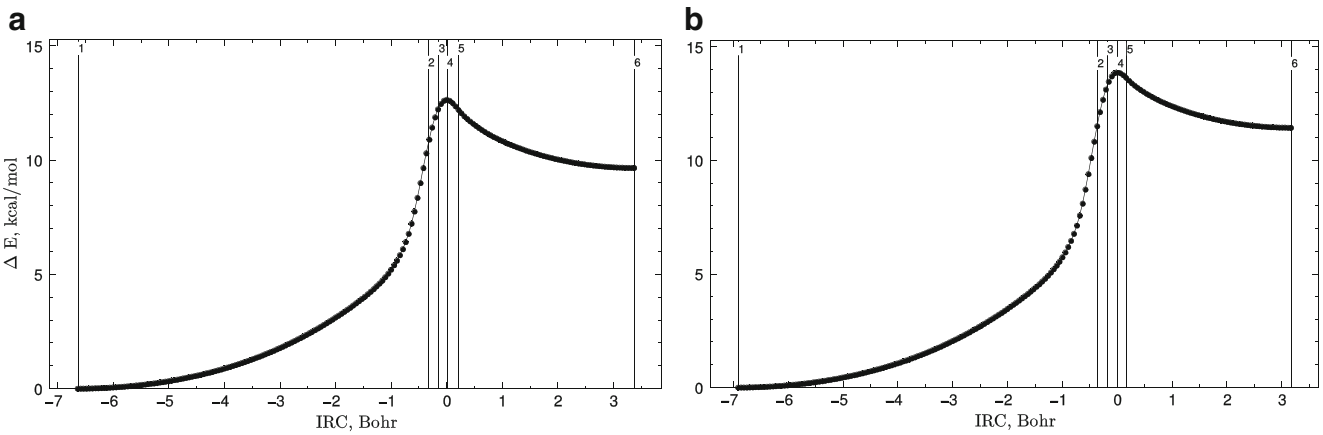

Fig. 12 Energy profiles of the Hyp·Hyp→Hyp*·Hyp* tautomerization along the IRC obtained at the DFT level of theory **a** *in vacuo* and **b** in the continuum with a low dielectric constant (ε=4). Coordinates of the six key points are listed in Figs. 11 and S33

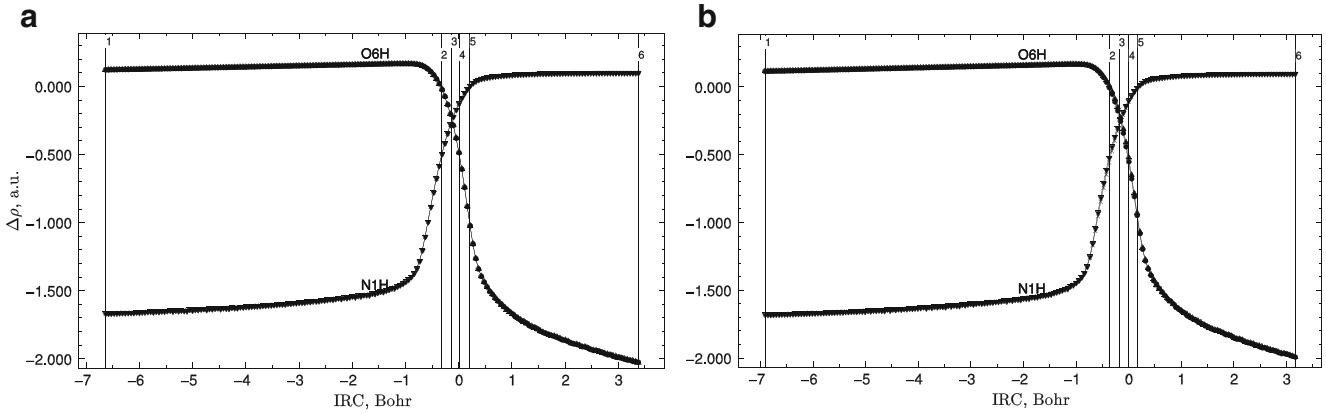

Fig. 13 Profiles of the Laplacian of the electron density $\Delta\rho$ at the BCPs along the IRC of the Hyp·Hyp→Hyp*·Hyp* tautomerization obtained at the DFT level of theory **a** *in vacuo* and **b** in the continuum with a low dielectric constant (ε=4)

IRC range -5.15÷-1.67 Bohr μ almost linearly decreases from 4.07 D (IRC=-5.15 Bohr) to 3.59 D (IRC=-1.67 Bohr) and then sharply increases almost twice to the value 7.99 D at the IRC=-0.30 Bohr in vacuum. Then a sharp decline is observed in vacuum followed by the monotonical decrease to the value 3.90 D at the IRC=3.23 Bohr corresponding to the Hyp*·Cyt* base pair. Along the IRC in the continuum with ε=4 the dipole moment μ monotonically increases from 5.09 D at the IRC=-11.15 Bohr to 5.51 D at the IRC=-7.59 Bohr. Then, a sharp rise (more than twice) of the dipole moment μ from this point to its maximum value 11.90 D at the IRC=-2.34 Bohr is observed. This is followed by the sharp decline of the value of the dipole moment μ to 4.93 D at the IRC=1.82 Bohr corresponding to the Hyp*·Cyt* base pair. This indicates that the tautomerization of the Hyp·Cyt base pair is accompanied by a substantial reorganization of its electronic structure. The profile of the dipole moment also gives a qualitative idea of the charge separation in the course of the Hyp·Cyt→Hyp*·Cyt* tautomerization along the IRC. This observation is interesting for understanding the particular influence of solvation on the tautomerization of the Hyp·Cyt base pair (Fig. S14).

It draws attention to the fact that the symmetry, namely C_s, of the tautomerized base pair remains unchanged during the tautomerization, but at this the Hyp·Cyt base pair significantly deformes, namely it compresses at the $TS_{Hyp\cdot Cyt\leftrightarrow Hyp^*\cdot Cyt^*}$ (see Figs. S15 and S16).

Tautomerization of the Hyp*·Thy mispair into the Hyp·Thy* mispair

The process of the Hyp*·Thy↔Hyp·Thy* tautomerization *via* the DPT along the IRC is also characterized by the nine key points (Figs. 6, 7 and S17), which energetic, geometric, electron-topological and polar characteristics are exhaustively presented in Figs. 6–10 and S17-S32.

Key point 1. The starting structure along the IRC pathway is the Hyp*·Thy base pair with Watson-Crick geometry. It is stabilized by the O6H…O4,

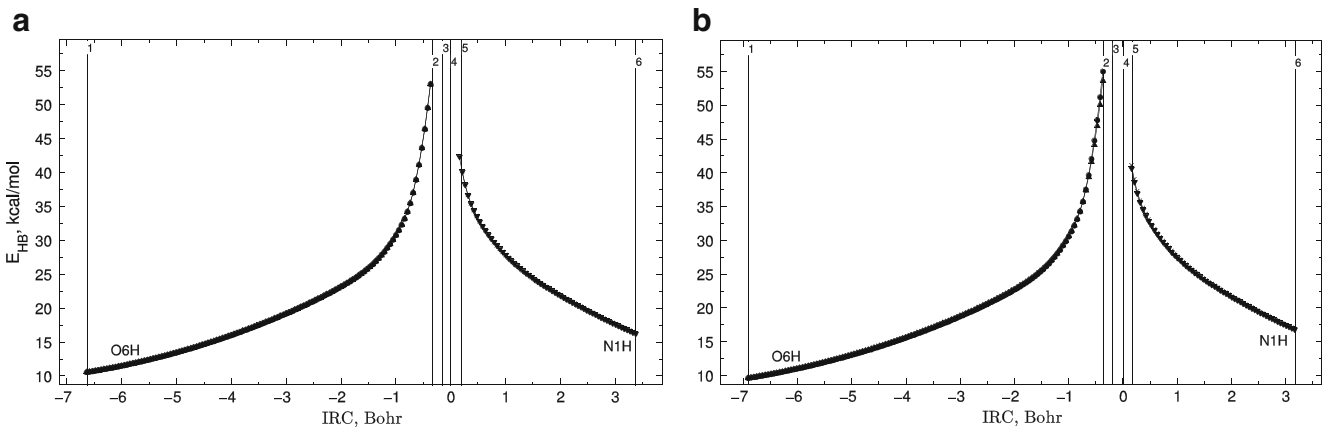

Fig. 14 Profiles of the energy of the H-bonds E_{HB}, estimated by the EML formula [72, 73], along the IRC of the Hyp·Hyp→Hyp*·Hyp* tautomerization obtained at the DFT level of theory **a** *in vacuo* and **b** in the continuum with a low dielectric constant (ε=4)

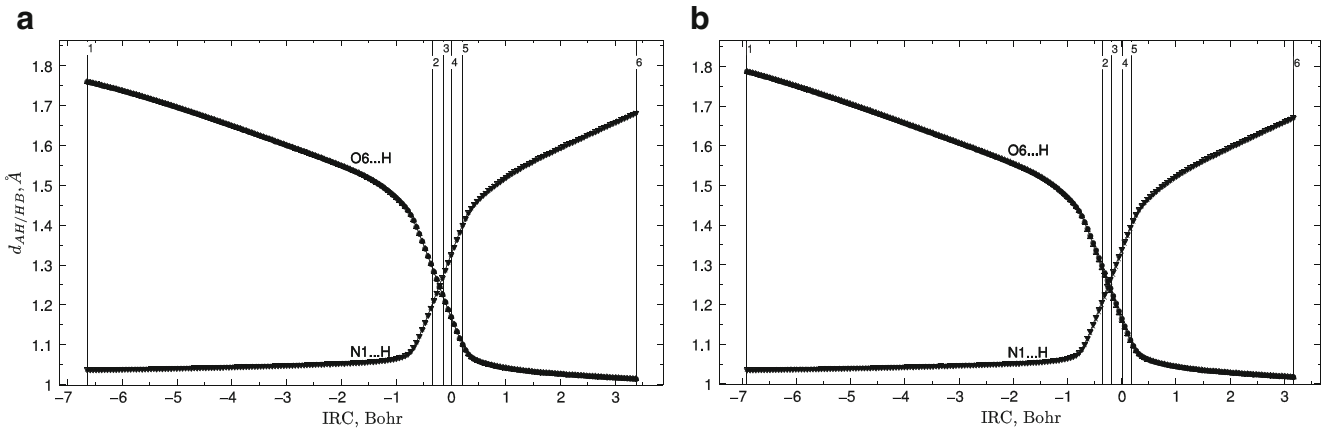

Fig. 15 Profiles of the distance $d_{AH/HB}$ between the hydrogen H and electronegative A or B atoms of the AH…B H-bond along the IRC of the Hyp·Hyp→Hyp*·Hyp* tautomerization obtained at the DFT level of theory **a** *in vacuo* and **b** in the continuum with a low dielectric constant (ε=4)

N3H…N1 and C2H…O2 H-bonds (Table 1, Figs. 6 and S17).

Key point 2. The structure of the base pair corresponding to the situation in which the N3-H chemical bond of the Thy base is significantly weakened and the N1…HH-bond actually becomes the N1-H covalent bond (Figs. 6 and S17). A characteristic feature of this structure is a zero value of the $\Delta\rho$ for the BCP of the N1…HH-bond (Fig. 8). The maximum value of the energy of the N1…HH-bond is attained at this key point (Fig. 9).

Key point 3. This structure is characterized by the equivalent loosened N3-H and N1-H covalent bonds that have equal values of the electron density ρ, the Laplacian of the electron density $\Delta\rho$ at the BCPs and the $d_{N3H/N1H}$ distances of the N3-H and N1-H covalent bonds (Figs. 6, 8, 10, S17 and S19). χ-like dependencies of electron-topological and geometric characteristics are observed for the loosened N3-H-N1 bridge (Figs. 8, 10 and S19).

Key point 4. The $TS_{Hyp^*\cdot Thy \leftrightarrow Hyp\cdot Thy^*}$ of the tautomerization *via* the DPT, stabilized by the O6H…O4 and C2H…O2 H-bonds and N3-H-N1 covalent bridge (Table 1, Figs. 6 and S17).

Key point 5. At this structure situated quite close to the $TS_{Hyp^*\cdot Thy \leftrightarrow Hyp\cdot Thy^*}$ the N3-H covalent bond of the Thy base begins to acquire characteristics of the N3…HH-bond (Figs. 6 and S17). A characteristic feature of this structure is a zero value of the $\Delta\rho$ for the BCP of the N3…HH-bond (Fig. 8). The maximum value of the energy of the N3…HH-bond is attained at this key point (Fig. 9).

Key point 6. The structure corresponding to the situation in which the O6-H chemical bond of the Hyp base is significantly weakened and the O4…HH-bond actually becomes the O4-H covalent bond (Figs. 6 and S17). A characteristic feature of this structure is a zero value of the $\Delta\rho$ for the BCP of the O4…HH-bond (Fig. 8). The maximum value of the energy of the O4…HH-bond is attained at this key point (Fig. 9).

Key point 7. The structure that possesses equivalent loosened O6-H and O4-H covalent bonds that have equal values of the electron density ρ, the Laplacian of the electron density $\Delta\rho$ at the BCPs and the $d_{O4H/O6H}$ distances of the O4-H and O6-H covalent bonds (Figs. 6, 8, 10, S17 and S19). χ-like dependencies of electron-topological and geometric characteristics are observed for the loosened O6-H-O4 bridge (Figs. 8, 10 and S19).

Key point 8. The structure corresponding to the situation in which the mispair containing rare tautomers of the Hyp and Thy bases begins to form, i.e., when the significantly weakened O6-H covalent bond begins to acquire characteristics of the O6…HH-bond (Figs. 6 and S17). A characteristic feature of this structure is a zero value of the $\Delta\rho$ for the BCP of the O6…HH-bond (Fig. 8). The maximum value of the energy of the O6…HH-bond is attained at this key point (Fig. 9).

Key point 9. The final structure is the tautomerized Hyp·Thy* base pair, stabilized by the

O6H…O4, N1H…N3 and C2H…O2 H-bonds (Table 1, Figs. 6 and S17).

These nine key points are used to define the reactant, transition state and product regions [27, 35, 36] of the Hyp*·Thy↔Hyp·Thy* tautomerization *via* the DPT (Fig. 7).

Interestingly, the extrema of the first derivative of the electron energy with respect to the IRC dE/dIRC are reached exactly at key points 2 and 8 (Fig. S18).

We obtained that the electronic energy necessary to bring the donor and acceptor atoms as close as possible to each other to activate the DPT reaction is 5.84/6.18 kcalmol^{-1}. Comparably small amount of energy 4.36/4.87 kcalmol^{-1} releases at the relaxation of key point 8 into key point 9.

Movement along the IRC of the Hyp*·Thy↔Hyp·Thy* tautomerization reaction in vacuum and in the continuum with ε=4 shows that this process is concerted including the concerted movement of two protons along the H-bridges that join the Hyp and Thy bases. Moreover, we established that the DPT in the Hyp*·Thy base pair in vacuum and in the continuum with ε=4 is the *concerted asynchronous* process, in which first the most acidic proton [80, 81], localized at the N3 nitrogen atom of the Thy base, migrates along the strongest N3H…N1 H-bond to the N1 nitrogen atom of the Hyp base inducing the formation of the unstable Hyp^{+}·Thy^{-} zwitterion (Figs. 6 and S17). This migration in its turn provokes the final transfer of the proton at the O6 oxygen atom of the Hyp to the O4 oxygen atom of the Thy base along the O6H…O4 H-bond, exposed in the major groove of double-stranded DNA, that is the limiting stage of this phenomenon, and leads to the formation of the Hyp·Thy* mispair. In the continuum with a low dielectric constant (ε=4) the mechanism of the DPT reaction remains the same for the Hyp*·Thy mispair. We revealed that the $TS_{Hyp^*\cdot Thy \rightarrow Hyp\cdot Thy^*}$ of the Hyp*·Thy→Hyp·Thy* tautomerization reaction is more similar to product – the Hyp·Thy* mispair (Figs. 6 and S17).

The middle N3H…N1 and the upper O6H…O4 H-bonds in the Hyp*·Thy base pair exist within the 1-2 and 1-6 structures, respectively, coherently becoming stronger during the tautomerization process; the middle N1H…N3 and the upper O4H…O6 H-bonds in the Hyp·Thy* base pair exist within the 5-9 and 8-9 structures, respectively, coherently becoming weaker during the tautomerization process in vacuum and in the continuum with ε=4 (Fig. 9). The transition from vacuum to the continuum with ε=4 practically has no impact on the energy of H-bonds (Fig. 9). It should be noted that on the graphs is shown only that energy of H-bonds corresponding to the value $\Delta\rho \geq 0$ (Fig. 8).

Characteristic structural feature of the Hyp*·Thy↔Hyp·Thy* tautomerization process is the slight disturbance of Watson-Crick geometry of the Hyp*·Thy base pair in the course of the tautomerization (Figs. 10, S21, S22, S31 and S32), namely the DPT reaction is accompanied by a contraction of the distance between the Hyp and Thy bases. The DPT reaction insignificantly affects the distance R(H-H) between the H1 and H9 glycosidic protons of the Hyp*·Thy base pair (Figs. S31 and S32), while accompanying by the distorsion of the α_1 and α_2 glycosidic angles that leads to their non-monotonical decreasing with the singularities in the region of the aforementioned key points. Graphs, reflecting the changes in glycosidic parameters and geometric characteristics of the H-bonds, especially $d_{AH/HB}$ distances (Figs. 10, S21, S22, S31 and S32), vividly demonstrate the "breathing" of the Hyp*·Thy base pair throughout the tautomerization process. The Hyp*·Thy base pair deforms in the range 50.4÷52.3/48.9÷52.6° and 9.925÷10.281/9.952÷10.378 Å (Figs. S31 and S32) of the glycosidic distance and angles.

First of all, it is connected with the change of distances between the pairs of electronegative atoms: N3 and N1, O4 and O6, C2 and O2 (Fig. S21). It should be noted that in vacuum and in the continuum with ε=4 the C2…O2 and N3…N1 distances monotonically decrease reaching the minimum at key point 2, coinciding with the minimum of the profile of the R(H-H) glycosidic distance, and then non-monotonically increase (Fig. S31). At the same time the O4…O6 distance non-monotonically decreases reaching its minimum at key point 8 in vacuum and in the continuum with ε=4. This process is represented by the changes of α_1 and α_2 glycosidic angles and R(H-H) glycosidic distance (Fig. S31 and S32).

Analysis of the energetic characteristics listed in Fig. 9 allows us to make a definite conclusion about the cooperativity of H-bonds [37, 38]. We obtained the following numerical relations of the H-bonds cooperativity in the Hyp*·Thy and Hyp·Thy* base pairs in vacuum using the EML formula [72, 73]: $dE_{N3H\ldots N1}/dE_{O6H\ldots O4}/dE_{C2H\ldots O2}$=1.00/-0.67/0.09 and $dE_{N1H\ldots N3}/dE_{O4H\ldots O6}/dE_{C2H\ldots O2}$=1.00/0.28/0.05, respectively. Thus, in contrast to the Hyp·Thy* base pair, in which all three H-bonds are significantly cooperative and mutually reinforce each other, in the Hyp*·Thy base pair O6H…O4 H-bond behaves anti-cooperatively, that is it becomes weakened, while two others become strengthened.

We established that the Hyp*·Thy→Hyp·Thy* tautomerization is assisted by the third C2H…O2 H-bond, that in contrast to the two others H-bonds continuously exists as the H-bond during the entire process of tautomerization and the profiles of which is presented in the Figs. S23-S29. The energy value of the C2H…O2 H-bond $E_{C2H\ldots O2}$ maximally increases by 55.4/64.2 % at the IRC=-0.41/-0.40 Bohr in comparison with the starting Hyp*·Thy base pair (Fig. S26). $E_{C2H\ldots O2}$ decreases to 0.56/0.49 kcalmol^{-1} in the final Hyp·Thy* mispair. The geometric characteristics of the C2H…O2

H-bond – C2…O2 and H…O2 distances reach their minima along the IRC at key point 2 (Figs. S27 and S28). The angle of the C2H…O2 H-bond non-monotonically decreases along the IRC (Fig. S29).

It also draws attention to the fact that maximum value of the energy of the C2H…O2 H-bond along IRC is realized in the vicinity of those points where maximum values of the ρ and the $\Delta\rho$ of the C2H…O2 H-bond are implemented (Figs. S23, S24 and S26).

Characteristically, the ellipticity ε for the C2H…O2 H-bond reaches its maximum value 0.19/0.22 exactly at the Hyp·Thy* mispair (Fig. S25). At the same time the values of the ellipticity ε for the C2H…O2 H-bond possess the same range as the classical H-bonds and lie within the same range (Figs. S25 and S20).

The dipole moment μ initially starts from the value 3.21/4.10 D at the IRC=-5.94/-6.37 Bohr for the Hyp* Thy base pair remaining almost constant value and then rapidly declines to the value 1.61/2.25 D at the IRC=-0.12/-0.06 Bohr (Fig. S30). Then, the slight rise to 1.67/2.41 D at the IRC=0.06/0.11 Bohr is followed by the slight slope to 1.63/2.30 D at the IRC=0.23/0.39 Bohr and then is accompanied by a rapid growth to the value 3.12/4.01 D at the IRC=3.58/3.55 Bohr. After this the value of the dipole moment insignificantly decreases to the 3.01/3.67 D at the IRC=4.82/5.52 Bohr corresponding to the Hyp·Thy* base pair.

It draws attention to the fact that the symmetry of the Hyp·Thy* base pair remains unchanged, namely C_s, but at this the base pair itself slightly deformes, namely it compresses at the $TS_{Hyp^*\cdot Thy \leftrightarrow Hyp\cdot Thy^*}$ (Figs. S31 and S32).

Tautomerization of the Hyp·Hyp homodimer into the Hyp*·Hyp* homodimer

In our opinion, the results obtained by us at the study of the physico-chemical mechanism of the centrosymmetric (C_{2h} symmetry) and therefore non-polar Hyp·Hyp homodimer tautomerization through the DPT are extremely interesting. The Hyp·Hyp homodimer associated by the two equivalent antiparallell N1H…O6 H-bonds with the energy 6.52/6.31 kcalmol^{-1} each (Table 1, Figs. 11–15 and S33-S39). The high-energy Hyp*·Hyp* homodimer (ΔG=5.30/6.83 kcalmol^{-1}, ΔE=5.64/7.04 kcalmol^{-1}), stabilized by the two equivalent antiparallel O6H…N1H-bonds with the energy 9.56/9.97 kcalmol^{-1} each, possesses the same C_{2h} symmetry and zero value of the dipole moment μ as the Hyp·Hyp homodimer. This is not surprising, because there are no fundamental physico-chemical grounds for the symmetry breaking during the Hyp·Hyp→Hyp*·Hyp* tautomerization process under the study.

So, we have first shown that the Hyp·Hyp→Hyp*·Hyp* tautomerization process does not change the symmetry of the Hyp·Hyp dimer. At the movement along the IRC the dipole moment μ remains equal to zero and the corresponding pairs of characteristics listed in Figs. 13–15 and S35 completely coincide. χ-like dependencies, which we discussed earlier, in this case coincide, that is they realize at the same point (IRC=0.16/0.16 Bohr). This clearly indicates that the tautomerization process through the DPT is *synchronous*, i.e., both protons move simultaneously toward each other without the time gap.

Another characteristic difference of the tautomerization of the Hyp Hyp complex in contrast to the previous two (Hyp·Cyt and Hyp*·Thy) is the six key points that characterize the Hyp·Hyp→Hyp*·Hyp* tautomerization instead of the nine key points (Figs. 11, 12 and S33). Degeneration of some points takes place as a consequence of the C_{2h} symmetry of the system.

It should be noted that, the degree of degeneracy of the key points can be larger and the number of points can be reduced to four in those cases, where the initial and final (tautomerized) structures of high-symmetry H-bonded complexes coincide.

Key point 1. The starting structure along the IRC pathway is the Hyp·Hyp homodimer, stabilized by the two equivalent antiparallel N1H…O6 H-bonds (Table 1, Figs. 11 and S33).

Key point 2. The structure of the base pair corresponding to the situation in which the two N1-H chemical bonds of the Hyp bases are significantly weakened and the O6…HH-bonds actually become the O6-H covalent bonds (Figs. 11 and S33). A characteristic feature of this structure is a zero value of the $\Delta\rho$ for the BCPs of the O6…HH-bonds (Fig. 13). The maximum values of the energy of the O6…HH-bonds are realized at this key point (Fig. 14).

Key point 3. This structure is characterized by the equivalent loosened N1-H and O6-H covalent bonds that have equal values of the electron density ρ, the Laplacian of the electron density $\Delta\rho$ at the BCPs and the $d_{N1H/O6H}$ distances of the N1-H and O6-H covalent bonds (Figs. 11, 13, 15, S33 and S35). χ-like dependencies of electron-topological and geometric characteristics are observed for the loosened N3-H-N1 bridges (Figs. 13, 15 and S35).

Key point 4. The $TS_{Hyp\cdot Hyp \leftrightarrow Hyp^*\cdot Hyp^*}$ of the Hyp·Hyp ↔Hyp*·Hyp* tautomerization *via* the DPT stabilized by the pair of the O6-H and N1-H non-equivalent covalent bonds (Table 1, Figs. 11 and S33).

Key point 5. The structure corresponding to the situation in which the mispair containing rare

tautomer of the Hyp base begins to form, i.e., when the significantly weakened N1-H covalent bonds begin to acquire characteristics of the N1…HH-bonds (Figs. 11 and S33). A characteristic feature of this structure is a zero value of the $\Delta\rho$ for the BCPs of the N1…HH-bonds (Fig. 13). The maximum values of the energies of the N1…HH-bonds are realized at this key point (Fig. 14).

Key point 6. The final structure – the tautomerized Hyp*·Hyp* homodimer, stabilized by the two equivalent antiparallel O6H…N1H-bonds (Table 1, Figs. 11 and S33).

These six key points are used to define the reactant, transition state and product regions [27, 35, 36] of the DPT in the Hyp·Hyp base pair (Fig. 12).

Interestingly, the extrema of the first derivative of the electron energy with respect to the IRC dE/dIRC are reached exactly at key points 2 and 5 (Fig. S34).

It quite logically follows from our data (sweeps) on the Laplacian of the electron density $\Delta\rho$ that *the reactant region* begins at key point 1 and ends at key point 2, after passing which Hyp bases lose their chemical individuality and the DPT chemical reaction starts (Fig. 13). We established that the electronic energy necessary to bring the donor and acceptor atoms as close as possible to each other to activate the DPT reaction is 10.90/11.51 kcalmol^{-1}. Actually, *the transition state region*, where the DPT occurs, is located between key points 2 and 5.

So, *the product region*, where the rare tautomer of the Hyp base does not lose its chemical individuality and where the relaxation to the final Hyp*·Hyp* homodimer takes place, begins at key point 5 and ends at key point 6. A comparably small amount of energy 2.54/2.20 kcalmol^{-1}, that is the difference between the energy of key points 5 and 6, releases at the relaxation of the base pair, corresponding to key point 5, under its tautomerization into the Hyp*·Hyp* reaction product, corresponding to key point 6.

Another characteristic feature of this process is a rather narrow zone of the essential chemical reaction, that lies within the range -0.32÷-0.21/-0.37÷0.16 Bohr (Fig. 12). At the same time there is a rather distended process of the mutual deformation and orientation of the Hyp bases until they lose their chemical individuality, energy of which for the Hyp·Hyp→Hyp*·Hyp* tautomerization reaction consists of 86.5/83.2 % of the $TS_{Hyp\cdot Hyp\leftrightarrow Hyp^*\cdot Hyp^*}$ energy!

The N1H…O6 H-bonds in the Hyp·Hyp homodimer exist within the 1 and 2 structures, coherently becoming stronger during the tautomerization process, and the O6H…N1H-bonds in the Hyp*·Hyp* homodimer exist within the 5 and 6 structures, coherently becoming weaker during the tautomerization process, in vacuum and in the continuum with a low dielectric constant (ε=4) (Fig. 14).

Finally, let's discuss the cooperativity [37, 38] of the antiparallel H-bonds in the Hyp·Hyp and Hyp*·Hyp* homodimers. As is clearly seen from Fig. 14, they reinforce each other and in the Hyp*·Hyp* dimer this process is stronger, since in this case the first derivative dE_{HB}/dIRC is ~three times greater than in the Hyp·Hyp dimer.

It should be mentioned, that Hyp monomers, which lost their chemical individuality at the region of the chemical reaction, are joined by the covalent bonds instead of the H-bonds (Figs. 11 and S33). Exactly this fact explains the minimization of the d_{AB} distances at the region of the chemical reaction (Fig. S37). Invariability of the dipole moment μ (namely zero value) during the course of the Hyp·Hyp homodimer tautomerization does not mean at all that this process is not accompanied by the electronic reorganization. This process has a rather high symmetry, which is confirmed by the dependence of the Mulliken charges (Fig. S39): the antibate dependence of the charge on the hydrogen atom H and on the electronegative atoms along the IRC takes place. They change most strongly at the region of the chemical reaction and it is quite natural.

Conclusions

The calculations at the DFT B3LYP/6-311++G(d,p) level of QM theory in combination with the quantum theory "atoms in molecules" (QTAIM) in the isolated state and in the continuum with a low dielectric constant (ε=4), typical for hydrophobic interfaces of specific protein-nucleic acid interactions [25, 27–34], were performed to study in detail a biologically important process of the tautomerization of the Hyp·Cyt, Hyp*·Thy and Hyp·Hyp base pairs to the Hyp*·Cyt*, Hyp·Thy* and Hyp*·Hyp* base pairs *via* the DPT.

To gain a deeper insight into the physico-chemical "anatomy" of the tautomerization of the base pairs, the calculations of the sweeps of the electron-topological, energetic, geometric and polar parameters, that describe the course of the tautomerization along the IRC, have been performed and analyzed.

The nine key points along the IRC of the Hyp·Cyt↔Hyp*·Cyt* and Hyp*·Thy↔Hyp·Thy* tautomerizations and the six key points of the Hyp·Hyp↔Hyp*·Hyp* tautomerization have been identified and fully characterized. These key points could be considered as electron-topological "fingerprints" of concerted asynchronous (for Hyp·Cyt and Hyp*·Thy) or synchronous (for Hyp·Hyp) tautomerization process *via* the DPT.

Based on the sweeps of the energetic, electron-topological, geometric and polar parameters, which describe

the course of the tautomerization along the IRC, it was proved that the tautomerization through the DPT is concerted asynchronous process for the Hyp·Cyt and Hyp*·Thy base pairs, while concerted synchronous process for the Hyp·Hyp homodimer.

It was demonstrated for the first time that in the Hyp*·Cyt*, Hyp·Thy*, Hyp·Hyp and Hyp*·Hyp* base pairs all H-bonds are significantly cooperative and mutually reinforce each other, while the C2H…O2 H-bond in the Hyp·Cyt base pair and the O6H…O4 H-bond in Hyp*·Thy base pair behave anti-cooperatively, i.e., they become weakened, while two others become strengthened.

It was established that the tautomerization process in the Hyp·Cyt and Hyp·Thy* base pairs is accompanied by the electronic reorderings and structural deviations of the geometries of the base pairs from the Watson-Crick geometry. Structural deviations arise primarily due to a decrease in bonding distances between electronegative atoms of H-bridges in the course of the tautomerization process.

H-bond energies in the investigated nucleobase pairs containing Hyp were estimated by QM calculations in vacuum and in the continuum with $\varepsilon=4$. Within the framework of the CPCM model, it was demonstrated that energies of the intrapair H-bonds are slightly disturbed under the transition from vacuum ($\varepsilon=1$) to the continuum with a low dielectric constant ($\varepsilon=4$).

The obtained results enable us to regard the proposed approach [27] not only as a powerful tool for the study of the mechanisms underlying the tautomerization of any H-bonded complex *via* the DPT, but also as a method for researching the cooperativity of the H-bonds [37, 38] that stabilize them.

Acknowledgments This work was partly supported by the State Fund for Fundamental Research of Ukraine within the Ukrainian-Russian (project № F40.4/039) and Ukrainian-Slovenian research bilateral projects for 2011-2012 years. O.O.B. was supported by the Grant of the President of Ukraine to support scientific research of young scientists for 2012 year from the State Fund for Fundamental Research of Ukraine (project № GP/F44/086) and by the Grant of the President of Ukraine for talented youth for 2012 year from the Ministry of Education and Science, Youth and Sports of Ukraine. Authors thank Bogolyubov Institute for Theoretical Physics of the National Academy of Sciences of Ukraine for providing calculation resources and software. This work was performed using computational facilities of joint computer cluster of SSI "Institute for Single Crystals" and Institute for Scintillation Materials of National Academy of Sciences of Ukraine incorporated into Ukrainian National Grid.

References

1. Friedberg EC, Walker GC, Siede W, Wood RD, Schultz RA, Ellenberger T (2006) DNA repair and mutagenesis. ASM, Washington DC
2. Auerbach C (1976) Mutation research: Problems, results, and perspectives. Wiley, New York
3. Dellarco VL, Erickson RP, Lewis SE, Shelby MD (1995) Mutagenesis and human genetic disease: an introduction. Environ Mol Mutagen 25:2–6
4. Brovarets' OO, Kolomiets' IM, Hovorun DM (2012) Elementary molecular mechanisms of the spontaneous point mutations in DNA: a novel quantum-chemical insight into the classical understanding. In: Tada T (ed) Quantum chemistry – molecules for innovations. Rijeka, In Tech Open Access, pp 59–102
5. Inge-Vechtomov SG (1989) Genetics and fundamentals of selection (Genetika s osnovami selektsii) (in Russian). Moscow, Vysshaya Shkola
6. Crick FHC (1966) Codon-anticodon pairing: the wobble hypothesis. J Mol Biol 19:548–555
7. Karran P, Lindahl T (1980) Hypoxanthine in deoxyribonucleic acid: generation by heat-induced hydrolysis of adenine residues and release in free form by a deoxyribonucleic acid glycosylase from calf thymus. Biochemistry 19:6005–6011
8. Lindahl T (1993) Instability and decay of the primary structure of DNA. Nature 362:709–715
9. Miao F, Bouziane M, O'Connor TR (1998) Interaction of the recombinant human methylpurine-DNA glycosylase (MPG protein) with oligodeoxyribonucleotides containing either hypoxanthine or abasic sites. Nucleic Acids Res 26:4034–4041
10. Voet D, Voet JG (1995) Biochemistry. Wiley, New York
11. Munns ARI, Tollin P (1970) The crystal and molecular structure of inosine. Acta Crystallogr B26:1101–1113
12. Schmalle HW, Hanggi G, Dubler E (1988) Structure of hypoxanthine. Acta Crystallogr C44:732–736
13. Kamiya H, Miura H, Kato H, Nishimura S, Ohtsuka E (1992) Induction of mutation of a synthetic c-Ha-ras gene containing hypoxanthine. Cancer Res 52:1836–1839
14. Hang B, Singer B, Margison GP, Elder RH (1997) Targeted deletion of alkylpurine-DNA-N-glycosylase in mice eliminates repair of 1, N6-ethenoadenine and hypoxanthine but not of 3, N4-ethenocytosine or 8-oxoguanine. Proc Natl Acad Sci USA 94:12869–12874
15. Hill-Perkins M, Jones MD, Karran P (1986) Site-specific mutagenesis *in vivo* by single methylated or deaminated purine bases. Mutat Res 162:153–163
16. Lesley RR, Craig AW, Stacey DW (2007) A computational characterization of the hydrogen-bonding and stacking interactions of hypoxanthine. Phys Chem Chem Phys 9:497–509
17. Rutledge LR, Wetmore SD (2012) A computational proposal for the experimentally observed discriminatory behavior of hypoxanthine, a weak universal nucleobase. Phys Chem Chem Phys 14:2743–2753
18. Saparbaev M, Laval J (1994) Excision of hypoxanthine from DNA containing dIMP residues by the *Escherichia coli*, yeast, rat, and human alkylpurine DNA glycosylases. Proc Natl Acad Sci USA 91:5873–5877
19. Schuster H (1960) The reaction of nitrous acid with deoxyribonucleic acid. Biochem Biophys Res Commun 2:320–323
20. Sun X, Lee JK (2007) Acidity and proton affinity of hypoxanthine in the gas phase versus in solution: intrinsic reactivity and biological implications. J Org Chem 72:6548–6555
21. Sun X, Lee JK (2010) The stability of DNA duplexes containing hypoxanthine (inosine): gas versus solution phase and biological implications. J Org Chem 75:1848–1854
22. Bass BL (2002) RNA editing by adenosine deaminases that act on RNA. Annu Rev Biochem 71:817–846
23. Hartman Z, Henrikson EN, Hartman PE, Cebula TA (1994) Molecular models that may account for nitrous acid mutagenesis in organisms containing double-stranded DNA. Environ Mol Mutagen 24:168–175

24. Wuenschell GE, O'Connor TR, Termini J (2003) Stability, miscoding potential, and repair of 2'-deoxyxanthosine in DNA: implications for nitric oxide-induced mutagenesis. Biochemistry 42:3608–3616
25. Brovarets' OO, Hovorun DM (2012) Prototropic tautomerism and basic molecular principles of hypoxanthine mutagenicity: an exhaustive quantum-chemical analysis. J Biomol Struct Dynam. doi:10.1080/07391102.2012.715041
26. Topal MD, Fresco JR (1976) Base pairing and fidelity in codon-anticodon interaction. Nature 263:289–293
27. Brovarets' OO, Hovorun DM (2012) Can tautomerization 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. doi:10.1080/07391102.2012.755795
28. Brovarets' OO, Yurenko YP, Dubey IY, Hovorun DM (2012) Can DNA-binding proteins of replisome tautomerize nucleotide bases? *Ab initio* model study. J Biomol Struct Dynam 29:1101–1109
29. Dwyer JJ, Gittis AG, Karp DA, Lattman EE, Spencer DS, Stites WE, García-Moreno EB (2000) High apparent dielectric constants in the interior of a protein reflect water penetration. Biophys J 79:1610–1620
30. García-Moreno BE, Dwyer JJ, Gittis AG, Lattman EE, Spencer DS, Stites WE (1997) Experimental measurement of the effective dielectric in the hydrophobic core of a protein. Biophys Chem 64:211–224
31. Bayley ST (1951) The dielectric properties of various solid crystalline proteins, amino acids and peptides. Trans Faraday Soc 47:509–517
32. Dewar MJS, Storch DM (1985) Alternative view of enzyme reactions. Proc Natl Acad Sci USA 82:2225–2229
33. Mertz EL, Krishtalik LI (2000) Low dielectric response in enzyme active site. Proc Natl Acad Sci USA 97:2081–2086
34. Petrushka J, Sowers LC, Goodman M (1986) Comparison of nucleotide interactions in water, proteins, and vacuum: model for DNA polymerase fidelity. Proc Natl Acad Sci USA 83:1559–1562
35. Lamsabhi AM, Mó O, Gutiérrez-Oliva S, Pérez P, Toro-Labbé A, Yáñez M (2009) The mechanism of double proton transfer in dimers of uracil and 2-thiouracil – the reaction force perspective. J Comput Chem 30:389–98
36. Herrera B, Toro-Labbé A (2007) The role of reaction force and chemical potential in characterizing the mechanism of double proton transfer in the adenine−uracil complex. J Phys Chem A 111:5921–5926
37. Kar T, Scheiner S (2004) Comparison of cooperativity in CH…O and OH…O hydrogen bonds. J Phys Chem A 108:9161–9168
38. Ziółkowski M, Grabowski SJ, Leszczynski J (2006) Cooperativity in hydrogen-bonded interactions: *ab initio* and "atoms in molecules" analyses. J Phys Chem A 110:6514–6521
39. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Pople JA et al. (2010) GAUSSIAN 09 (Revision B.01), Gaussian Inc. Wallingford, CT
40. Tirado-Rives J, Jorgensen WL (2008) Performance of B3LYP density functional methods for a large set of organic molecules. J Chem Theory Comput 4:297–306
41. Parr RG, Yang W (1989) Density-functional theory of atoms and molecules. Oxford University Press, Oxford
42. Lee C, Yang W, Parr RG (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys Rev B 37:785–789
43. Frisch MJ, Head-Gordon M, Pople JA (1990) Semi-direct algorithms for the MP2 energy and gradient. Chem Phys Lett 166:281–289
44. Brovarets' OO, Hovorun DM (2010) How stable are the mutagenic tautomers of DNA bases? Biopolym Cell 26:72–76
45. Brovarets' OO, Hovorun DM (2010) Stability of mutagenic tautomers of uracil and its halogen derivatives: the results of quantum-mechanical investigation. Biopolym Cell 26:295–298
46. Brovarets' OO, Hovorun DM (2010) Quantum-chemical investigation of tautomerization ways of Watson-Crick DNA base pair guanine-cytosine. Ukr Biochem J 82:55–60
47. Brovarets' OO, Hovorun DM (2010) Molecular mechanisms of transitions induced by cytosine analogue: comparative quantum-chemical study. Ukr Biochem J 82:51–56
48. Brovarets' OO, Hovorun DM (2010) Quantum-chemical investigation of the elementary molecular mechanisms of pyrimidine-purine transversions. Ukr Biochem J 82:57–67
49. Brovarets' OO, Zhurakivsky RO, Hovorun DM (2010) Is there adequate ionization mechanism of the spontaneous transitions? Quantum-chemical investigation. Biopolym Cell 26:398–405
50. Brovarets' OO, Hovorun DM (2010) By how many characters is the genetic information written in DNA? Repts Natl Acad Sci Ukraine 6:175–179
51. Brovarets' OO, Hovorun DM (2011) IR Vibrational spectra of H-bonded complexes of adenine, 2-aminopurine and 2-aminopurine^{+} with cytosine and thymine: quantum-chemical study. Opt Spectrosc 111:750–757
52. Brovarets' OO, Hovorun DM (2011) Intramolecular tautomerization and the conformational variability of some classical mutagens – cytosine derivatives: quantum chemical study. Biopolym Cell 27:221–230
53. Pelmenschikov A, Hovorun DM, Shishkin OV, Leszczynski J (2000) A density functional theory study of vibrational coupling between ribose and base rings of nucleic acids with ribosyl guanosine as a model system. J Chem Phys 113:5986–5990
54. Shishkin OV, Pelmenschikov A, Hovorun DM, Leszczynski J (2000) Theoretical analysis of low-lying vibrational modes of free canonical 2 -deoxyribonucleosides. Chem Phys 260:317–325
55. Yurenko YP, Zhurakivsky RO, Ghomi M, Samijlenko SP, Hovorun DM (2007) Comprehensive conformational analysis of the nucleoside analogue 2′-β-deoxy-6-azacytidine by DFT and MP2 calculations. J Phys Chem B 111:6263–6271
56. Yurenko YP, Zhurakivsky RO, Ghomi M, Samijlenko SP, Hovorun DM (2008) *Ab initio* comprehensive conformational analysis of 2′-deoxyuridine, the biologically significant DNA minor nucleoside, and reconstruction of its low-temperature matrix infrared spectrum. J Phys Chem B 112:1240–1250
57. Yurenko YP, Zhurakivsky RO, Samijlenko SP, Hovorun DM (2011) Intramolecular CH···O hydrogen bonds in the AI and BI DNA-like conformers of canonical nucleosides and their Watson-Crick pairs. Quantum chemical and AIM analysis. J Biomol Struct Dynam 29:51–65
58. Matta CF (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. J Comput Chem 31:1297–1311
59. Lozynski M, Rusinska-Roszak D, Mack H-G (1998) Hydrogen bonding and density functional calculations: the B3LYP approach as the shortest way to MP2 results. J Phys Chem A 102:2899–2903
60. Yurenko YP, Zhurakivsky RO, Ghomi M, Samijlenko SP, Hovorun DM (2007) How many conformers determine the thymidine low-temperature matrix infrared spectrum? DFT and MP2 quantum chemical study. J Phys Chem B 111:9655–9663
61. Samijlenko SP, Krechkivska OM, Kosach DA, Hovorun DM (2004) Transitions to high tautomeric states can be induced in adenine by interactions with carboxylate and sodium ions: DFT calculation data. J Mol Struct 708:97–104
62. Cossi M, Rega N, Scalmani G, Barone V (2003) Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. J Comput Chem 24:669–681

63. Barone V, Cossi M (1998) Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. J Phys Chem A 102:1995–2001
64. Peng C, Schlegel HB (1993) Combining synchronous transit and quasi-Newton methods to find transition states. Isr J Chem 33:449–454
65. Peng C, Ayala PY, Schlegel HB, Frisch MJ (1996) Using redundant internal coordinates to optimize equilibrium geometries and transition states. J Comput Chem 17:49–56
66. Hratchian HP, Schlegel HB (2004) Accurate reaction paths using a Hessian based predictor-corrector integrator. J Chem Phys 120:9918–9924
67. Hratchian HP, Schlegel HB (2005) Finding minima, transition states, and following reaction pathways on *ab initio* potential energy surfaces. In: Dykstra CE, Frenking G, Kim KS, Scuseria G (eds) Theory and applications of computational chemistry: The first 40 years. Elsevier, Amsterdam, pp 195–249
68. Hratchian HP, Schlegel HB (2005) Using Hessian updating to increase the efficiency of a Hessian based predictor-corrector reaction path following method. J Chem Theory Comput 1:61–69
69. Atkins PW (1998) Physical chemistry. Oxford University Press, Oxford
70. Bader RFW (1990) Atoms in molecules: a quantum theory. Oxford University Press, Oxford
71. Keith TA (2011) AIMAll (Version 11.12.19). Retrieved from http://aim.tkgristmill.com
72. Espinosa E, Molins E, Lecomte C (1998) Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. Chem Phys Lett 285:170–173
73. Mata I, Alkorta I, Espinosa E, Molins E (2011) Relationships between interaction energy, intermolecular distance and electron density properties in hydrogen bonded complexes under external electric fields. Chem Phys Lett 507: 185–189
74. Nikolaienko TY, Bulavin LA, Hovorun DM (2012) Bridging QTAIM with vibrational spectroscopy: the energy of intramolecular hydrogen bonds in DNA-related biomolecules. Phys Chem Chem Phys 14:7441–7447
75. Iogansen AV (1999) Direct proportionality of the hydrogen bonding energy and the intensification of the stretching ν(XH) vibration in infrared spectra. Spectrochim Acta A Mol Biomol Spectrosc 55:1585–1612
76. Saenger W (1984) Principles of nucleic acid structure. Springer, New York
77. Grabowski SJ (2011) Red- and blue-shifted hydrogen bonds: the bent rule from quantum theory of atoms in molecules perspective. J Phys Chem A 115:12789–12799
78. Desiraju GR, Steiner T (2001) The weak hydrogen bond in structural chemistry and biology (International union of crystallography monographs on crystallography). Oxford University Press, New York
79. Hovorun DM, Gorb L, Leszczynski J (1999) From the nonplanarity of the amino group to the structural nonrigidity of the molecule: A post-Hartree–Fock *ab initio* study of 2-aminoimidazole. Int J Quantum Chem 75:245–253
80. Gorb L, Podolyan Y, Dziekonski P, Sokalski WA, Leszczynski J (2004) Double-proton transfer in adenine−thymine and guanine−cytosine base pairs. A post-Hartree-Fock *ab initio* study. J Am Chem Soc 126:10119–10129
81. Hovorun DM, Kondratyuk IV, Zheltovsky NV (1995) Nucleotide bases as CH-acids. Biopolym Cell 11:15–20