Electronic Supplementary Information (ESI) for:

Novel physico-chemical mechanism of the mutagenic tautomerisation of the Watson-Crick-like A·G and C·T DNA base mispairs: a quantum-chemical picture

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

Department of Molecular and Quantum Biophysics, Institute of Molecular Biology and Genetics, National Academy of Sciences of Ukraine, 150 Akademika Zabolotnoho Str., 03680 Kyiv, Ukraine
Department of Molecular Biotechnology and Bioinformatics, Institute of High Technologies, Taras Shevchenko National University of Kyiv, 2-h Akademika Hlushkova Ave., 03022 Kyiv, Ukraine
✉Corresponding author. Email: dhovorun@imbg.org.ua

Table S1. Energetic characteristics of the investigated DNA base mispairs and TSs of their tautomerisation *via* the sequential DPT obtained at the different levels of theory for the geometry calculated at the B3LYP/6-311++G(d,p) level of theory.

Complex	Level of theory							
	MP2/6-311++G(2df,pd)		MP2/6-311++G(3df,2pd)		MP2/cc-pVTZ		MP2/cc-pVQZ	
	ΔG^a	ΔE^b	ΔG^a	ΔE^b	ΔG^a	ΔE^b	ΔG^a	ΔE^b
A·G(WC)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$A{\cdot}G^*_{\downarrow}(w)$	3.27	5.70	3.63	6.06	3.88	6.31	3.76	6.19
$A{\cdot}G^*_{\uparrow}(w)$	6.14	7.34	6.36	7.56	6.34	7.54	6.28	7.48
$A^*{\cdot}G_{\uparrow}(w)$	14.16	13.96	14.11	13.91	14.03	13.83	14.29	14.09
$TS^{A^{+}\cdot G^{-}}_{A\cdot G(WC)\leftrightarrow A\cdot G^*\downarrow(w)}$	16.88	16.94	16.79	16.86	17.73	17.79	17.01	17.07
$A^*{\cdot}G_{\downarrow}(w)$	21.01	22.54	20.73	22.26	21.39	22.92	21.11	22.65
$TS^{A^{+}\cdot G^{-}}_{A\cdot G(WC)\leftrightarrow A^*\cdot G\uparrow(w)}$	24.95	24.05	25.01	24.11	25.46	24.56	25.29	24.39
$TS^{A^{-}\cdot G^{+}}_{A\cdot G(WC)\leftrightarrow A\cdot G^*\uparrow(w)}$	25.48	25.37	25.97	25.86	25.35	25.23	25.79	25.68
$TS^{A^{-}\cdot G^{+}}_{A\cdot G(WC)\leftrightarrow A^*\cdot G\downarrow(w)}$	49.23	49.70	48.94	49.41	49.53	50.00	49.20	49.67
C·T(WC)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$C^*{\cdot}T_{\uparrow}(w)$	0.66	0.64	0.58	0.57	0.42	0.41	0.52	0.50
$C^*{\cdot}T_{\downarrow}(w)$	3.31	3.42	2.97	3.08	3.09	3.20	3.08	3.20
$C{\cdot}T^*_{O2\uparrow}(w)$	6.17	6.50	6.40	6.74	6.31	6.64	6.25	6.58
$TS_{C^*\cdot T\uparrow(w)\leftrightarrow C\cdot T^*O2\uparrow(w)}$	6.33	8.95	6.74	9.36	6.72	9.33	6.67	9.28
C*·T*(WC)	8.97	8.81	9.27	9.10	8.82	8.66	9.13	8.97
$TS_{C\cdot T(WC)\leftrightarrow C^*\cdot T^*(WC)}$	9.08	10.91	9.56	11.39	9.36	11.19	9.53	11.36
$C{\cdot}T^*_{\downarrow}(w)$	11.84	14.34	12.10	14.60	12.06	14.56	12.05	14.56
$C{\cdot}T^*_{\uparrow}(w)$	15.60	16.97	15.94	17.31	15.64	17.02	15.69	17.06
$TS^{C^{+}\cdot T^{-}}_{C^*\cdot T^*(WC)\leftrightarrow C^*\cdot T\uparrow(w)}$	17.35	17.66	17.13	17.44	17.36	17.68	17.03	17.34
$TS^{C^{+}\cdot T^{-}}_{C^*\cdot T^*(WC)\leftrightarrow C\cdot T^*\downarrow(w)}$	26.71	25.39	26.69	25.37	27.15	25.83	26.62	25.31
$TS^{C^{-}\cdot T^{+}}_{C\cdot T(WC)\leftrightarrow C^*\cdot T\downarrow(w)}$	34.41	35.70	34.27	35.57	34.66	35.95	34.39	35.68
$TS^{C^{-}\cdot T^{+}}_{C\cdot T(WC)\leftrightarrow C\cdot T^*\uparrow(w)}$	46.51	47.72	46.88	48.09	46.66	47.87	46.74	47.95

[a]The relative Gibbs free energy of the investigated complexes ($\Delta G_{A\cdot G(WC)/C\cdot T(WC)}$=0.00; T=298.15 K), kcal·mol^{-1}

[b]The relative electronic energy of the investigated complexes ($\Delta E_{A\cdot G(WC)/C\cdot T(WC)}$=0.00), kcal·mol^{-1}

Table S2. Interbase interaction energies (in kcal·mol^{-1}) for the investigated DNA base mispairs and TSs of their conversions *via* the sequential DPT obtained at the MP2/6-311++G(3df,2pd)//B3LYP/6-311++G(d,p) level of theory.

Complex	$-\Delta E_{int}$[a]	ΣE_{HB}[b]	$\Sigma E_{HB}/\lvert\Delta E_{int}\rvert$, %	$-\Delta G_{int}$[c]
A·G(WC)[§]	17.54	12.87	73.4	3.57
A·G*$_{\uparrow}$(w)	10.40	9.72	93.5	-2.34
A·G*$_{\downarrow}$(w)	12.41	9.83	79.2	0.90
A*·G$_{\uparrow}$(w)	15.24	11.00	72.2	2.83
A*·G$_{\downarrow}$(w)	7.28	3.60	49.5	-3.40
$TS^{A-\cdot G+}_{A\cdot G(WC)\leftrightarrow A\cdot G^{*}\uparrow(w)}$	143.62	26.60	18.5	128.29
$TS^{A+\cdot G-}_{A\cdot G(WC)\leftrightarrow A\cdot G^{*}\downarrow(w)}$	117.62	14.63	12.4	105.01
$TS^{A+\cdot G-}_{A\cdot G(WC)\leftrightarrow A^{*}\cdot G\uparrow(w)}$	111.19	14.63	13.2	97.62
$TS^{A-\cdot G+}_{A\cdot G(WC)\leftrightarrow A^{*}\cdot G\downarrow(w)}$	115.33	15.72	13.6	100.59
C·T(WC)[§§]	13.86	11.84	85.4	1.54
C*·T$_{\uparrow}$(w)	14.43	12.58	87.2	2.84
C*·T$_{\downarrow}$(w)	11.34	9.92	87.5	-0.12
C·T*$_{O2\uparrow}$(w)	27.55	17.65	64.1	15.17
C*·T*(WC)	21.01	17.54	83.5	9.16
C·T*$_{\downarrow}$(w)	11.26	9.16	81.4	1.36
C·T*$_{\uparrow}$(w)	8.28	8.32	100.5	-2.74
$TS^{C+\cdot T-}_{C^{*}\cdot T^{*}(WC)\leftrightarrow C^{*}\cdot T\uparrow(w)}$	118.41	13.70	11.6	106.83
$TS^{C+\cdot T-}_{C^{*}\cdot T^{*}(WC)\leftrightarrow C\cdot T^{*}\downarrow(w)}$	114.18	16.58	14.5	100.97
$TS^{C-\cdot T+}_{C\cdot T(WC)\leftrightarrow C^{*}\cdot T\downarrow(w)}$	136.14	17.93	13.2	124.04
$TS^{C-\cdot T+}_{C\cdot T(WC)\leftrightarrow C\cdot T^{*}\uparrow(w)}$	136.11	23.17	17.0	123.93

[a]The BSSE-corrected electronic interaction energy

[b]The total energy of the intermolecular H-bonds (see Table 1)

[c]The BSSE-corrected Gibbs free energy of interaction (T=298.15 K)

[§,§§]Data are taken from the works [Brovarets' O.O., Zhurakivsky R.O., Hovorun D.M. (2014). Is the DPT tautomerisation of the long A·G Watson-Crick DNA base mispair a source of the adenine and guanine mutagenic tautomers? A QM and QTAIM response to the biologically important question. *Journal of Computational Chemistry, 35*, 451-466] and [Brovarets', O.O., & Hovorun, D.M. (2013). Atomistic understanding of the C·T mismatched DNA base pair tautomerization *via* the DPT: QM and QTAIM computational approaches. *Journal of Computational Chemistry, 34*, 2577-2590], respectively.

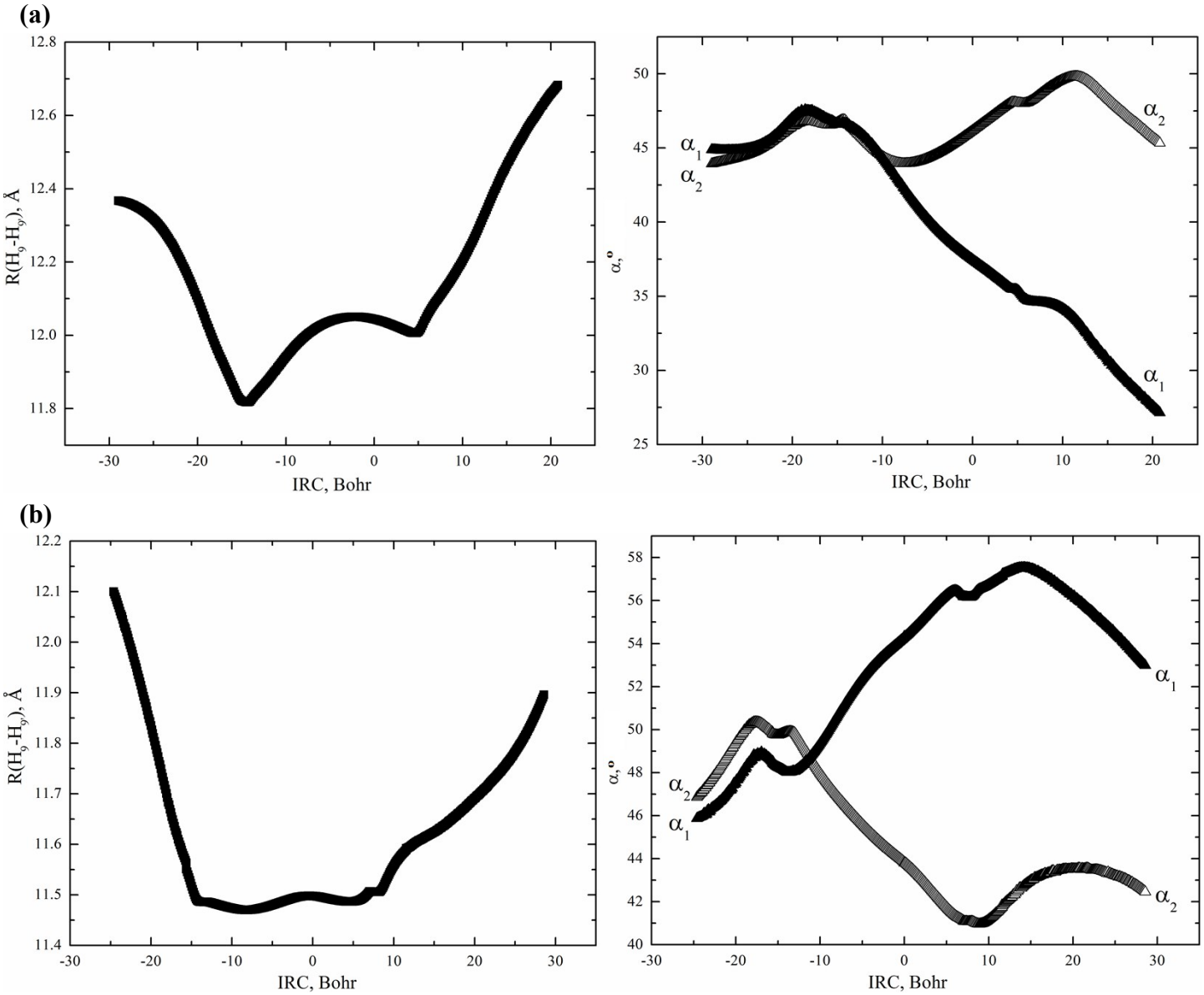
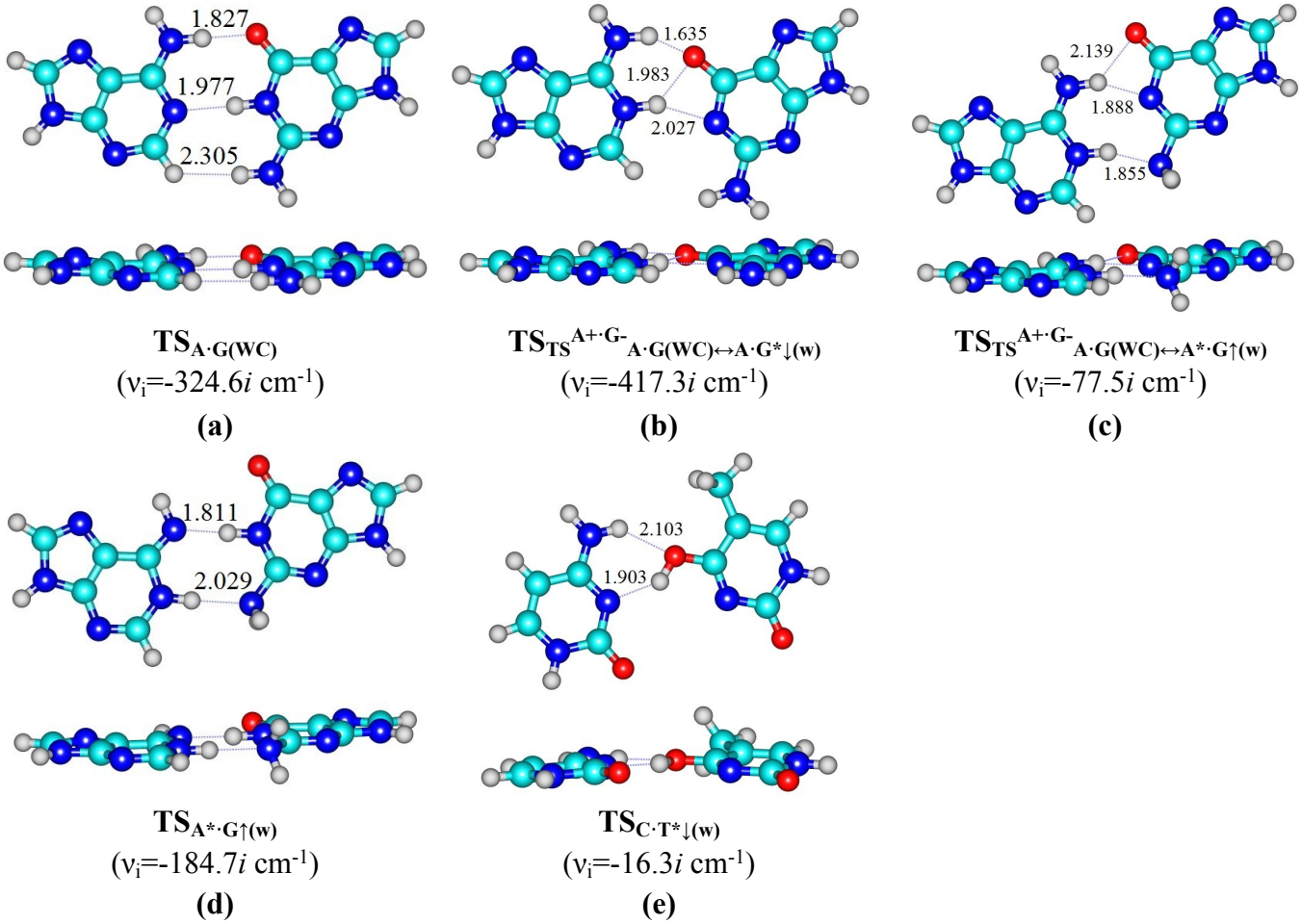


Fig. S1. Profiles of the R(H_9-$H_{9'}$) distances between the H_9 and $H_{9'}$ glycosidic hydrogens and the α_1 (∠N9H_9(G)$H_{9'}$(A)) or α_2 (∠N9$H_{9'}$(A)H_9(G)) glycosidic angles of the G and A DNA bases, respectively, along the IRC of the (**a**) A·G(WC)↔A·G*$_{\downarrow}$(w) and (**b**) A·G(WC)↔A*·G$_{\uparrow}$(w) tautomerisations *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory (see Figs. 1 and 2).



$TS_{A\cdot G(WC)}$ (ν_i=-324.6*i* cm^{-1}) **(a)**

$TS_{TS}^{A+\cdot G-}{}_{A\cdot G(WC)\leftrightarrow A\cdot G^*\downarrow(w)}$ (ν_i=-417.3*i* cm^{-1}) **(b)**

$TS_{TS}^{A+\cdot G-}{}_{A\cdot G(WC)\leftrightarrow A^*\cdot G\uparrow(w)}$ (ν_i=-77.5*i* cm^{-1}) **(c)**

$TS_{A^*\cdot G\uparrow(w)}$ (ν_i=-184.7*i* cm^{-1}) **(d)**

$TS_{C\cdot T^*\downarrow(w)}$ (ν_i=-16.3*i* cm^{-1}) **(e)**

Fig. S2. Transition states of the conformational interconversions of the mirror-symmetric enantiomers of the **(a)** A·G(WC), **(b)** $TS^{A+\cdot G-}{}_{A\cdot G(WC)\leftrightarrow A\cdot G^*\downarrow(w)}$, **(c)** $TS^{A+\cdot G-}{}_{A\cdot G(WC)\leftrightarrow A^*\cdot G\uparrow}(w)$, **(d)** $A^*\cdot G_{\uparrow}(w)$ and **(e)** $TS_{C\cdot T^*\downarrow(w)}$ biologically important complexes. Dotted lines indicate AH···B H-bonds (their lengths are presented in angstroms). Carbon atoms are in light-blue, nitrogen – in dark-blue, hydrogen – in grey and oxygen – in red. ν_i – imaginary frequency.

Table S3. Energetic, vibrational, kinetic and geometric characteristics of the conformational interconversions of the mirror-symmetric enantiomers of the biologically important DNA base mispairs and TSs of their tautomerisation *via* the sequential DPT calculated at the MP2/cc-pVQZ//B3LYP/6-311++G(d,p) level of theory.

Transition state of the interconversion	$\Delta\Delta G_{TS}$[a]	$\Delta\Delta E_{TS}$[b]	$\tau_{99.9\%}$[c]	ν_i[d]	β[e]
$TS_{A\cdot G(WC)}$	1.99	1.56	$1.60\cdot10^{-11}$	-324.6*i*	∠C6N1(G)N1C6(A)=±21.9
$TS_{TS}^{A+\cdot G-}{}_{A\cdot G(WC)\leftrightarrow A\cdot G^*\downarrow(w)}$	2.99	2.81	$8.66\cdot10^{-11}$	-417.3*i*	∠C6N1(G)N1C6(A)=±39.9
$TS_{TS}^{A+\cdot G-}{}_{A\cdot G(WC)\leftrightarrow A^*\cdot G\uparrow(w)}$	2.81	2.34	$6.42\cdot10^{-11}$	-77.5*i*	∠C6N1(G)N1C6(A)=±7.1
$TS_{A^*\cdot G\uparrow(w)}$	7.27	7.16	$1.19\cdot10^{-7}$	-184.7*i*	∠C6N1(G)N1C6(A)=±93.5
$TS_{C\cdot T^*\downarrow(w)}$	4.69	2.33	$1.53\cdot10^{-9}$	-16.3*i*	∠C5N3(T)C2N3(C)=±34.3

[a]The Gibbs free energy barrier of the conformational interconversion under normal conditions, kcal·mol^{-1}

[b]The electronic energy barrier of the conformational interconversion, kcal·mol^{-1}

[c]The time necessary to reach 99.9% of the equilibrium concentration between the mirror-symmetric enantiomers, s

[d]Imaginary frequency in the transition state of the conformational interconversion, cm^{-1}

[e]Dihedral angle that characterize the non-planarity of the complex, degree; signs "±" correspond to the mirror-symmetric enantiomers

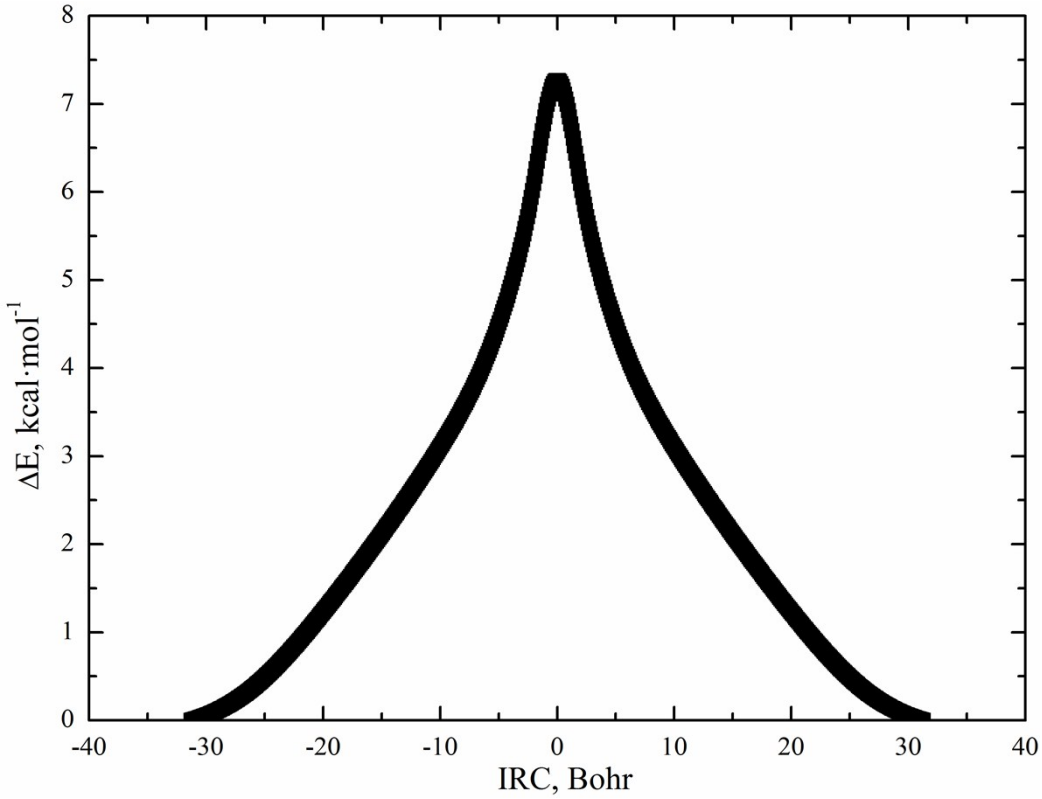
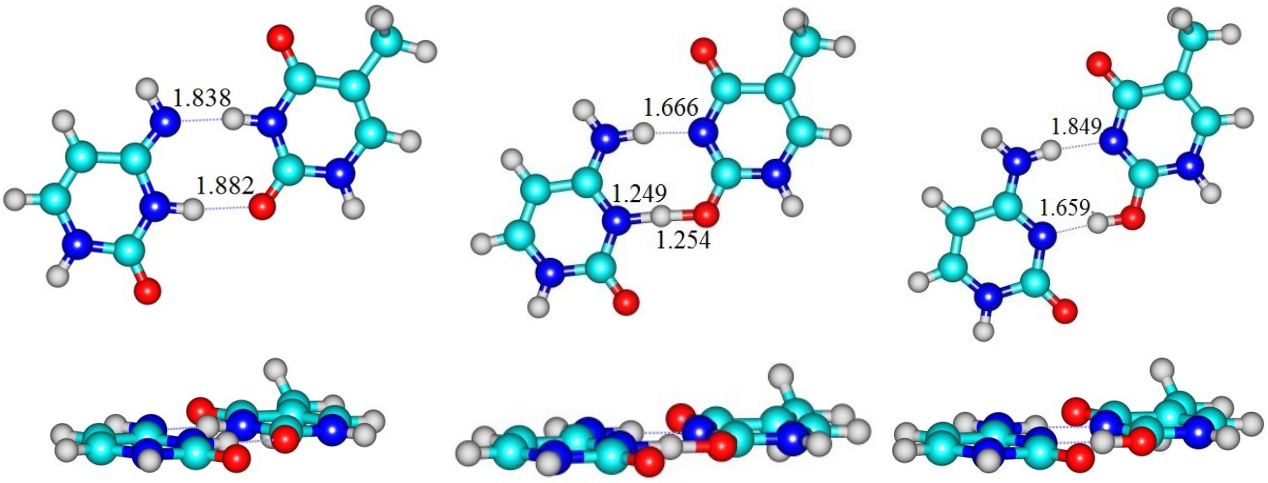


Fig. S3. Profile of the relative electronic energy ΔE of the conformational interconversion of the mirror-symmetric enantiomers of the A*·$G_{↑}$(w) base mispair *via* the rotation of the NH_2 amino group of the G base along the IRC obtained at the B3LYP/6-311++G(d,p) level of theory.



C*·T↑(w) **TS$_{C^{*}\cdot T\uparrow(w)\leftrightarrow C\cdot T^{*}O2\uparrow(w)}$** ($\nu_i$=-807.1*i* cm^{-1}) **C·T*$_{O2\uparrow}$(w)**

Fig. S4. Structures corresponding to the stationary points on the reaction pathway of the C*·T$_{\uparrow}$(w)↔C·T*$_{O2\uparrow}$(w) conversion *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory. Dotted lines indicate AH···B H-bonds and continuous lines – loosened A-H-B covalent bridges (their lengths are presented in angstroms). Carbon atoms are in light-blue, nitrogen – in dark-blue, hydrogen – in grey and oxygen – in red. ν_i – imaginary frequency.

Table S4. Energetic and kinetic characteristics of the C*·T$_{\uparrow}$(w)↔C·T*$_{O2\uparrow}$(w) tautomerisation *via* the sequential DPT obtained at the different levels of theory for the geometry calculated at the B3LYP/6-311++G(d,p) level of theory.

Level of theory	ΔG^a	ΔE^b	$\Delta\Delta G_{TS}^c$	$\Delta\Delta E_{TS}^d$	$\Delta\Delta G^e$	$\Delta\Delta E^f$	$\tau_{99.9\%}^g$
MP2/6-311++G(2df,pd)	5.51	5.86	5.67	8.30	0.16	2.44	$9.21\cdot10^{-13}$
MP2/6-311++G(3df,2pd)	5.83	6.17	6.16	8.79	0.33	2.62	$1.23\cdot10^{-12}$
MP2/cc-pVTZ	5.88	6.23	6.30	8.93	0.41	2.69	$1.41\cdot10^{-12}$
MP2/cc-pVQZ	5.71	6.05	6.12	8.75	0.42	2.70	$1.42\cdot10^{-12}$

[a]The Gibbs free energy of the product relatively the reactant of the tautomerisation reaction ($\Delta G_{C^{*}\cdot T\uparrow(w)}$=0.00 kcal·mol^{-1}; T=298.15 K), kcal·mol^{-1}

[b]The electronic energy of the product relatively the reactant of the tautomerisation reaction ($\Delta E_{C^{*}\cdot T\uparrow(w)}$=0.00 kcal·mol^{-1}), kcal·mol^{-1}

[c]The Gibbs free energy barrier for the forward reaction of tautomerisation, kcal·mol^{-1}

[d]The electronic energy barrier for the forward reaction of tautomerisation, kcal·mol^{-1}

[e]The Gibbs free energy barrier for the reverse reaction of tautomerisation, kcal·mol^{-1}

[f]The electronic energy barrier for the reverse reaction of tautomerisation, kcal·mol^{-1}

[g]The 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 S1 for the Gibbs and electronic energies of the DNA mispairs and TSs relatively the global minimum – the short Watson-Crick-like C·T(WC) DNA base mispair.

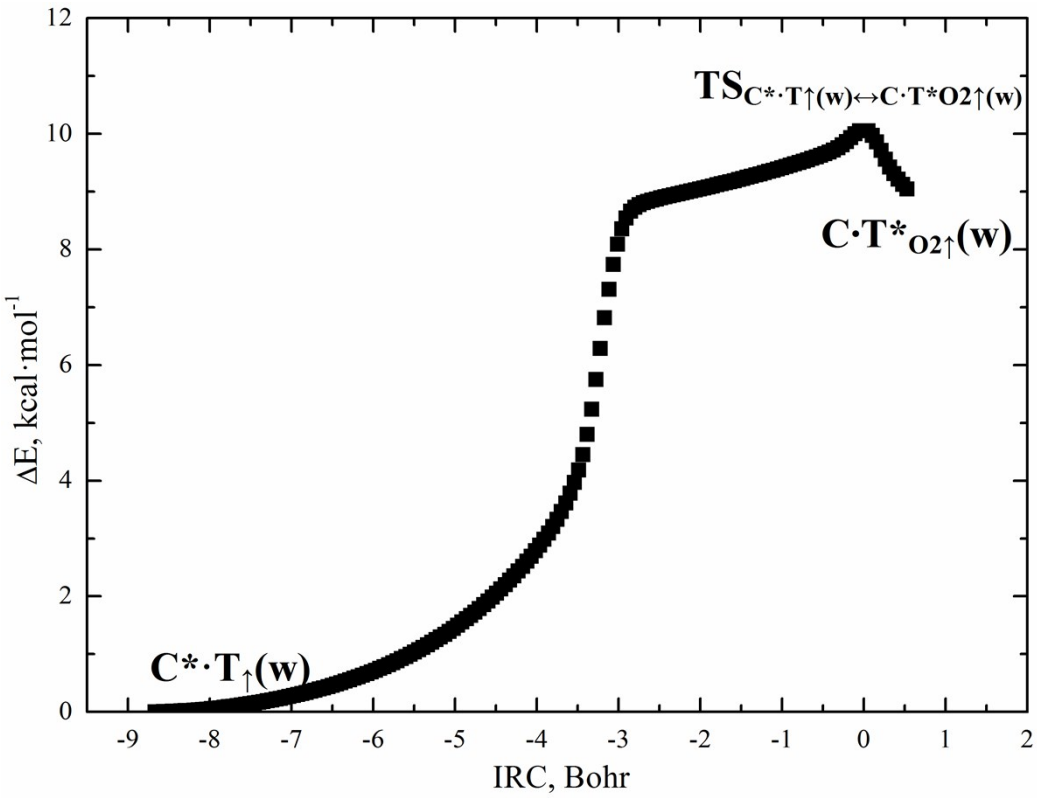


Fig. S5. Profile of the relative electronic energy ΔE of the DNA base mispair along the IRC of the $C^*{\cdot}T_{\uparrow}(w) \leftrightarrow C{\cdot}T^*_{O2\uparrow}(w)$ tautomerisations *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory.

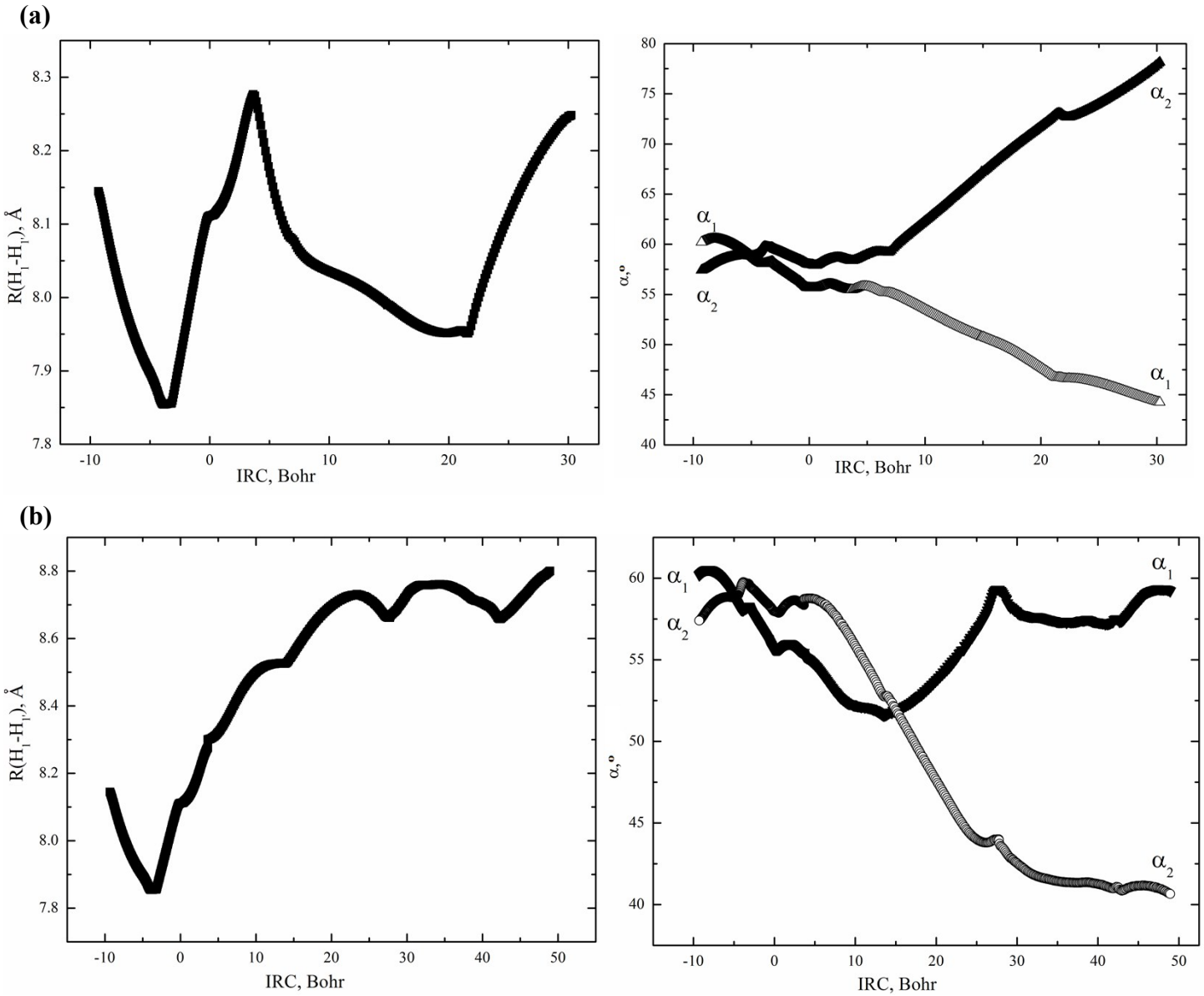


Fig. S6. Profiles of the $R(H_1\text{-}H_{1'})$ distances between the H_1 and $H_{1'}$ glycosidic hydrogens and the α_1 ($\angle N1H_1(C)H_{1'}(T)$) or α_2 ($\angle N1H_{1'}(T)H_1(C)$) glycosidic angles of the C and T DNA bases, respectively, along the IRC of the (**a**) C·T(WC)↔C*·T$_{\uparrow}$(w) and (**b**) C·T(WC)↔C·T*$_{\downarrow}$(w) tautomerisations *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory (see Figs. 4 and 5).