A Quantum-Mechanical Looking Behind the Scene of the Classic G·C Nucleobase Pairs Tautomerization

Ol'ha O. Brovarets'[1]*, Alona Muradova[2] and Dmytro M. Hovorun[1,2]

[1] Department of Molecular and Quantum Biophysics, Institute of Molecular Biology and Genetics, National Academy of Sciences of Ukraine, Kyiv, Ukraine, [2] Department of Molecular Biotechnology and Bioinformatics, Institute of High Technologies, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

For the first time, at the MP2/6-311++G(2df,pd)//B3LYP/6-311++G(d,p) level of theory, a comprehensive quantum-mechanical investigation of the physico-chemical mechanism of the tautomeric wobblization of the four biologically-important G·C nucleobase pairs by the participation of the monomers in rare, in particular mutagenic, tautomeric forms (marked with an asterisk) was provided. These novel tautomeric transformations (wobblization or shifting of the bases within the pair) are intrinsically inherent properties of the G·C nucleobase pairs. In this study, we have obtained intriguing results, lying far beyond the existing representations. Thus, it was shown that Löwdin's G*·C*(WC) base pair does not tautomerize according to the wobblization mechanism. Tautomeric wobblization of the G*·C*(rWC) (relative Gibbs free energy $\Delta G = 0.00$/relative electronic energy $\Delta E = 0.00$ kcal·mol^{-1}) ("r"—means the configuration of the base pair in reverse position; "WC"—the classic Watson-Crick configuration) and G*^{t}·C*(H) ($\Delta G = -0.19/\Delta E = 0.29$ kcal·mol^{-1}) ("H"—Hoogsteen configuration;"t" denotes the O6H hydroxyl group in the *trans* position) base pairs are preceded by the stages of the base pairs tautomerization by the single proton transfer (SPT). It was established that the G*^{t}·C*(rH) ($\Delta G = 2.21/\Delta E = 2.81$ kcal·mol^{-1}) base pair can be wobbled through two different pathways *via* the traditional one-stage mechanism through the TSs, which are tight G$^+$·C$^-$ ion pairs, stabilized by the participation of only two intermolecular H-bonds. It was found out that the G·C base pair is most likely incorporated into the DNA/RNA double helix with parallel strands in the G*·C*(rWC), G·C*(rw$_{wc}$), and G*·C(rw$_{wc}$) ("w"—wobble configuration of the pair) tautomeric forms, which are in rapid tautomeric equilibrium with each other. It was proven that the G*·C*(rWC) nucleobase pair is also in rapid tautomeric equilibrium with the eight tautomeric forms of the so-called Levitt base pair. It was revealed that a few cases of tautomerization *via* the DPT of the nucleobase pairs by the participation of the C8H group of the guanine had occurred. The biological role of the obtained results was also made apparent.

Keywords: DNA, RNA, G●C base pair, tautomerization mechanism, wobblization, wobble base pair, Levitt base pair, Löwdin's base pair

INTRODUCTION

Shortly after the establishment of the spatial organization of the DNA molecule by James Watson and Francis Crick (Watson and Crick, 1953a,b), the tautomeric hypothesis was formulated (Watson and Crick, 1953b; Crick and Watson, 1954), which considers the transformation or transition of the nucleotide bases from the main (canonical) into the rare (mutagenic) tautomeric form as the main source of the origin of spontaneous point mutations. Since that time, the topic of tautomerism has remained active over the decades to the present day (Löwdin, 1963, 1966; Topal and Fresco, 1976; Florian et al., 1994; Gorb et al., 2004; Brovarets' et al., 2014; Godbeer et al., 2015; Turaeva and Brown-Kennerly, 2015).

However, up until recently it was considered that only a few unusual tautomers existed for the G·C Watson-Crick nucleobase pair (Pous et al., 2008; Alvey et al., 2014; Brovarets' and Hovorun, 2014a; Nikolova et al., 2014; Poltev et al., 2016; Szabat and Kierzek, 2017; Brovarets' et al., 2019a,b; Srivastava, 2019). In particular, tautomerization *via* the double proton transfer (DPT) has been carefully investigated in the reverse Löwdin $G^*·C^*(rWC)$, Hoogsteen (H) $G^{*\prime}·C^*(H)$, and reverse Hoogsteen $G^{*\prime}·C^*(rH)$ base pairs (Brovarets' et al., 2019b), leading to the novel structures: $G·C^*_{O2}(rWC)$, $G^*_{N2}·C(rWC)$, $G^{*\prime}_{N2}·C(rWC)$, $G^*_{N7}·C(H)$, and $G^{*\prime}_{N7}·C(rH)$.

Eventually a great contribution into the further development of the tautomeric hypothesis was made by Per-Orlov Löwdin (Löwdin, 1963, 1966) and Topal and Fresco (Topal and Fresco, 1976; Brovarets' et al., 2014). Thus, Per-Orlov Löwdin expressed the revolutionary, non-trivial opinion that the ability of the nucleotide bases to transform into the rare tautomeric form is provided by the electronic structure of the canonical DNA base pairs and qualitatively substantiated this assumption from the position of quantum mechanics. Subsequently, Topal and Fresco elaborated this approach in more detail, by using simple and visual models, and extended it for the explanation of the limited accuracy of codon-anticodon recognition (Topal and Fresco, 1976; Brovarets' et al., 2014).

In recent years, an alternative view in this area of the research was suggested, which could be characterized as the "Renaissance" of the tautomeric hypothesis [see Chapter Brovarets' and Hovorun (2018) and bibliography provided there]. According to this investigation the new, unusual pathways of the tautomeric interconversions between wobble (w) and Watson-Crick (WC) base pairs have been provided (Brovarets' and Hovorun, 2009, 2015a,b,c,d,e,f):

- for usual A·T and G·C DNA base pairs: $A·T(WC)↔A^*·T(w)/A·T^*_{O2}(w)/A·T^*(w)$ and $G·C(WC)↔G·C^*_↑(w)/G^*·C_↓(w)/G·C^*_↓(w)/G^*·C_↑(w)$ (Brovarets' and Hovorun, 2015a);
- for unusual purine-pyrimidine wobble A·C and G·T DNA base pairs: $A·C(w)↔A·C^*(WC)$ and $G·T(w)↔G^*·T(WC)$ (Brovarets' and Hovorun, 2009, 2015b,c);
- for incorrect purine-purine A·A, G·G, and A·G DNA base pairs: $A·A(w)↔A^*·A(WC)$, $G·G(w)↔G^*·G(WC)$,

$A·G(WC)↔A·G^*_↓(w)$, $A·G(WC)↔A^*·G_↑(w)$ (Brovarets' and Hovorun, 2015d,e);
- for pyrimidine-pyrimidine base pairs: $C·T(WC)↔C^*·T_↑(w)$, $C·T(WC)↔C·T^*_↓(w)$, $T·T(w)↔T·T^*(WC)$, $C·C(w)↔C·C^*$ (WC) (Brovarets' and Hovorun, 2015f).

Thus, by utilizing modern quantum-mechanical (QM) methods, the mechanisms of the mutagenic tautomerization of the pairs of nucleotide bases were investigated in detail, which were revealed to be active players in the field of spontaneous point mutagenesis (Brovarets' and Hovorun, 2018). It was established, in which cases Löwdin's approach was adequate and in which cases another approach should be reconsidered and supplemented.

Thus, it was suggested that the mechanism of the mutagenic tautomerization of the DNA base pairs, in particular classic Watson-Crick pairs, are accompanied by the mutual shifting (wobblization) of the bases one relative to the other into the minor or major DNA grooves at the intrapair sequential proton transfer (Brovarets' and Hovorun, 2015a; Brovarets' et al., 2019a). This valuable finding enables researchers to figure out, how the incorrect DNA base pairs, which architecture is different from the Watson-Crick configuration, can acquire the enzymatically-competent conformation, that guarantees their successful chemical incorporation into the composition of the main carrier of the genetic information—DNA—by the high-fidelity DNA-polymerase. Notably, even though these theoretical approaches have been realized in quite basic model objects, they correctly reflect the real state-of-affairs at the macromolecular level, since they have been experimentally confirmed for macromolecular objects.

In this research, the objects of the investigation have been extended—except the Watson-Crick (WC) nucleobase pair, to the other biologically-important G·C nucleobase pairs—reverse Watson-Crick G·C(rWC), Hoogsteen G·C(H), and reverse Hoogsteen G·C(rH). Also, it was exactly established why the classic A·T(WC) DNA base pair was selected for the construction of the genetic material (Brovarets' and Hovorun, 2009, 2015a,b,c,d,e,f; Brovarets' et al., 2018a). The novel mechanism of the mutagenic tautomerization of the biologically-important A·T DNA base pairs through the quasi-orthogonal transition state and also through the protonated amino-group (Brovarets' et al., 2018b,c,d,e,f) was revealed for the first time. Based on these data an assumption was expressed about their possible biological role.

At the same time, investigations into the mechanisms of the mutagenic tautomerization of the pairs of nucleotide bases seemed to be quite a complicated issue, which may not be evident at a first glance. Thus, recent investigations into the tautomerization mechanisms of the biologically-important G·C nucleobase pairs, in which monomers are in the rare, in particular mutagenic, tautomeric form, continue to challenge researchers by its mystery (Brovarets' et al., 2019a,b).

It is still not possible to formulate simple physico-chemical rules, that would predict the course of these biologically important processes. Obviously, this is due to the fact that despite the enormous theoretical and experimental efforts of

researchers, the present material remains insufficient for its final generalization.

This work aims to deepen the existing ideas about the microstructural mechanisms of the tautomerization of the biologically important pairs of nucleobases using the example of the G·C base pair (Brovarets' and Hovorun, 2014a), for which both monomers are in the rare tautomeric form.

Such a task is completely substantiated—we have investigated a few surprising tautomerizations, which significantly expand the existing ideas on tautomerization mechanisms and their biological applications. They will be outlined and discussed in more detail below.

COMPUTATIONAL METHODS

Density Functional Theory Calculations of the Geometry and Vibrational Frequencies

Equilibrium geometries of the investigated nucleobase pairs and the transition states (TSs) of their mutual tautomeric transformations, as well as their harmonic vibrational frequencies have been calculated at the B3LYP/6-311++G(d,p) level of QM theory (Hariharan and Pople, 1973; Krishnan et al., 1980; Lee et al., 1988; Parr and Yang, 1989; Tirado-Rives and Jorgensen, 2008), using the Gaussian'09 program package (Frisch et al., 2010). An applied level of theory has proved itself to be successful for the calculations of similar systems (Brovarets' and Hovorun, 2010a,b, 2015g; Matta, 2010; Brovarets' et al., 2015). A scaling factor that is equal to 0.9668 has been applied in the present work for the correction of the harmonic frequencies of all complexes and TSs of their tautomeric transitions (Palafox, 2014; Brovarets' and Hovorun, 2015g; Brovarets' et al., 2015; El-Sayed et al., 2015). We have confirmed the local minima and TSs, localized by a synchronous transit-guided quasi-Newton method (Peng et al., 1996), on the potential energy landscape by the absence or presence, respectively, of the imaginary frequency in the vibrational spectra of the complexes. We applied standard TS theory for the estimation of the activation barriers of the tautomerization reaction (Atkins, 1998).

All calculations have been carried in the continuum with $\varepsilon = 1$, that adequately reflects the processes occurring in real biological systems without deprivation of the structurally functional properties of the bases in the composition of DNA/RNA and satisfactorily models the substantially hydrophobic recognition pocket of the DNA-polymerase machinery as a part of the replisome (Bayley, 1951; Dewar and Storch, 1985; Petrushka et al., 1986; García-Moreno et al., 1997; Mertz and Krishtalik, 2000; Brovarets' and Hovorun, 2014a,b).

Single Point Energy Calculations

We continued geometry optimizations with electronic energy calculations as single point calculations at the MP2/6-311++G(2df,pd) level of theory (Frisch et al., 1990; Kendall et al., 1992).

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

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

where E_{el}−electronic energy, while E_{corr}−thermal correction.

Evaluation of the Interaction Energies

Electronic interaction energies ΔE_{int} have been calculated at the MP2/6-311++G(2df,pd) level of theory as the difference between the total energy of the base pair and energies of the monomers, which have been corrected for the basis set superposition error (BSSE) (Boys and Bernardi, 1970; Gutowski et al., 1986) through the counterpoise procedure (Sordo et al., 1988; Sordo, 2001).

QTAIM Analysis

Bader's quantum theory of atoms in molecules (QTAIM) (Bader, 1990; Matta and Hernández-Trujillo, 2003; Matta et al., 2006; Cukrowski and Matta, 2010; Keith, 2010; Matta, 2014; Lecomte et al., 2015), using the program package AIMAll (Keith, 2010), was applied to analyze the electron density distribution. The presence of the bond critical point (BCP), namely the so-called (3,-1) BCP, and a bond path between the hydrogen donor and acceptor, as well as the positive value of the Laplacian at this BCP ($\Delta \rho > 0$), were considered as criteria for the H-bond formation (Bader, 1990; Matta and Hernández-Trujillo, 2003; Matta et al., 2006; Cukrowski and Matta, 2010; Matta, 2014; Lecomte et al., 2015). Wave functions were obtained at the level of QM theory used for geometry optimization.

The atomic numbering scheme used for the nucleobases is conventional (Saenger, 1984). In this study mutagenic or rare tautomeric forms of the bases (Brovarets' and Hovorun, 2009, 2014a, 2015a,b,c,d,e,f; Brovarets' and Hovorun, 2018; Brovarets' et al., 2018a,b,c,d,e,f; Brovarets' et al., 2019a,b) are denoted by the asterisk.

OBTAINED RESULTS AND THEIR DISCUSSION

So, based on the obtained data, let us firstly formulate the basic results, which have been obtained for the first time and which have the closest connection to the structural biology and molecular biophysics (**Figures 1**, **2**, **Table 1**).

Before providing the discussion of the investigated material, let us firstly give attention to the novel mechanisms of the G*·C*(rWC) tautomerization, which complement the results of the previous work (Brovarets' et al., 2019a).

1. So, in the G*·C*(rWC) base pair, the non-usual DPT-tautomerization was fixed by the participation of the protons at the N3(C) and N2(G) atoms (**Figure 1**, part I): G*·C*(rWC)↔G*$_{N2}^{*t}$·C(rWC). This process is unusual, since the transfer of the proton from the C* to the G* base along the intermolecular (C)N3H...N1(G) H-bond provokes the rotation of the amino group of the G base into the *trans*-position relative to the C2=N3 double bond. As a result, a significantly non-planar TS$_{G^*·C^*(rWC)↔G_{N2}^{*t}·C(rWC)}$ of the tautomerization reaction is formed, which proceeds through the asynchronous mechanism and the significantly non-planar product of the tautomerization—the G$_{N2}^{*t}$·C(rWC) base pair, which is stabilized by the three intermolecular H-bonds (G)O6H...O2(C),   (G)N1H...N3(C),   and   (C)N4H...N2(G).

I. G*·C*(rWC)↔G*t_N2·C (rWC) {DPT}

G*·C*(rWC)

(ΔE_int=-20.21 / ΔG_int=-7.40 / ΔG=0.00 / ΔE=0.00)

TS_{G*·C*(rWC)↔G*t_N2·C(rWC)}
(ν_i=1243.4 i cm⁻¹)
(ΔG=27.43 / ΔE=29.29)

G*t_N2·C(r_WC)

(ΔE_int=-33.41 / ΔG_int=-18.99 / ΔG=24.74 / ΔE=24.58)

II. G*·C*(rWC)↔G⁺·C⁻(rWC)↔G·C*_O2(rWC)↔G·C*(rw_WC)↔G*·C(rw_WC)

G*·C*(rWC)↔G⁺·C⁻(rWC) {SPT}
(Brovarets' et al., 2019a)

G⁺·C⁻(rWC)↔G·C*_O2(rWC) {SPT}
(Brovarets' et al., 2019a)

G*·C*(rWC)

(ΔE_int=-20.21 / ΔG_int=-7.40 / ΔG=0.00 / ΔE=0.00)

TS_{G*·C*(rWC)↔G⁺·C⁻(rWC)}
(ν_i=1132.1 i cm⁻¹)
(ΔG=4.38 / ΔE=6.72)

G⁺·C⁻(rWC)

(ΔE_int=-161.61 / ΔG_int=-146.76 / ΔG=4.44 / ΔE=5.68)

TS_{G⁺·C⁻(rWC)↔G·C*_O2(rWC)}
(ν_i=383.9 i cm⁻¹)
(ΔG=3.64 / ΔE=5.71)

G·C*_O2(rWC)

(ΔE_int=-39.67 / ΔG_int=-26.53 / ΔG=3.51 / ΔE=3.80)

G·C*_O2(rWC)↔G·C*(rw_WC) {WC↔w}

G·C*(rw_WC)↔G*·C(rw_WC) {DPT}

TS_{G⁺·C⁻ G·C*_O2(rWC)↔G·C*(rw_WC)}
(ν_i=159.8 i cm⁻¹)
ΔE_int=-139.96 / ΔG_int=-125.22 /
ΔG=19.47 / ΔE=20.83)

G·C*(rw_WC)

(ΔE_int=-17.44 / ΔG_int=-5.43 /
ΔG=1.39 / ΔE=2.29)

TS_{G·C*(rw_WC)↔G*·C(rw_WC)}
(ν_i=544.9 i cm⁻¹)
(ΔG=6.42 / ΔE=9.46)

G*·C(rw_WC)

(ΔE_int=-17.45 / ΔG_int=-4.85 /
ΔG=1.40 / ΔE=2.33)

III. G*·C*(rWC)↔G⁺·C⁻(rWC)↔G*_N2·C*(rw_WC)↔G*·C*_O2(rw_WC)

G*·C*(rWC)↔G⁺·C⁻(rWC) {SPT}
(Brovarets' et al., 2019a)

G⁺·C⁻(rWC)↔G*_N2·C*(rw_WC) {WC↔w}

G*·C*(rWC)

(ΔE_int=-20.21 / ΔG_int=-7.40 /
ΔG=0.00 / ΔE=0.00)

TS_{G*·C*(rWC)↔G⁺·C⁻(rWC)}
(ν_i=1132.1 i cm⁻¹)
(ΔG=4.38 / ΔE=6.72)

G⁺·C⁻(rWC)

(ΔE_int=-161.61 / ΔG_int=-146.76 / ΔG=4.44 / ΔE=5.68)

TS_{G⁺·C⁻(rWC)↔G*_N2·C*(rw_WC)}
(ν_i=763.3 i cm⁻¹)
(ΔG=17.94 / ΔE=21.81)

G*_N2·C*(rw_WC)

(ΔE_int=-23.88 / ΔG_int=-11.01 /
ΔG=25.70 / ΔE=26.96)

G*_N2·C*(rw_WC)↔G*·C*_O2(rw_WC) {DPT}

TS_{G*_N2·C*(rw_WC)↔G*·C*_O2(rw_WC)}
(ν_i=1036.8 i cm⁻¹)
(ΔG=24.41 / ΔE=27.80)

G*·C*_O2(rw_WC)

(ΔE_int=-20.69 / ΔG_int=-10.05 /
ΔG=17.31 / ΔE=20.03)

FIGURE 1 | Continued

IV. $G^* \cdot C^*_{O2}(rw_{WC})$ ↔ $G^{*t} \cdot C^*_{O2}(rw_{WC})$

             ↕   ↕

   $G^* \cdot C^{*t}_{O2}(rw_{WC})$ ↔ $G^{*t} \cdot C^{*t}_{O2}(rw_{WC})$

$G^* \cdot C^*_{O2}(rw_{WC})$

(ΔE_{int}=-20.69 / ΔG_{int}=-10.05 /
ΔG=17.31 / ΔE=20.03)

$TS_{G^* \cdot C^*O2(rwWC) \leftrightarrow G^{*t} \cdot C^*O2(rwWC)}$
(ν_i=509.6 i cm^{-1})
(ΔE_{int}=-21.98 / ΔG_{int}=-10.01 /
ΔG=24.53 / ΔE=26.98)

$G^{*t} \cdot C^*_{O2}(rw_{WC})$

(ΔE_{int}=-21.89 / ΔG_{int}=-10.03 /
ΔG=17.16 / ΔE=18.65)

$TS_{G^* \cdot C^*O2(rwWC) \leftrightarrow G^* \cdot C^{*t}O2(rwWC)}$
(ν_i=1048.4 i cm^{-1})
(ΔE_{int}=-19.19 / ΔG_{int}=-8.54 /
ΔG=43.53 / ΔE=48.58)

$TS_{G^{*t} \cdot C^*O2(rwWC) \leftrightarrow G^{*t} \cdot C^{*t}O2(rwWC)}$
(ν_i=1052.1 i cm^{-1})
(ΔE_{int}=-20.66 / ΔG_{int}=-8.66 /
ΔG=43.65 / ΔE=47.35)

$G^* \cdot C^{*t}_{O2}(rw_{WC})$

(ΔE_{int}=-17.32 / ΔG_{int}=-5.79 /
ΔG=18.09 / ΔE=19.76)

$TS_{G^* \cdot C^{*t}O2(rwWC) \leftrightarrow G^{*t} \cdot C^{*t}O2(rwWC)}$
(ν_i=504.9 i cm^{-1})
(ΔE_{int}=-18.60 / ΔG_{int}=-6.64 /
ΔG=24.73 / ΔE=27.03)

$G^{*t} \cdot C^{*t}_{O2}(rw_{WC})$

(ΔE_{int}=-19.19 / ΔG_{int}=-8.54 /
ΔG=17.26 / ΔE=18.61)

V. $G^*_{N2} \cdot C^*(rw_{WC})$ ↔ $G^{*t}_{N2} \cdot C^*(rw_{WC})$

             ↕   ↕

   $G^*_{N2} \cdot C^{*t}(rw_{WC})$ ↔ $G^{*t}_{N2} \cdot C^{*t}(rw_{WC})$

$G^*_{N2} \cdot C^*(rw_{•C})$

(ΔE_{int}=-23.88 / ΔG_{int}=-11.01 /
ΔG=25.70 / ΔE=26.96)

$TS_{G^*N2 \cdot C^*(rwWC) \leftrightarrow G^{*t}N2 \cdot C^*(rwWC)}$
(ν_i=476.4 i cm^{-1})
(ΔE_{int}=-8.88 / ΔG_{int}=2.07 /
ΔG=31.54 / ΔE=33.54)

$G^{*t}_{N2} \cdot C^*(rw_{•C})$

(ΔE_{int}=-18.43 / ΔG_{int}=-6.46 /
ΔG=24.36 / ΔE=25.54)

$TS_{G^*N2 \cdot C^*(rwWC) \leftrightarrow G^*N2 \cdot C^{*t}(rwWC)}$
(ν_i=1016.3 i cm^{-1})
(ΔE_{int}=-24.01 / ΔG_{int}=-10.84 /
ΔG=49.94 / ΔE=53.48)

$TS_{G^{*t}N2 \cdot C^*(rwWC) \leftrightarrow G^{*t}N2 \cdot C^{*t}(rwWC)}$
(ν_i=1024.3 i cm^{-1})
(ΔE_{int}=-18.14 / ΔG_{int}=-5.72 /
ΔG=48.27 / ΔE=52.56)

$G^*_{N2} \cdot C^{*t}(rw_{WC})$

(ΔE_{int}=-26.87 / ΔG_{int}=-13.79 /
ΔG=24.54 / ΔE=25.73)

$TS_{G^*N2 \cdot C^{*t}(rwWC) \leftrightarrow G^{*t}N2 \cdot C^{*t}(rwWC)}$
(ν_i=476.4 i cm^{-1})
(ΔE_{int}=-21.47 / ΔG_{int}=-9.42 /
ΔG=30.80 / ΔE=32.76)

$G^{*t}_{N2} \cdot C^{*t}(rw_{WC})$

(ΔE_{int}=-20.97 / ΔG_{int}=-7.90 /
ΔG=23.56 / ΔE=24.76)

FIGURE 1 | Continued

VI. a) G*·C*$_{O2}$(rw$_{WC}$)↔G*$_{N2}$·C*(rw$_{WC}$) {*DPT*}

G*·C*$_{O2}$(rw$_{WC}$)

(ΔE$_{int}$=-20.69 / ΔG$_{int}$=-10.05 /
ΔG=17.31 / ΔE=20.03)

TS$_{G*·C*O2(rwWC)↔G*N2·C*(rwWC)}$
(ν$_i$=1036.8 *i* cm^{-1})
(ΔG=24.41 / ΔE=27.80)

G*$_{N2}$·C*(rw$_{WC}$)

(ΔE$_{int}$=-23.88 / ΔG$_{int}$=-11.01 /
ΔG=25.70 / ΔE=26.96)

b) G*^{t}·C*$^t_{O2}$ (rw$_{WC}$)↔G*$^t_{N2}$·C*^{t}(rw$_{WC}$) {*DPT*}

G*^{t}·C*$^t_{O2}$ (rw$_{WC}$)

(ΔE$_{int}$=-19.19 / ΔG$_{int}$=-8.54 /
ΔG=17.26 / ΔE=18.61)

TS$_{G*^t·C*O2 (rwWC)↔G*^t_{N2}·C*^t(rwWC)}$
(ν$_i$=995.3 *i* cm^{-1})
(ΔG=24.69 / ΔE=28.58)

G*$^t_{N2}$·C*^{t}(rw$_{WC}$)

(ΔE$_{int}$=-20.97 / ΔG$_{int}$=-7.90 /
ΔG=23.56 / ΔE=24.76)

c) G*^{t}·C*$_{O2}$(rw$_{WC}$)↔G*$^t_{N2}$·C*(rw$_{WC}$) {*DPT*}

G*^{t}·C*$_{O2}$(rw$_{WC}$)

(ΔE$_{int}$=-21.89 / ΔG$_{int}$=-10.03 /
ΔG=17.16 / ΔE=18.65)

TS$_{G*^t·C*O2(rwWC) ↔ G*^t_{N2}·C*(rwWC)}$
(ν$_i$=945.7 *i* cm^{-1})
(ΔG=24.19 / ΔE=27.84)

G*$^t_{N2}$·C*(rw$_{WC}$)

(ΔE$_{int}$=-18.43 / ΔG$_{int}$=-6.46 /
ΔG=24.36 / ΔE=25.54)

d) G*·C*$^t_{O2}$ (rw$_{WC}$)↔G*$_{N2}$·C*^{t}(rw$_{WC}$) {*DPT*}

G*·C*$^t_{O2}$ (rw$_{WC}$)

(ΔE$_{int}$=-17.32 / ΔG$_{int}$=-5.79 /
ΔG=18.09 / ΔE=19.76)

TS$_{G*·C*^t_{O2} (rwWC) ↔ G*N2·C*^t(rwWC)}$
(ν$_i$=1144.3 *i* cm^{-1})
(ΔG=24.26 / ΔE=28.23)

G*$_{N2}$·C*^{t}(rw$_{WC}$)
(ΔE$_{int}$=-26.87 / ΔG$_{int}$=-13.79 /
ΔG=24.54 / ΔE=25.73)

FIGURE 1 | Continued

VII. G*t·C*(H)↔G*N7·C(H)↔G*t·C*O2(wH)↔G*N7·C*(wH)

G*t·C*(H)↔G*N7·C(H) {DPT}

G*t·C*(H)
(ΔEint=-21.24 / ΔGint=-8.91 /
ΔG=0.00 / ΔE=0.00)

TSG*t·C*(H)↔G*N7·C(H)
(νi=675.7 i cm⁻¹)
(ΔG=4.01 / ΔE=6.65)

G*N7·C(H)
(ΔEint=-35.66 / ΔGint=-22.05 /
ΔG=3.20 / ΔE=3.07)

G*N7·C(H)↔G*t·C*O2(wH) {H↔wH}   G*t·C*O2(wH)↔G*N7·C*(wH) {DPT}

TSG+·C⁻G*N7·C(H)↔G*t·C*O2(wH)
(νi=196.2 i cm⁻¹)
(ΔEint=-140.75 / ΔGint=-127.54 /
ΔG=25.24 / ΔE=27.66)

G*t·C*O2(wH)
(ΔEint=-25.11 / ΔGint=-12.97 /
ΔG=17.07 / ΔE=17.80)

TSG*t·C*O2(wH)↔G*N7·C*(wH)
(νi=548.9 i cm⁻¹)
(ΔG=18.88 / ΔE=22.22)

G*N7·C*(wH)
(ΔEint=-21.91 / ΔGint=-10.05 /
ΔG=16.43 / ΔE=17.33)

VIII. G*t·C*O2(wH)↔G*t·C*tO2(wH)↔G*N7·C*t(wH)↔G*N7·C*(wH)

G*t·C*O2(wH)↔G*t·C*tO2(wH) {TRANS}   G*t·C*tO2(wH)↔G*N7·C*t(wH) {DPT}

G*t·C*O2(wH)
(ΔEint=-25.11 / ΔGint=-12.97 /
ΔG=17.07 / ΔE=17.80)

TSG*t·C*O2(wH)↔G*t·C*tO2(wH)
(νi=1039.8 i cm⁻¹)
(ΔEint=-25.87 / ΔGint=-14.56 /
ΔG=41.11 / ΔE=44.34)

G*t·C*tO2(wH)
(ΔEint=-24.84 / ΔGint=-12.80 /
ΔG=13.70 / ΔE=14.37)

TSG*t·C*tO2(wH)↔G*N7·C*t(wH)
(νi=1063.0 i cm⁻¹)
(ΔG=13.63 / ΔE=18.49)

G*N7·C*t(wH)
(ΔEint=-29.40 / ΔGint=-16.73 /
ΔG=12.04 / ΔE=12.25)

G*N7·C*t(wH)↔G*N7·C*(wH) {TRANS}

TSG*N7·C*t(wH)↔G*N7·C*(wH)
(νi=463.0 i cm⁻¹)
(ΔEint=-24.49 / ΔGint=-11.98 /
ΔG=39.13 / ΔE=41.95)

G*N7·C*(wH)
(ΔEint=-21.91 / ΔGint=-10.05 /
ΔG=16.43 / ΔE=17.33)

IX. G*t·C(wH)↔G*·C(wH)↔G*N7·C*(wH)↔G*tN7·C*(wH)

G*t·C(wH)↔G*·C(wH) {TRANS}   G*·C(wH)↔G*tN7·C*(wH) {DPT}

G*t·C(wH)
(ΔEint=-10.75 / ΔGint=0.05 /
ΔG=5.16 / ΔE=7.41)

TSG*t·C(wH)↔G*·C(wH)
(νi=500.7 i cm⁻¹)
(ΔEint=-12.83 / ΔGint=-1.10 /
ΔG=11.43 / ΔE=13.83)

G*·C(wH)
(ΔEint=-13.06 / ΔGint=-1.82 /
ΔG=3.25 / ΔE=5.07)

TSG*·C(wH)↔G*N7·C*(wH)
(νi=1137.9 i cm⁻¹)
(ΔG=30.48 / ΔE=33.75)

G*N7·C*(wH)
(ΔEint=-17.97 / ΔGint=-6.26 /
ΔG=30.33 / ΔE=30.90)

G*N7·C*(wH)↔G*tN7·C*(wH) {TRANS}

TSG*N7·C*(wH)↔G*tN7·C*(wH)
(νi=463.0 i cm⁻¹)
(ΔEint=-18.30 / ΔGint=-6.55 /
ΔG=38.05 / ΔE=39.57)

G*tN7·C*(wH)
(ΔEint=-20.05 / ΔGint=-8.70 /
ΔG=34.30 / ΔE=35.29)

FIGURE 1 | Continued

FIGURE 1 | Investigated pathways of the tautomeric wobblization of the biologically-important G·C nucleobase pairs – G*·C*(rWC), G*·C*(H) and G*·C*(rH) pairs obtained at the MP2/6-311++G(2df,pd)//B3LYP/6-311++G(d,p) level of QM theory. ΔG-relative Gibbs free energy and ΔE-electronic energy (in kcal·mol^{-1}; MP2/6-311++G(2df,pd)//B3LYP/6-311++G(d,p) level of QM theory); ΔE$_{int}$-electronic and ΔG$_{int}$-Gibbs free energies of the interaction (MP2/6-311++G(2df,pd)//B3LYP/6-311++G(d,p) level of QM theory, in kcal·mol^{-1}). ν$_i$–imaginary frequency. Intermolecular AH…B H-bonds are designated by dotted lines, their lengths H…B are presented in angstroms.
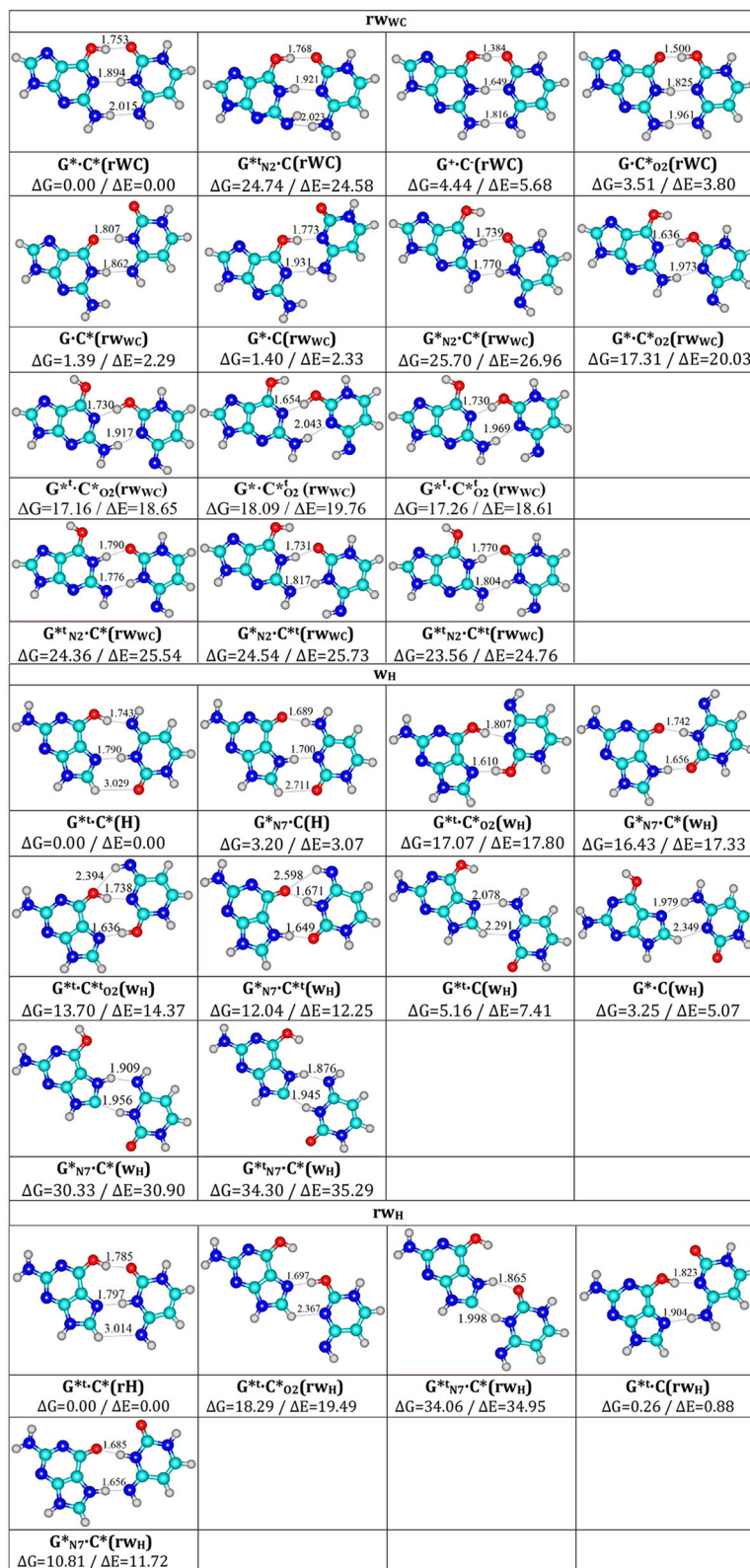
FIGURE 2 | Total geometries of the investigated G·C nucleobase pairs, corresponding to the local minima, presented altogether with their relative Gibbs free energies ΔG and electronic energies (ΔE in kcal·mol^{-1} under normal conditions) obtained at the MP2/6-311++G(2df,pd)//B3LYP/6-311++G(d,p) level of QM theory.

TABLE 1 | Energetic characteristics of the tautomers of the G·C nucleobase pairs obtained at the MP2/6-311++G(2df,pd)//B3LYP/6-311++G(d,p) level of QM theory in vacuum ($\varepsilon = 1$) (see **Figure 2**).

N	G·C complex	ΔG^a	ΔE^b
rWC/rw_WC			
1	$G^* \cdot C^*(rWC)$	0.00	0.00
2	$G \cdot C^*(rw_{WC})$	1.39	2.29
3	$G^* \cdot C(rw_{WC})$	1.40	2.33
4	$G \cdot C^*_{O2}(rWC)$	3.51	3.80
5	$G^+ \cdot C^-(rWC)$	4.44	5.68
6	$G^{*t} \cdot C^*_{O2}(rw_{WC})$	17.16	18.65
7	$G^{*t} \cdot C^{*t}_{O2}(rw_{WC})$	17.26	18.61
8	$G^* \cdot C^*_{O2}(rw_{WC})$	17.31	20.03
9	$G^* \cdot C^{*t}_{O2}(rw_{WC})$	18.09	19.76
10	$G^{*t}_{N2} \cdot C^*(rw_{WC})$	24.36	25.54
11	$G^*_{N2} \cdot C^{*t}(rw_{WC})$	24.54	25.73
12	$G^t_{N2} \cdot C^{*t}(rw_{WC})$	23.56	24.76
13	$G^*_{N2} \cdot C(rWC)$	24.74	24.58
14	$G^*_{N2} \cdot C^*(rw_{WC})$	25.70	26.96
H/w_H			
15	$G^{*t} \cdot C^*(H)$	0.00	0.00
16	$G^*_{N7} \cdot C(H)$	3.20	3.07
17	$G^* \cdot C(w_H)$	3.25	5.07
18	$G^{*t} \cdot C(w_H)$	5.16	7.41
19	$G^*_{N7} \cdot C^{*t}(w_H)$	12.04	12.25
20	$G^{*t} \cdot C^{*t}_{O2}(w_H)$	13.70	14.37
21	$G^*_{N7} \cdot C^*(w_H)$	16.43	17.33
22	$G^{*t} \cdot C^*_{O2}(w_H)$	17.07	17.80
23	$G^*_{N7} \cdot C^*(w_H)$	30.33	30.90
24	$G^t_{N7} \cdot C^*(w_H)$	34.30	35.29
rHrw_H			
25	$G^{*t} \cdot C^*(rH)$	0.00	0.00
26	$G^{*t} \cdot C(rw_H)$	0.26	0.88
27	$G^*_{N7} \cdot C^*(rw_H)$	10.81	11.72
28	$G^{*t} \cdot C^*_{O2}(rw_H)$	18.29	19.49
29	$G^{*t}_{N7} \cdot C^*(rw_H)$	34.06	34.95

[a]*Relative Gibbs free energy (T=298.15 K), kcal·mol⁻¹.*
[b]*Relative electronic energy, kcal·mol⁻¹.*

Its characteristic structural specificity has significant non-planarity and out-of-plane deformation of the purine ring of the O6H, N1H, and N2H atomic groups with *trans*-orientation relatively to the neighboring C2N3 bond.

2. Further, it was found out that Löwdin's $G^* \cdot C^*(WC)$ DNA base pair, which is formed from the classic G·C(WC) DNA base pair through the DPT and is stabilized by the participation of the three intermolecular (G)O6H...N4(C), (C)N3H...N1(G), and (G)N2H...O2(C) H-bonds (Brovarets' and Hovorun, 2014a), does not tautomerize in the wobble manner.

In this case all localized transition states of tautomerization in this manner and its pathways are the same as in the case of the wobble mutagenic tautomerization of the G·C(WC) DNA base pair, which has been investigated and described earlier (Brovarets'

and Hovorun, 2015a). In other words, in order to tautomerize in the wobble-manner, the Löwdin's $G^* \cdot C^*(WC)$ DNA base pair should revert back to the classic G·C(WC) configuration (form) (Brovarets' and Hovorun, 2015a, 2018).

This bright fact allows us to claim that the functional role of the tautomeric $G \cdot C(WC) \rightarrow G^* \cdot C^*(WC)$ transition consists in the removal of the steric obstacles for the conformational $G \cdot C(WC) \rightarrow G^* \cdot C^*(rWC)$ transition (Brovarets' et al., 2019a) and is not directly related to the origin of the spontaneous point mutations—transitions and transversions, as it was suggested earlier (please, refer to work (Brovarets' and Hovorun, 2014a) and references provided therein for more details).

This aforementioned conformational transition, in its turn, guarantees the integration of the G·C(WC) nucleobase pair into the DNA/RNA with parallel strands.

3. Opposite to the previously considered methods both the so-called correct and incorrect DNA base pairs (Brovarets' and Hovorun, 2009, 2015a,b,c,d,e,f; Brovarets' and Hovorun, 2018), the process of the tautomeric wobblization in the investigated $G^* \cdot C^*(rWC)$ (**Figure 1**, parts II and III), and $G^{*t} \cdot C^*(H)$ (**Figure 1**, part X) base pairs is preceded by the process of the tautomerization *via* the single proton transfer (SPT). At this point, both processes of the wobblization of the $G^* \cdot C^*(rWC)$ base pair occur through the joint intermediate—tight $G^+ \cdot C^-$ ion pair, which is stabilized by the participation of the three $(G)O6^+H...O2^-(C)$, $(G)N1^+H...N3^-(C)$, and $(G)N2^+H...N4^-H(C)$ H-bonds (Brovarets' and Hovorun, 2014a). This dynamically non-stable intermediate is associated with the local minimum on potential (electronic) energy surface (PES). This situation is observed for the first time. Up until now the commonly accepted idea, that mutagenic tautomerization of the classic DNA base pairs is assisted by the intermediate corresponding to the local minimum on the PES, has not been confirmed.

The first process of the tautomeric wobblization of the $G^* \cdot C^*(rWC)$ base pair (**Figure 1**, part II)— $G^* \cdot C^*(rWC) \leftrightarrow G \cdot C^*_{O2}(rWC) \leftrightarrow G \cdot C^*(rw_{WC}) \leftrightarrow G^* \cdot C(rw_{WC})$—is *most likely tightly connected with the incorporation of the G·C(WC) base pair into the DNA/RNA with parallel strands* (Watson and Crick, 1953b).

Another tautomerization process (**Figure 1**, part III)— $G^* \cdot C^*(rWC) \leftrightarrow G^*_{N2} \cdot C^*(rw_{wc}) \leftrightarrow G^* \cdot C^*_{O2}(rw_{wc})$, which proceeds through the unique $TS_{G+ \cdot C-(rWC) \leftrightarrow G^* \cdot C^* O2(rwwc)}$ path with the (G)N1-H-O2(C) covalent bridge, is most probably concerned with the mechanisms of maintaining the RNA spatial architecture due to the incorporation of the non-stable (in the main tautomeric state) Levitt base pair (Crick and Watson, 1954; Levitt, 1969). This suggestion is based on the established structural mechanism of the tautomeric interconversion of the $G^* \cdot C^*(rWC)$ pair into the eight *stable planar tautomeric forms of the Levitt base pair* (Watson and Crick, 1953a) (**Figure 1**, parts IV-VI)—$G^{*t} \cdot C^*_{O2}(rw_{WC})$, $G^* \cdot C^*_{O2}(rw_{WC})$, $G^{*t} \cdot C^{*t}_{O2}(rw_{WC})$, $G^* \cdot C^{*t}_{O2}(rw_{WC})$, $G^*_{N2} \cdot C^*(rw_{WC})$, $G^{*t}_{N2} \cdot C^*(rw_{WC})$, $G^{*t}_{N2} \cdot C^{*t}(rw_{WC})$, and $G^{*t}_{N2} \cdot C^{*t}(rw_{WC})$ (**Figure 2**,

Table 1)—and in principle, allows us to understand the dynamic of the formation of the Levitt base pair, which has not been considered before in the literature. It would be interesting to investigate how the tautomers of the Levitt base pair is stabilized in RNA by the H-bonds and surrounding environment further in the future (Oliva et al., 2007).

4. A quite interesting situation is observed for the tautomeric wobblization of the $G^{*t} \cdot C^*(H)$ base pair (**Figure 1**, part VII): $G^{*t} \cdot C^*(H) \leftrightarrow G^*_{N7} \cdot C(C) \leftrightarrow G^{*t} \cdot C^*_{O2}(w_H) \leftrightarrow G^*_{N7} \cdot C^*(w_H)$.

The transition of the C^*_{O2} tautomer of the cytosine (C) within the $G^{*t} \cdot C^*_{O2}(w_H)$ base pairs with *cis*-orientation of the N4H C-imino group into the *trans*-orientation through its inversion leads to the decreasing of the energy in the tautomerization (**Figure 1**, part VIII): $G^{*t} \cdot C^*_{O2}(w_H) \leftrightarrow G^{*t} \cdot C^{*t}_{O2}(w_H) \leftrightarrow G^*_{N7} \cdot C^{*t}(w_H) \leftrightarrow G^*_{N7} \cdot C^*(w_H)$. This decreasing of energy occurs when the affinity of the C^{*t}_{O2} tautomer according to the "complementary" G^{*t} tautomer is higher than the C^*_{O2} tautomer. This decreasing of the energy with excess overrides the increasing of the internal energy of the C^*_{O2} tautomer at its tautomerization $C^*_{O2} \rightarrow C^{*t}_{O2}$.

In the another pathway of the tautomeric wobblization of the $G^{*t} \cdot C^*(H)$ base pair (**Figure 1**, part VII, VIII) the decreasing of energy in the course of the process is achieved by the conformational transition of the G^{*t} tautomer within the $G^{*t} \cdot C(w_H)$ complex into the low-energy mutagenic tautomeric form G^*_{N7}, which is zwitterion.

5. At this, the $G^{*t} \cdot C(w_H) \leftrightarrow G^{*t}_{N7} \cdot C^*(w_H)$ DPT tautomerization process does not really occur, since its barrier is negative under normal conditions (**Figure 1**, part IX): $G^{*t} \cdot C(w_H) \leftrightarrow G^* \cdot C(w_H) \leftrightarrow G^*_{N7} \cdot C^*(w_H) \leftrightarrow G^{*t}_{N7} \cdot C^*(w_H)$.
   The same situation is also observed for the $G^{*t} \cdot C^*(H) \leftrightarrow G^*_{N7} \cdot C(H) \leftrightarrow G^{*t} \cdot C(w_H) \leftrightarrow G^{*t}_{N7} \cdot C^*(w_H)$ DPT tautomerization (**Figure 1**, part X).

6. Tautomeric wobblization of the $G^{*t} \cdot C^*(rH)$ base pair (**Figure 1**, part XI) occurs through the two traditional pathways without any preparatory SPT stages through the TSs, which represent themselves as the covalently bonded tight $G^+ \cdot C^-$ ion pairs in reverse Hoogsteen conformation, which are only supported by two H-bonds: $G^{*t} \cdot C^*(rH) \leftrightarrow G^{*t} \cdot C^*_{O2}(rw_H) \leftrightarrow G^*_{N7} \cdot C^*(rw_H)$ and $G^{*t} \cdot C^*(rH) \leftrightarrow G^{*t} \cdot C^*(rw_H) \leftrightarrow G^*_{N7} \cdot C^*(rw_H)$. The transition of the G^{*t} tautomer within the $G^{*t} \cdot C^*_{O2}(rw_H)$ complex into the G^* mutagenic tautomer through the orthogonal TS decreases the energy of the further process of tautomerization.

7. Also, in addition to the previously revealed processes, DPT tautomerization was also fixed by the participation of the proton at the C8 carbon atom of G, which lead to the dynamically-stable, but short-lived, complexes by the participation of the yilidic forms of the G base (**Figure 1**, parts IX-XI).

8. Finally, there are three more fixed mysteries, which deserve more attention. Several G·C base pairs, in which both bases were in the rare tautomeric form and their energy of stabilization significantly exceeded the analogical values for the classic G·C(WC) DNA base pair were fixed.

Despite the structural softness of the heterocycles of the G and C bases for the out-of-plane deformational bending (Hovorun et al., 1999), it was not revealed that there was any deviation from the plane in the investigated processes of the tautomerization of the base pairs.

Obtained data convincingly show that among all possible tautomeric wobblizations of the $G^* \cdot C^*(rWC)$, $G^{*t} \cdot C^*(H)$, and $G^{*t} \cdot C^*(rH)$ DNA base pairs, which possess Watson-Crick, Hoogsteen, and reverse Hoogsteen configurations and both monomers of which are in the rare tautomeric form, at least one non-dissociative transition was absent, which would recover the tautomeric status of both the G^*/G^{*t} and C^* bases to the canonical G and C bases, correspondingly. This fact altogether with the results, obtained in our previous work (Brovarets' et al., 2019a), soundly exhibits why the Watson-Crick DNA base pairs were chosen for the building of genetic material (Brovarets' et al., 2018a).

CONCLUSION

Concluding the obtained results, we arrived to a summation after providing an investigation of the tautomeric wobblization of the biologically-important G·C(WC), $G^* \cdot C^*(WC)$, $G^* \cdot C^*(rWC)$, $G^{*t} \cdot C^*(H)$, and $G^{*t} \cdot C^*(rH)$ nucleobase pairs and extended the existing thoughts about the microstructural mechanisms of these processes, as well as about their functional roles. Thus, it was established that the G·C base pair is the most likely to be incorporated into the DNA/RNA double helix with parallel strands in the form of the $G^* \cdot C^*_{O2}(rWC)$, $G \cdot C^*(rw_{WC})$, and $G^* \cdot C(rw_{WC})$ tautomers, which are in rapid tautomeric equilibrium with each other.

For the first time we have formulated rules, defining these biologically-important processes.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

OB: idea formulation, setting of the task, calculation of the data, building of the graphs, data extrapolation, preparing, and proofreading of the draft of the manuscript. AM: idea formulation, calculation of the data, building of the graphs, preparing, and proofreading of the draft of the manuscript. DH: idea formulation, preparing, and proofreading of the draft of the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

REFERENCES

Alvey, H. S., Gottardo, F. L., Nikolova, E. N., and Al-Hashimi, H. M. (2014). Widespread transient Hoogsteen base-pairs in canonical duplex DNA with variable energetics. *Nat. Comm.* 5, 4786–4794. doi: 10.1038/ncomms5786

Atkins, P. W. (1998). *Physical Chemistry*. Oxford: Oxford University Press.

Bader, R. F. W. (1990). *Atoms in Molecules: A Quantum Theory*. Oxford: Oxford University Press.

Bayley, S. T. (1951). The dielectric properties of various solid crystalline proteins, amino acids and peptides. *Trans. Faraday Soc.* 47, 509–517. doi: 10.1039/tf9514700509

Boys, S. F., and Bernardi, F. (1970). The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* 19, 553–566. doi: 10.1080/00268977000101561

Brovarets', O. O., and Hovorun, D. M. (2009). Physicochemical mechanism of the wobble DNA base pairs Gua·Thy and Ade·Cyt transition into the mismatched base pairs Gua*·Thy and Ade·Cyt* formed by the mutagenic tautomers. *Ukr. Bioorg. Acta* 8, 12–18.

Brovarets', O. O., and Hovorun, D. M. (2010a). Quantum-chemical investigation of tautomerization ways of Watson-Crick DNA base pair guanine-cytosine. *Ukr. Biochem. J.* 82, 55–60.

Brovarets', O. O., and Hovorun, D. M. (2010b). Quantum-chemical investigation of the elementary molecular mechanisms of pyrimidine·purine transversions. *Ukr. Biochem. J.* 82, 57–67.

Brovarets', O. O., and Hovorun, D. M. (2014a). Why the tautomerization of the G·C Watson–Crick base pair *via* the DPT does not cause point mutations during DNA replication? QM and QTAIM comprehensive analysis. *J. Biomol. Struct. Dynam.* 32, 1474–1499. doi: 10.1080/07391102.2013.822829

Brovarets', O. O., and Hovorun, D. M. (2014b). Can tautomerisation of the A·T Watson-Crick base pair *via* double proton transfer provoke point mutations during DNA replication? A comprehensive QM and QTAIM analysis. *J. Biomol. Struct. Dynam.* 32, 127–154. doi: 10.1080/07391102.2012.755795

Brovarets', O. O., and Hovorun, D. M. (2015a). New structural hypostases of the A·T and G·C Watson-Crick DNA base pairs caused by their mutagenic tautomerisation in a wobble manner: a QM/QTAIM prediction. *RSC Adv.* 5, 99594–99605. doi: 10.1039/C5RA19971A

Brovarets', O. O., and Hovorun, D. M. (2015b). Tautomeric transition between wobble A·C DNA base mispair and Watson-Crick-like A·C* mismatch: microstructural mechanism and biological significance. *Phys. Chem. Chem. Phys.* 17, 15103–15110. doi: 10.1039/C5CP01568E

Brovarets', O. O., and Hovorun, D. M. (2015c). How many tautomerisation pathways connect Watson-Crick-like G*·T DNA base mispair and wobble mismatches? *J. Biomol. Struct. Dynam.* 33, 2297–2315. doi: 10.1080/07391102.2015.1046936

Brovarets', O. O., and Hovorun, D. M. (2015d). Wobble↔Watson-Crick tautomeric transitions in the homo-purine DNA mismatches: a key to the intimate mechanisms of the spontaneous transversions. *J. Biomol. Struct. Dynam.* 33, 2710–2715. doi: 10.1080/07391102.2015.1077737

Brovarets', O. O., and Hovorun, D. M. (2015e). Novel physico-chemical mechanism of the mutagenic tautomerisation of the Watson–Crick-like A·G and C·T DNA base mispairs: a quantum-chemical picture. *RSC Adv.* 5, 66318–66333. doi: 10.1039/C5RA11773A

Brovarets', O. O., and Hovorun, D. M. (2015f). A novel conception for spontaneous transversions caused by homo-pyrimidine DNA mismatches: a QM/QTAIM highlight. *Phys. Chem. Chem. Phys.* 17, 21381–21388. doi: 10.1039/C5CP03211C

Brovarets', O. O., and Hovorun, D. M. (2015g). The nature of the transition mismatches with Watson-Crick architecture: the G*·T or G·T* DNA base mispair or both? A QM/QTAIM perspective for the biological problem. *J. Biomol. Struct. Dynam.* 33, 925–945. doi: 10.1080/07391102.2014.924879

Brovarets', O. O., and Hovorun, D. M. (2018). "Renaissance of the tautomeric hypothesis of the spontaneous point mutations in DNA: new ideas and computational approaches," in *Mitochondrial DNA—New Insights*, ed Herve Seligmann (London: IntechOpen), 31–55. doi: 10.5772/intechopen.77366

Brovarets', O. O., Oliynyk, T. A., and Hovorun, D. M. (2019a). Novel tautomerisation mechanisms of the biologically important conformers of the reverse Löwdin, Hoogsteen, and reverse Hoogsteen G*·C* DNA base

pairs *via* proton transfer: a quantum-mechanical survey. *Front. Chem.* 7:597. doi: 10.3389/fchem.2019.00597

Brovarets', O. O., Oliynyk, T. A., and Hovorun, D. M. (2019b). "Novel conformers of the G·C DNA base pair and their mutual interconversions *via* the proton transfer: a quantum-mechanical study" in *Joint 12th EBSA Congress and 10th ICBP-IUPAP Congress*, Vol. 48 (Madrid), S90. doi: 10.1007/s00249-019-01373-4

Brovarets', O. O., Tsiupa, K. S., Dinets, A., and Hovorun, D. M. (2018c). Unexpected routes of the mutagenic tautomerization of the T nucleobase in the classical A·T DNA base pairs: A QM/QTAIM comprehensive view. *Front. Chem.* 6:532. doi: 10.3389/fchem.2018.00532

Brovarets', O. O., Tsiupa, K. S., and Hovorun, D. M. (2018a). The A·T(rWC)/A·T(H)/A·T(rH) ↔ A·T*(rwWC)/A·T*(wH)/A·T*(rwH) mutagenic tautomerization *via* sequential proton transfer: a QM/QTAIM study. *RSC Adv.* 8, 13433–13445. doi: 10.1039/C8RA01446A

Brovarets', O. O., Tsiupa, K. S., and Hovorun, D. M. (2018b). Unexpected A·T(WC)↔A·T(rWC)/A·T(rH) and A·T(H)↔A·T(rH)/A·T(rWC) conformational transitions between the classical A·T DNA base pairs: A QM/QTAIM comprehensive study. *Int. J. Quantum. Chem.* 118:e25674. doi: 10.1002/qua.25692

Brovarets', O. O., Tsiupa, K. S., and Hovorun, D. M. (2018d). Non-dissociative structural transitions of the Watson-Crick and reverse Watson-Crick A·T DNA base pairs into the Hoogsteen and reverse Hoogsteen forms. *Sci. Repts.* 8:10371. doi: 10.1038/s41598-018-28636-y

Brovarets', O. O., Tsiupa, K. S., and Hovorun, D. M. (2018e). Novel pathway for mutagenic tautomerization of classical A·T DNA base pairs *via* sequential proton transfer through quasi-orthogonal transition states: a QM/QTAIM investigation. *PLoS ONE* 13:e0199044. doi: 10.1371/journal.pone.0199044

Brovarets', O. O., Tsiupa, K. S., and Hovorun, D. M. (2018f). Surprising conformers of the biologically important A·T DNA base pairs: QM/QTAIM proofs. *Front. Chem.* 6:8. doi: 10.3389/fchem.2018.00008

Brovarets', O. O., Zhurakivsky, R. O., and Hovorun, D. M. (2014). Does the tautomeric status of the adenine bases change upon the dissociation of the A*·A_syn Topal-Fresco DNA mismatch? A combined QM and QTAIM atomistic insight. *Phys. Chem. Chem. Phys.* 16, 3715–3725. doi: 10.1039/c3cp54708f

Brovarets', O. O., Zhurakivsky, R. O., and Hovorun, D. M. (2015). DPT tautomerisation of the wobble guanine·thymine DNA base mispair is not mutagenic: QM and QTAIM arguments. *J. Biomol. Struct. Dynam.* 33, 674–689. doi: 10.1080/07391102.2014.897259

Crick, F. H. C., and Watson, J. D. (1954). The complementary structure of deoxyribonucleic acid. *Proc. Roy. Soc.* A223, 80–96. doi: 10.1098/rspa.1954.0101

Cukrowski, I., and Matta, C. F. (2010). Hydrogen–hydrogen bonding: a stabilizing interaction in strained chelating rings of metal complexes in aqueous phase. *Chem. Phys. Lett.* 499, 66–69. doi: 10.1016/j.cplett.2010.09.013

Dewar, M. J. S., and Storch, D. M. (1985). Alternative view of enzyme reactions. *Proc. Natl. Acad. Sci. U. S. A.* 82, 2225–2229. doi: 10.1073/pnas.82.8.2225

El-Sayed, A. A., Tamara Molina, A., Alvarez-Ros, M. C., and Alcolea Palafox, M. (2015). Conformational analysis of the anti-HIV Nikavir prodrug: comparisons with AZT and thymidine, and establishment of structure-activity relationships/tendencies in other 6'-derivatives. *J. Biomol. Struct. Dynam.* 33, 723–748. doi: 10.1080/07391102.2014.909743

Florian, J., Hrouda, V., and Hobza, P. (1994). Proton transfer in the adenine-thymine base pair. *J. Am. Chem. Soc.* 116,1457–1460. doi: 10.1021/ja00083a034

Frisch, M. J., Head-Gordon, M., and Pople, J. A. (1990). Semi-direct algorithms for the MP2 energy and gradient. *Chem. Phys. Lett.* 166, 281–289. doi: 10.1016/0009-2614(90)80030-H

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2010). *GAUSSIAN 09 (Revision B.01)*. Wallingford CT: Gaussian Inc.

García-Moreno, B. E., Dwyer, J. J., Gittis, A. G., Lattman, E. E., Spencer, D. S., and Stites, W. E. (1997). Experimental measurement of the effective dielectric in the hydrophobic core of a protein. *Biophys. Chem.* 64, 211–224. doi: 10.1016/S0301-4622(96)02238-7

Godbeer, A. D., Al-Khalili, J. S., and Stevenson, P. D. (2015). Modelling proton tunnelling in the adenine-thymine base pair. *Phys. Chem. Chem. Phys.* 7, 13034–13044. doi: 10.1039/C5CP00472A

Gorb, L., Podolyan, Y., Dziekonski, P., Sokalski, W. A., and 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. doi: 10.1021/ja049155n

Gutowski, M., Van Lenthe, J. H., Verbeek, J., Van Duijneveldt, F. B., and Chalasinski, G. (1986). The basis set superposition error in correlated electronic structure calculations. *Chem. Phys. Lett.* 124, 370–375. doi: 10.1016/0009-2614(86)85036-9

Hariharan, P. C., and Pople, J. A. (1973). The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta* 28, 213–222. doi: 10.1007/BF00533485

Hovorun, D. M., Gorb, L., and 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. doi: 10.1002/(SICI)1097-461X(1999)75:3<245::AID-QUA14>3.0.CO;2-0

Keith, T. A. (2010). *AIMAll (Version 10.07.01).* Available online at: aim.tkgristmill.com (accessed October 23, 2020).

Kendall, R. A., Dunning, J.r,., T. H., and Harrison, R.J. (1992). Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* 96, 6796–6806. doi: 10.1063/1.462569

Krishnan, R., Binkley, J. S., Seeger, R., and Pople, J. A. (1980). Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* 72, 650–654. doi: 10.1063/1.438955

Lecomte, C., Espinosa, E., and Matta, C. F. (2015). On atom–atom "short contact" bonding interactions in crystals. *IUCrJ* 2, 161–163. doi: 10.1107/S2052252515002067

Lee, C., Yang, W., and Parr, R. G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B.* 37, 785–789. doi: 10.1103/PhysRevB.37.785

Levitt, M. (1969). Detailed molecular model for transfer ribonucleic acid. *Nature* 224, 759–763. doi: 10.1038/224759a0

Löwdin, P.-O. (1963). Proton tunneling in DNA and its biological implications. *Rev. Mod. Phys.* 35, 724–732. doi: 10.1103/RevModPhys.35.724

Löwdin, P.-O. (1966). "Quantum genetics and the aperiodic solid: some aspects on the biological problems of heredity, mutations, aging, and tumors in view of the quantum theory of the DNA molecule," in *Advances in Quantum Chemistry*, ed P.-O. Löwdin (New York, NY; London: Academic Press), 213–360. doi: 10.1016/S0065-3276(08)60076-3

Matta, C. F. (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. doi: 10.1002/jcc.21417

Matta, C. F. (2014). Modeling biophysical and biological properties from the characteristics of the molecular electron density, electron localization and delocalization matrices, and the electrostatic potential. *J. Comput. Chem.* 35, 1165–1198. doi: 10.1002/jcc.23608

Matta, C. F., Castillo, N., and Boyd, R. J. (2006). Atomic contributions to bond dissociation energies in aliphatic hydrocarbons. *J. Chem. Phys.* 125:204103. doi: 10.1063/1.2378720

Matta, C. F., and Hernández-Trujillo, J. (2003). Bonding in polycyclic aromatic hydrocarbons in terms of the electron density and of electron delocalization. *J. Phys. Chem. A* 107, 7496–7504. doi: 10.1021/jp034952d

Mertz, E. L., and Krishtalik, L. I. (2000). Low dielectric response in enzyme active site. *Proc. Natl. Acad. Sci. U. S. A.* 97, 2081–2086. doi: 10.1073/pnas.050316997

Nikolova, E. N., Zhou, H., Gottardo, F. L., Alvey, H. S., Kimsey, I. J., and Al-Hashimi, H. M. (2014). A historical account of Hoogsteen base pairs in duplex DNA. *Biopolymers* 99, 955–968. doi: 10.1002/bip.22334

Oliva, R., Tramontano, A., and Cavallo, L. (2007). Mg^{2+} binding and archaeosine modification stabilize the G15-C48 Levitt base pair in tRNAs. *RNA* 13, 1427–1436. doi: 10.1261/rna.574407

Palafox, M. A. (2014). Molecular structure differences between the antiviral nucleoside analogue 5-iodo-2'-deoxyuridine and the natural nucleoside 2'-deoxythymidine using MP2 and DFT methods: conformational analysis, crystal simulations, DNA pairs and possible behavior. *J. Biomol. Struct. Dynam.* 32, 831–851. doi: 10.1080/07391102.2013.789402

Parr, R. G., and Yang, W. (1989). *Density-Functional Theory of Atoms and Molecules.* Oxford: Oxford University Press.

Peng, C., Ayala, P. Y., Schlegel, H. B., and Frisch, M. J. (1996). Using redundant internal coordinates to optimize equilibrium geometries and transition states. *J. Comput. Chem.* 17, 49–56. doi: 10.1002/(SICI)1096-987X(19960115)17:1<49::AID-JCC5>3.0.CO;2-0

Petrushka, J., Sowers, L. C., and Goodman, M. (1986). Comparison of nucleotide interactions in water, proteins, and vacuum: model for DNA polymerase fidelity. *Proc. Natl. Acad. Sci. U. S. A.* 83, 1559–1562. doi: 10.1073/pnas.83.6.1559

Poltev, V. I., Anisimov, V. M., Sanchez, C., Deriabina, A., Gonzalez, E., Garcia, D., et al. (2016). Analysis of the conformational features of Watson–Crick duplex fragments by molecular mechanics and quantum mechanics methods. *Biophysics* 61, 217–226. doi: 10.1134/S0006350916020160

Pous, J., Urpi, L., Subirana, J. A., Gouyette, C., Navaza, J., and Campos, J. L. (2008). Stabilization by extra-helical thymines of a DNA duplex with Hoogsteen base pairs. *J. Am. Chem. Soc.* 130, 6755–6760. doi: 10.1021/ja078022+

Saenger, W. (1984). *Principles of Nucleic Acid Structure.* New York, NY: Springer. doi: 10.1007/978-1-4612-5190-3

Sordo, J. A. (2001). On the use of the Boys–Bernardi function counterpoise procedure to correct barrier heights for basis set superposition error. *J. Mol. Struct.* 537, 245–251. doi: 10.1016/S0166-1280(00)00681-3

Sordo, J. A., Chin, S., and Sordo, T. L. (1988). On the counterpoise correction for the basis set superposition error in large systems. *Theor. Chim. Acta* 74, 101–110. doi: 10.1007/BF00528320

Srivastava, R. (2019). The role of proton transfer on mutations. *Front Chem.* 7:536. doi: 10.3389/fchem.2019.00536

Szabat, M., and Kierzek, R. (2017). Parallel-stranded DNA and RNA duplexes: structural features and potential applications. *FEBS J.* 284, 3986–3998. doi: 10.1111/febs.14187

Tirado-Rives, J., and Jorgensen, W. L. (2008). Performance of B3LYP density functional methods for a large set of organic molecules. *J. Chem. Theory Comput.* 4, 297–306. doi: 10.1021/ct700248k

Topal, M. D., and Fresco, J. R. (1976). Complementary base pairing and the origin of substitution mutations. *Nature* 263, 285–289. doi: 10.1038/263285a0

Turaeva, N., and Brown-Kennerly, V. (2015). Marcus model of spontaneous point mutation in DNA. *Chem. Phys.* 461, 106–110. doi: 10.1016/j.chemphys.2015.09.005

Watson, J. D., and Crick, F. H. C. (1953a). Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* 171, 737–738. doi: 10.1038/171737a0

Watson, J. D., and Crick, F. H. C. (1953b). The structure of DNA. *Cold. Spring Harb. Symp. Quant. Biol.* 18, 123–131. doi: 10.1101/SQB.1953.018.01.020