Structural and energetic properties of the potential HIV-1 reverse transcriptase inhibitors d4A and d4G: a comprehensive theoretical investigation

Alla G. Ponomareva, Yevgen P. Yurenko, Roman O. Zhuravik, Tanja van Mourik and Dmytro M. Hovorun

A comprehensive quantum-chemical investigation of the conformational landscapes of two nucleoside HIV-1 reverse transcriptase inhibitors, 2′,3′-didehydro-2′,3′-dideoxyadenosine (d4A), and 2′,3′-didehydro-2′,3′-dideoxyguanosine (d4G), has been performed at the MP2/6-311++G(d,p)//B3LYP/6-31G(d,p) level of theory. It was found that d4A can adopt 21 conformers within a 5.17 kcal/mol Gibbs free energy range, whereas d4G has 20 conformers within 6.23 kcal/mol at T = 298.15 K. Both nucleosides are shaped by a sophisticated network of specific noncovalent interactions, including conventional (OH···O, NH···O) and weak (CH···O, CH···N) hydrogen bonds, as well as dihydrogen (CH···HC) contacts. For the OH···O, NH···O, and CH···O hydrogen bonds, natural bond orbital analysis revealed hyperconjugative interactions between the oxygen lone pairs and the antibonding orbital of the donor group. For the CH···HC contacts, the electron density migrates from the antibonding orbital, corresponding to the CH group of the sugar residue, to the bonding orbital relative to the same group in the nucleobase. The results confirm the current belief that the biological activity of d4A and d4G is connected with the termination of the DNA chain synthesis in the 5′-3′ direction. Thus, these nucleosides act as competitive HIV-1 reverse transcriptase inhibitors.

**Keywords:** 2′,3′-didehydro-2′,3′-dideoxyadenosine; 2′,3′-didehydro-2′,3′-dideoxyguanosine; d4A; d4G; conformational analysis; hydrogen bonds; H-bonds; biological activity; nucleoside reverse transcriptase inhibitors; density functional theory; electron density topological analysis; NBO analysis; compliance constants

**Introduction**

The advent of computer technology coupled with improvements in computational chemistry algorithms has given impetus to *in silico* design of new efficient drugs against acquired immunodeficiency syndrome. HIV-1 reverse transcriptase (hereafter referred to as RT) is a multifunctional enzyme that catalyzes two different processes: (i) DNA synthesis using RNA or DNA as the template; (ii) endonucleolytic ribonuclease H (RNase H) activity that specifically degrades the RNA strand of RNA:DNA duplexes (see e.g. Sluis-Cremer, Temiz, & Bahar, 2004 and refs. therein). Due to its paramount role in the HIV-1 life cycle, RT has been the main target in the fight against this virus (Skowron & Ogden, 2006). In general, RT inhibitors can be classified into two groups: nucleoside RT inhibitors (NRTIs), which affect the DNA synthesis by blocking the growth of a polynucleotide chain (Jeang, 2000), and non-nucleoside RT inhibitors (NNRTIs), which are alloteric and noncompetitive inhibitors (Sahlberg & Zhou, 2008).

The conformational properties of NRTIs play a key role in their binding to target enzymes (see e.g. Herdewijn, 2008 and refs. therein). When introduced into a cell, a nucleoside passes through several phases of biomolecular recognition by nucleoside transporters (King, Ackley, Cass, Young, & Baldwin, 2006), kinases/5′-nucleotidases (Van Rompay, Johansson, & Karlsson, 2003) and, finally, polymerases (Alber & Carloni, 2000). At
each phase, the nucleoside can adopt a unique three-dimensional structure or conformation. The structural properties of NRTIs, particularly, those governed by intramolecular interactions, determine to a great extent their potential therapeutic efficiency. For instance, efficient phosphorylation depends largely on the spatial structure of the nucleoside (el Kouni, 2002). Therefore, an extensive conformational analysis identifying all minima on the potential energy surface can be considered as a first step in the design of nucleoside analogs as anti-HIV drugs.

The 2′,3′-didehydro-2′,3′-dideoxynucleosides are promising drug candidates for anti-retroviral treatment. The structural features that distinguish them from the canonical nucleosides are the presence of a double bond in the sugar residue, making the sugar almost planar, and the lack of a hydroxyl group at the 3′-position of the base. It is thought that, after activation by cellular kinases and interaction with the RT nucleoside triphosphate binding site, the 2′,3′-didehydro-2′,3′-dideoxy nucleosides act as terminators of viral replication (Everaert et al., 1993). So far, only 2′,3′-didehydro-2′,3′-dideoxythymidine (d4T, stavudine) has been widely used as part of combined antiretroviral therapy (Hitchcock, 1991), due to its favorable risk/benefit profile (Russell, Whiterock, Marrero, & Klunk, 1989). The other 2′,3′-didehydro-2′,3′-dideoxy nucleosides containing canonical DNA and RNA bases are also considered as potential RT inhibitors (Everaert et al., 1993). However, the full mechanisms of their biological activity are still unknown.

This work is devoted to a comprehensive conformational analysis of two nucleoside analogs, 2′,3′-didehydro-2′,3′-dideoxyadenosine (d4A), and 2′,3′-didehydro-2′,3′-dideoxyguanosine (d4G). In earlier work, we studied the d4T (Ponomareva, Yurenko, Zhurakivsky, van Mourik, & Hovorun, 2013), d4U (2′,3′-didehydro-2′,3′-dideoxyuridine) and d4C (2′,3′-didehydro-2′,3′-dideoxyctydidine) (Ponomareva, Yurenko, Zhurakivsky, van Mourik, & Hovorun, 2012) nucleosides and thus, the current work completes our extensive investigation of the conformational families of the 2′,3′-didehydro-2′,3′-dideoxy analogs of the canonical nucleosides. The d4A and d4G nucleosides are prospective drug candidates. For instance, it was earlier established that d4A and its derivatives exhibit a noticeable activity with respect to HIV-1 (Mitsuya, Yarchoan, Kageyama, & Broder, 1991) as well as against the hepatitis B virus (Balzarini et al., 1997). In contrast to d4A, the d4G nucleoside exhibits relatively low activity against HIV-1 (Chu et al., 1988). However, slight chemical modifications of d4G such as, addition of a cyclopropylamino group at the 6-position of the base, result in increased stability, lipophilicity, and solubility, as well as decreased toxicity relative to unmodified d4G (Ray, Yang, Chu, & Anderson, 2002).

X-ray structural analysis of d4A (Hutcheon & James, 1974) and d4G (van Roey & Chu, 1992) attested that both nucleosides adopt high-anti conformations in the crystal state, which differ by the orientation of the γ (C4′-C5′) torsion angle. These high-anti conformations are energetically unfavorable in the gas phase. Therefore, a theoretical investigation of the d4A and d4G conformational landscapes is essential for determining the full set of biologically active conformations.

Several earlier works were devoted to the conformational analysis of 2′,3′-didehydro-2′,3′-dideoxynucleoside analogs. Conformational energy calculations of d4A (Rao, 1998a) and d4G (Rao, 1998b) were performed by Rao employing different force fields. It was established that the preferred combination for the glycosidic (χ) and C4′-C5′ bond (γ) conformations are (χ ∈ anti, γ ∈ gauche−) and (χ ∈ anti, γ ∈ trans) for d4A and d4G, respectively. However, the consistency of the calculated preferred conformations with the X-ray data (Hutcheon & James, 1974; van Roey & Chu, 1992) appeared to be very sensitive to the force field employed. In a recent study, Palafoux, Iza, de la Fuente, and Navarro (2009) computed structures and energies of selected conformers of stavudine (d4T), in the isolated phase as well as in the presence of solvent, to analyze the effect of hydration on the d4T molecular structure and energetics.

In recent works (Yurenko, Zhurakivsky, Ghomi, Samijlenko, & Hovorun, 2007a, 2007b, 2008; Yurenko, Zhurakivsky, Samijlenko, Ghomi, & Hovorun, 2007; Yurenko, Zhurakivsky, Samijlenko, & Hovorun, 2011), we developed a novel approach to conformational analysis of nucleic acid constituents. This methodology allows identification of all conformers without performing a full potential energy surface scan. In particular, the approach permits a reliable reconstruction of the vibrational spectra of the nucleosides (Yurenko et al., 2007, 2008). A comprehensive conformational investigation of d4T (Ponomareva et al., 2013) showed that 19 d4T conformers are within 5.51 kcal/mol of each other and are governed by eight different types of specific intramolecular interactions.

In the current work, we apply this methodology to the d4A and d4G nucleosides, since, to the best of our knowledge, their conformational space has not yet been comprehensively explored. As the biological activity of nucleosides is structure-dependent, the results will be useful, for example, for molecular docking investigations of enzyme-substrate systems.

**Computational methodology**

The chemical structures of d4A and d4G are presented in Figure 1, together with the atom numbering and definition of the dihedral angles (Saenger, 1984). The sugar ring pseudorotational concept by Altona and Sundaralingam (1972) was employed in this work. A conformational analysis was performed by density functional theory (DFT) using the B3LYP nonlocal hybrid
Results and discussion

Conformational characteristics

Tables 1 and 2 list the energetic and structural properties of the d4A and d4G conformers located, respectively. The most energetically favorable four conformers of d4A and d4G are displayed in Figures 2 and 3, respectively. The d4A and d4G global minima are both syn/O4′-endo, whereas the most stable d4G conformer is syn/O4′-endo. The d4A global minimum is shaped by a single O5′H···N3 H-bond, whereas the d4G global minimum contains two rather strong O5′H···N3 and N2H1···O5 H-bonds. X-ray diffraction data (Hutcheon & James, 1974; van Roey & Chu, 1992) (Table 3), shows that in the crystal state d4A and d4G adopt structures that are high-energy conformers in the gas phase, with very different β, γ, and χ torsion angles than in DNA (Hartmann & Lavery, 1996). Some differences between the crystal structures and isolated conformers, especially in the χ angle values, are due to strong intermolecular H-bonds and crystal packing forces.

The full conformational families of d4A and d4G contain only one DNA-like conformer (conformer 5 for d4A and conformer 4 for d4G, χ ∈ anti, χ ∈ ε; see Tables 1 and 2). As the sugar in d4A and d4G lacks the O3′H hydroxyl group occurring in DNA, these nucleosides do not have the δ and ε torsion angles associated with this group. Both d4A conformer 5 and d4G conformer 4 are governed by C8H···O5′ H-bonds.

The full ground-state conformational families of the 2′,3′-didehydro-2′,3′-dideoxy analogs of the natural purine nucleosides are relatively small: At 298.15 K 21 d4A conformers lie within about 5 kcal/mol Gibbs free energy of each other (Table 1), whereas 20 d4G conformers lie within about 6 kcal/mol of each other (Table 2). The number of conformers in the d4A and d4G conformational families is much smaller than the number of starting structures (36). We think this is due to the existence of relatively strong H-bonds of the O5′H···N3, N2H1···O5′ and C6H···O5′ types, which have a stabilizing effect, as well as to other types of intramolecular interactions (such as Van der Waals forces and anomeric effects (Thibaudeau, Acharya, & Chattopadhyaya, 2005)). The combination of the preference for formation of these interactions with steric hindrances prevents the formation of some conformers.

The d4A and d4G families are much smaller than those of the corresponding canonical nucleosides (dA: 88 conformers; dG: 97 conformers (Yurenko,
Zhurakivsky, & Hovorun, 2009)). The considerable reduction in the number of conformers on going from the canonical nucleosides to their 2′,3′-didehydro-2′,3′-dideoxy analogs was also observed for dT, dU, and dC (88 dT vs. 19 d4T conformers (Ponomareva et al., 2013); 94 dU (Yurenko et al., 2008) vs. 20 d4U

Table 1. Energetic and structural properties of the d4A conformers calculated at the MP2/6-311++G(d,p)//B3LYP/6-31G(d,p) level of theory.a

| Conformer | ΔG, kcal/mol | μ, D | ν_max, degrees | γ, degrees | β, degrees | τ_max, degrees | R_CyNO, Å |
|-----------|--------------|-----|----------------|------------|------------|----------------|----------|
| 1         | 0.00         | 3.34| 7.7            | 55.3       | 35.4       | 54.5           | 0.8      |
| 2         | 1.77         | 2.54| 3.3            | 106.1      | 62.5       | 62.9           | -0.6     |
| 3         | 2.14         | 3.43| 1.5            | 68.8       | 175.5      | -59.4          | 0.9      |
| 4         | 2.15         | 3.06| 0.5            | 103.6      | 174.5      | -54.1          | 0.6      |
| 5         | 2.49         | 4.72| 1.7            | 120.8      | 45.5       | 163.3          | -0.6     |
| 6         | 2.78         | 2.08| 2.0            | 167.9      | 49.6       | 56.2           | -1.5     |
| 7         | 3.03         | 3.60| 2.8            | 104.4      | -66.1      | -177.2         | 0.6      |
| 8         | 3.23         | 1.25| 1.1            | 67.4       | -69.5      | 168.7          | 0.9      |
| 9         | 3.45         | 4.16| 2.4            | 134.2      | 50.9       | -60.3          | -0.7     |
| 10        | 3.60         | 2.67| 2.9            | 104.6      | -65.3      | -74.7          | 0.6      |
| 11        | 3.63         | 2.58| 2.3            | 175.5      | -67.7      | 177.9          | -0.9     |
| 12        | 3.99         | 2.78| 1.2            | 66.8       | -67.7      | 73.5           | 0.9      |
| 13        | 4.10         | 2.24| 1.4            | 68.8       | -66.0      | 90.0           | 1.0      |
| 14        | 4.17         | 2.96| 9.1            | 178.5      | 178.6      | -176.2         | -0.7     |
| 15        | 4.24         | 2.87| 1.7            | 176.7      | -66.5      | -74.2          | -0.8     |
| 16        | 4.38         | 4.46| 1.6            | 105.3      | -170.4     | -167.2         | 0.7      |
| 17        | 4.65         | 3.44| 2.4            | 105.5      | -67.9      | 80.4           | 0.6      |
| 18        | 4.71         | 3.40| 8.5            | 178.2      | 169.3      | 57.3           | -0.5     |
| 19        | 4.83         | 0.98| 3.7            | 173.8      | -67.6      | 81.5           | -1.0     |
| 20        | 5.14         | 4.10| 4.2            | 108.8      | 179.7      | 60.1           | 0.7      |
| 21        | 5.17         | 3.17| 1.7            | 68.5       | -179.4     | 62.8           | 1.1      |

aΔG = relative Gibbs free energy at 298.15 K; μ = dipole moment; ν_max = maximum degree of pucker; ν_max = 2ν'/max (Saenger, 1984). γ, β are the backbone torsion angles (see Figure 1); τ_max = maximum torsion angle in the base; R_CyNO is the glycosidic bond length. All torsion angles are in degrees. Structural characteristics were obtained at the B3LYP/6-31G(d,p) level of theory; relative electronic energies were calculated at the MP2/6-311++G(d,p)/B3LYP/6-31G(d,p) level. The conformers are numbered in order of increasing Gibbs free energy. The puckering parameters were calculated using an in-house Scilab routine.

Table 2. Energetic and structural properties of the d4G conformers calculated at the MP2/6-311++G(d,p)//B3LYP/6-31G(d,p) level of theory.a

| Conformer | ΔG, kcal/mol | μ, D | ν_max, degrees | γ, degrees | β, degrees | τ_max, degrees | R_CyNO, Å |
|-----------|--------------|-----|----------------|------------|------------|----------------|----------|
| 1         | 0.00         | 5.65| 8.5            | 53.5       | 39.1       | 44.9           | 1.0      |
| 2         | 2.00         | 7.72| 3.0            | -103.2     | 59.6       | 64.4           | 1.8      |
| 3         | 2.55         | 7.11| 1.1            | -100.5     | 174.9      | -55.7          | 1.7      |
| 4         | 3.00         | 6.72| 1.0            | -115.8     | 45.0       | 158.8          | 1.9      |
| 5         | 3.11         | 5.45| 1.0            | -98.9      | -66.2      | -178.8         | 1.8      |
| 6         | 3.19         | 6.57| 2.7            | 69.4       | 174.5      | -56.3          | 1.5      |
| 7         | 3.72         | 7.82| 1.1            | 173.1      | 47.6       | 60.7           | 1.2      |
| 8         | 4.03         | 7.57| 1.5            | -100.4     | -65.8      | -75.2          | 1.7      |
| 9         | 4.31         | 7.20| 2.7            | 71.7       | -67.5      | -176.1         | 2.0      |
| 10        | 4.42         | 5.53| 1.6            | -100.1     | -170.1     | -163.3         | 1.9      |
| 11        | 4.54         | 7.10| 1.2            | 69.5       | -65.4      | -72.5          | 1.7      |
| 12        | 4.55         | 6.42| 0.6            | -100.0     | -67.6      | 83.0           | 1.8      |
| 13        | 4.78         | 8.75| 2.3            | -132.3     | 50.5       | -61.6          | 2.1      |
| 14        | 5.00         | 5.62| 4.9            | 30.1       | 156.6      | -62.0          | 2.7      |
| 15        | 5.43         | 7.75| 9.0            | -177.8     | 178.3      | -173.6         | 1.8      |
| 16        | 5.53         | 7.94| 3.3            | -172.3     | -68.3      | 83.6           | 2.0      |
| 17        | 5.61         | 6.71| 2.4            | 69.0       | -67.8      | 78.4           | 1.6      |
| 18        | 5.66         | 8.18| 1.2            | -103.4     | -179.0     | 61.8           | 1.9      |
| 19        | 6.01         | 9.10| 8.4            | -174.6     | 168.9      | 56.8           | 1.9      |
| 20        | 6.23         | 8.29| 2.2            | 67.8       | -179.3     | 60.4           | 1.6      |

aSee footnote, Table 1.
conformers (Ponomareva et al., 2012); and 88 dC (Yurenko et al., 2011; Zhurakivsky & Hovorun, 2006) vs. 19 d4C conformers (Ponomareva et al., 2012)). This is caused by the rigid 2,5-dihydrofuran sugar residue, which also remains planar in the crystalline state as attested by X-ray diffraction experiments (Van Roey, Taylor, Chu, & Schinazi, 1993). For all d4A and d4G conformers, the maximum amplitude of the sugar pucker \( \nu_{\text{max}} \) (see footnote, Table 1 for its definition) is smaller than 10°. It is noteworthy that conformers 14 (d4A, Table 1) and 15 (d4G, Table 2) have the largest pucker (9°) in the two conformational families. These large puckers result from a rather strong dipole–dipole interaction between the base and the sugar moieties in these

Figure 2. Spatial structure of the most energetically favorable conformers of d4A. Non-covalent intramolecular interactions are marked by dotted lines.

Figure 3. Spatial structure of the most energetically favorable conformers of d4G. Non-covalent interactions are marked by dotted lines.

---

734  A.G. Ponomareva et al.
conformers. In an earlier study (Gelbin et al., 1996), it was shown that the average $\nu_{\text{max}}$ values of the DNA and RNA sugar moieties, deoxyribose, and ribose, are 35° and 37°, respectively, i.e. much larger than those calculated for the d4A and d4G conformers. Thus, unlike their canonical counterparts, the 2',3'-didehydro-2',3'-dideoxyribonucleosides are conformationally restrained.

The $\chi$ angle of the syn conformers of d4A and d4G varies from 55° to 69° for d4A and from 30° to 72° for d4G, the second interval being much larger. The $\chi$ angle of the anti conformers ranges from 168° to 179° for d4A and from $\leq 116°$ to $\geq 99°$ for d4G (see Tables 1 and 2). The corresponding ranges for the high-anti conformers are $134°$–$103°$ for d4A and $116°$ to $99°$ for d4G. Thus, like d4T, d4C, and d4U (Yurenko et al., 2007, 2011), but in contrast to the canonical nucleosides, d4A and d4G can adopt the high-anti base orientation. All nine possible combinations of the $\gamma$ and $\beta$ torsion angles, which determine the orientation of the methoxy group on the 5'-end of the sugar residue are observed in the conformational families of the d4A and d4G nucleosides. Both angles present a trimodal distribution (Figure 4), falling into three sectors ($g^+$, t and $g^-$).

In addition to the distribution of torsion angles, we also examined the structural changes induced by nucleoside formation in the d4A and d4G conformers, including changes in the bond lengths and angles and changes in the structure of the base. The adenine and guanine heterocycles in the minimum-energy structures of d4A and d4G are, strictly speaking, (slightly) nonplanar: the values of the intracycle torsion angles range from .5° to 3°. It was earlier demonstrated that isolated DNA bases are rather flexible molecules (Hovorun, Gorb, & Leszczynski, 1999; Nikolaienko, Bulavin, & Hovorun, 2011), with the result that non-planar deformations of the heterocycle of the bases occur upon nucleoside formation.

Another structural peculiarity of the d4A and d4G conformers is that the glycosidic bond (C1'N9) deviates only slightly (up to 4°) from the mean plane of the imidazole ring. The mean base plane was calculated for each conformer by a least-squares fit algorithm. In addition, the values of the angles between the C2N2 bond in d4A or the C6N6 bond in d4G and the respective base planes varies from 13° to 22° for d4A and from 34° to 40° for d4G. These data indicate that the exocyclic fragments in d4A and d4G are essentially pyramidalized, as was also observed earlier for d4C (Ponomareva et al., 2012).

A statistical analysis allowed us to establish correlations between the angles $\nu_1$ and $\nu_4$ (linear correlation coefficient of .92), $\nu_1$ and $\nu_3$ (.81), $\nu_2$ and $\chi$ ($-.74$), $\nu_3$ and $\nu_4$ (.98), and between $\nu_4$ and $\nu_0$ ($-.99$).

Correlation analysis confirmed that the conformational variables $\gamma$, $\beta$, and $\chi$ do not intercorrelate (the absolute values of the corresponding correlation coefficients do not exceed .26). This is expected, as these conformational variables are basically independent.

![Figure 4. Conformation wheels for three structural parameters in d4A and d4G: dihedral angles $\chi$, $\gamma$, $\beta$.](image-url)
According to the Boltzmann distribution, the conformational equilibrium of d4A and d4G at 298.15 K is shifted toward syn conformers, with syn:anti ratios of 88:12% (d4A) and 94:6% (d4G). In this respect, d4A and d4G differ drastically from stavudine (d4T), which favors the anti-base orientation in the free state (Ponomareva et al., 2013). The distributions of the d4A and d4G sugar conformational subfamilies differ strongly from those of the corresponding canonical nucleosides (Yurenko et al., 2007, 2011). The \( \gamma \) angle preferentially adopts the \( g^+ \) conformation: \( \gamma \in g^+ \) in 94% of d4A and 98% of d4G conformers (whereas less than 1% of d4A and d4G conformers contain the \( g^- \) conformation for the \( \gamma \) angle). A similar distribution is observed for the \( \beta \) angle: \( \beta \in g^+ \) in 93% d4A and 97% d4G conformers, whereas only \( \sim 1\% \) of d4A and d4G conformers adopt the \( t \) conformation of the \( \beta \) angle.

The deoxyribose unit was shown to be the major source of the conformational flexibility of DNA (Foloppe & MacKerell, 1999; Poltev et al., 2010) and, thus, the overall DNA structure is to a large extent determined by the conformational properties of 2'-deoxyribonucleosides. Due to the tightness of the active sites of polymerases (Kim, Delaney, Essigmann, & Kool, 2005), the nucleoside triphosphate should possess a DNA-like conformation to allow incorporation into DNA. To investigate possible conformational effects on the biological activity of d4A and d4G, we optimized two structures for each nucleoside with the \( \chi \) angle fixed at the “ideal” (Foloppe & MacKerell, 1999) A- and B-DNA forms, \( -158.9^\circ \) and \( -101.9^\circ \), respectively. The results, shown in Table 4, indicate that, like d4T (Ponomareva et al., 2013), d4U and d4C (Ponomareva et al., 2012), d4A and d4G can adopt A as well as B conformations. This is an important requirement to allow their incorporation into double-helical DNA. All four “ideal” DNA-like conformations of d4A and d4G are stabilized by a C8H-..-O5' H-bond, which plays an important role in maintaining the mutual orientation of the base and sugar residues in DNA (Wahl & Sundaralingam, 1997). These results show that there are no energetic or conformational barriers for the incorporation of d4A and d4G into double-helical DNA. This confirms the earlier suggestion that the biological activity of the 2',3'-didehydro-2',3'-dideoxy analogs of the natural nucleosides is directly related to the lack of a hydroxyl group at the C3'-position of the sugar moiety (Hamamoto et al., 1987). As a result, they function as competitive RT inhibitors through termination of the DNA biosynthesis in the 5'-3' direction. We are currently performing calculations to investigate the interactions of the triphosphates of the 2',3'-didehydro-2',3'-dideoxy nucleosides with HIV-1 RT, with the aim to provide a more detailed understanding of the biological activity of d4A and d4G.

**Specific intramolecular interactions.**

In general, 21 specific intramolecular contacts involving a hydrogen atom (hereafter called H-contacts), were identified in d4A and 23 in d4G by means of QTAIM (quantum theory of atoms in molecules) electron density topological analysis (Bader, 1990), see Tables SI and SII (Supplementary Materials). Most conformers (19 d4A and 14 d4G conformers) are shaped by only one H-contact, whereas one d4A and four d4G conformers are stabilized by two H-contacts.

The specific intramolecular contacts are sensitive to the nucleoside conformation. In particular, the C8H-..-O5', C2'H-..-N3, O5'H-..-C8 and C8H-..-H1/2C5' interactions only occur in conformers with an anti-oriented base, whereas O5'H-..-N3, N2H1-..-O5', and C5' H1/2-..-N3 contacts can be formed in syn-conformers. Like in d4T, d4U, and d4C (Ponomareva et al., 2012; 2013), the orientation of \( \gamma \) determines which of the two hydrogen atoms of the C5'H group (H1 or H2, see Figure 1) is involved in H-bonding: if \( \gamma \in t \) is H1, and if \( \gamma \in g^- \) it is H2 that participates in H-bonding. All specific intramolecular contacts are formed exclusively between the base and the sugar residue, in contrast to the canonical 2'-deoxyribonucleosides, where intra-sugar H-bonds of the C2'H-..-O5' type are also observed (Yurenko et al., 2011). The N3 atom of the nucleobase is the most common hydrogen-bonding acceptor in d4A and d4G (11 interactions in d4A and 9 in d4G).

### Table 4. Energetic and conformational characteristics of the DNA-like conformers of d4A and d4G (5 and 4, respectively) and the partially optimized structures (with fixed \( \chi \) angles) that correspond to the Bi- (\( \chi = -158.9^\circ \)) and Al-(\( \chi = -101.9^\circ \)) DNA forms.

| Nucleoside | Conformer | \( \Delta G \), kcal/mol | \( \mu \), D | \( \nu \)max, degrees | \( \chi \), degrees | \( \gamma \), degrees | \( \beta \), degrees | \( \beta \), degrees | \( t \)max, degrees | \( R_{C1UN5'} \), Å |
|-----------|-----------|----------------|-----------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| d4A       | 5         | 2.49           | 4.72      | 1.7            | -120.9         | 45.4           | 163.3          | 0.6            | 1.459          |
| A-DNA     | 3.91      | 4.83           | 4.1       | -158.9         | 50.3           | 172.9          | 2.0            | 1.468          |
| B-DNA     | 2.75      | 4.42           | 1.4       | -101.9         | 46.1           | 152.4          | 1.5            | 1.458          |
| d4G       | 4         | 3.01           | 6.72      | 1.0            | -115.8         | 45.0           | 155.8          | 1.9            | 1.459          |
| A-DNA     | 4.23      | 7.46           | 4.1       | -158.9         | 49.4           | 170.4          | 2.0            | 1.468          |
| B-DNA     | 3.22      | 6.74           | 1.5       | -101.9         | 45.1           | 180.0          | 2.0            | 1.458          |

*See footnote, Table 1.

*These values have been frozen during geometry optimization.
Properties of the potential HIV-1 RT inhibitors d4A and d4G

Table 5. Vibrational and energetic characteristics of the intramolecular OH· ·N and NH· ·O H-bonds in d4A conformer 5 and d4G conformers 1 and 14.

| Nucleoside | Conformer | H-bond | Δd[^a^], Å | Δν[^b^], cm[^−1] | I/I[^c^] 0 | Δν[^d^], cm[^−1] | E[^e^] HB[^c^], e, kcal/mol |
|------------|-----------|--------|-------------|----------------|------------|----------------|----------------------------|
| d4A        | 5         | O5′H· ·N3 | .019       | −332          | 52.0       | 110            | 5.64                      |
|            |           | O5′H· ·N3 | .021       | −425          | 52.7       | 148            | 6.48                      |
|            |           | N2H1· ·O5′ | .004       | −142          | 4.8        | 79             | 3.33                      |
| d4G        | 1         | O5′H· ·N3 | .017       | −245          | 14.5       | 98             | 4.72                      |
|            | 14        | O5′H· ·N3 | .017       | −245          | 14.5       | 98             | 4.72                      |
|            |           | N2H1· ·O5′ | .004       | −127          | 3.2        | 68             | 3.08                      |

[^a^] Difference between the OH/NH bond length in the OH(NH)· ·O/N H-bond and in the nonbonded state.
[^b^] Decrease in the OH/NH stretching mode frequency upon H-bonding.
[^c^] Increase in the OH/NH out-of-plane mode frequency upon H-bonding.
[^d^] H-bond energy (in kcal/mol), calculated by logansen’s empirical formula (Equation [3] in Supplementary Materials).

The specific intramolecular interactions in d4A and d4G can be divided into three categories: conventional (OH· ·N, NH· ·O) H-bonds, non-conventional (CH· ·O, OH· ·C, CH· ·N) H-bonds, and hydrogen–hydrogen contacts (CH· ·HC) (Bakhmutov, 2008; Matta, 2006; Matta, Hernández-Trujillo, Tang, & Bader, 2003). The conventional OH· ·N and NH· ·O H-bonds in d4A and d4G possess distinct spectral characteristics (Table 5), including a perceptible red shift of the stretching mode frequency ν (OH/NH), an increase in the stretching mode intensity and an increase in the out-of-plane γ (OH/NH) vibrational frequency. Based on these results, the energy of the conventional H-bonds was evaluated by means of logansen’s empirical formula (Equation [3] in Supplementary Materials) to range from 3.08 to 6.48 kcal/mol.

The magnitude of the electron density ρ at the bond critical points of the noncovalent interactions in d4A and d4G (Tables SI and SII, Supplementary Materials) are in good agreement with the H-bonding criteria of (Koch & Popelier, 1995). All H-contacts, except the CH· ·HC and O5′H· ·C8 interactions, meet Bondi’s geometric requirements for H-bonding (Bondi, 1964), in particular the requirement that the distance between the donor and acceptor atoms should be smaller than the sum of the van der Waals atomic radii, Σ[^v^]dw (2.72 Å for H· ·O and H· ·N and 2.2 Å for H· ·H). In general, there is a good correlation between the Σ[^v^]dw values and the H· ·B distance (where B is the acceptor atom), with a correlation coefficient of .988.

In this work, we used Grunenberg’s compliance constant theory (Brandhorst & Grunenberg, 2008, 2010; Grunenberg, 2004) to evaluate the individual strengths of the intramolecular interactions. Compliance constants (C[^i^][^j^]) are elements of the inverse force constant matrix. Unlike the elements of the regular force constant (Hessian) matrix, they depend only on the definition of the two internal coordinates pertaining to it and are independent of the definition of all other coordinates. Tables SI and SII, and Figure S1 (Supplementary Materials) clearly show that smaller values of C[^i^][^j^] correspond to stronger contacts, as quantified by the energies of the specific interactions E[^HB^] calculated by the Espinosa-Molins-Lecomte formula (Equation (1) in Supplementary Materials). The correlation is evidently not linear, as was assumed in our previous work (Ponomareva et al., 2012, 2013). The solid line in Figure S1 corresponds to a least-squares power fit (C[^i^][^j^] = 53.365 × E[^HB^]^{−1.2285}, R = −.93). Some of the CH· ·HC contacts are characterized by very high C[^i^][^j^] values (> 50 Å/mDyn), indicating the relatively weak nature of these contacts.

The CH· ·HC contacts in d4A and d4G deserve special attention. These H-bonds were earlier detected in canonical nucleosides (Yurenko et al., 2007) and some amino acids (Matta & Bader, 2000) by means of Bader’s topological analysis (Bader, 1990). To find the H-bond donating (donor) and accepting (acceptor) groups in these interactions (assuming these are genuine H-bonds (Matta, 2006; Matta et al., 2003)), an integration (in natural coordinates) over Bader atomic basins was performed for each hydrogen atom (Tables SV and SVI in Supplementary Materials). The hydrogen on the base moiety bears a partial positive charge and has less negative atomic energy than the hydrogen of the sugar residue, which is negatively charged and more energetically stable. (Note that the charges are calculated with the QTAIM method.) Thus, the H-bond donating group is located on the sugar residue, whereas the acceptor is on the base unit. This is in agreement with the results for d4T (Ponomareva et al., 2013), d4U, and d4C (Ponomareva et al., 2012).

Conclusions
The conformational landscapes of two HIV-1 RT inhibitors, the nucleoside analogs 2′,3′-didehydro-2′, 3′-dideoxyadenosine (d4A), and 2′,3′-didehydro-2′,3′-dideoxyguanosine (d4G) were explored using a comprehensive set of modern computational techniques (DFT and MP2 calculations, Bader’s electron density topological analysis, Natural Bond Orbital (NBO) analysis, and compliance constant theory).
We identified 21 d4A and 20 d4G conformers. The conformers are shaped by a sophisticated network of specific interactions of various types (NH···O, OH···N, CH···O, OH···C, and CH···HC). The relative strengths of these interactions can be efficiently predicted using Grunenberg’s compliance constants theory (Brandhorst & Grunenberg, 2008, 2010; Grunenberg, 2004). NBO analysis (Supplementary Materials) revealed the electronic nature of the noncovalent interactions. For the OH···O, NH···O, and CH···O H-bonds hyperconjugative interactions between the oxygen lone pairs and the antibonding orbital of the donor group are observed. The electron density of the CH···HC contacts migrates from the antibonding orbital, corresponding to the CH group of the sugar residue, to the bonding orbital relative to the same group in the nucleobase.

The results indicate that there are no energetic or conformational obstacles to the incorporation of d4A and d4G into double-helical DNA. This provides further support for the suggestion that the biological activity of 2′,3′-dideoxy-2′,3′-dideoxy analogs is directly related to the lack of a hydroxyl group at the C3′-position of the sugar moiety (Hamamoto et al., 1987). As a result, these nucleosides act as competitive RT inhibitors through termination of the DNA biosynthesis in the 5′-3′ direction.

The biological relevance of the results presented in this work lies in the considerable flexibility of the 2′,3′-dideoxy-2′,3′-dideoxy-nucleosides; as biological function of these nucleosides is structure-dependent, a conformational analysis is the initial, but very important, step in elucidating their biological activity. A detailed knowledge of the full conformational families of these nucleosides is essential for establishing the biologically relevant conformations that are formed in the course of biomolecular recognition by different enzymes (transporters, kinases/kinase, and polymerases). In addition, we have identified and characterized among all the conformers the one that can be introduced in the active site and potentially block the DNA biosynthesis. A study of the interactions of the triphosphates of these nucleosides with HIV-1 RT could provide a more detailed understanding of the biological activity of d4A and d4G. Such calculations are currently being performed in our research groups.

**Acknowledgements**

The authors greatly appreciate useful discussion of the results with Prof. M. Gomis and Dr B. Hernandez (Université Paris 13). Y.P.Y. is deeply grateful to Gaussian Inc. for granting a G03 molecular modeling package for the Win32 platform. This work was partly supported by the State Fund for Fundamental Research of Ukraine within Ukrainian-Slovenian bilateral research project (No M/451-2011).

**Supplementary material**

The supplementary material for this paper is available online at http://dx.doi.org/10.1080/07391102.2013.789401.

**References**

Alber, F., & Carloni, P. (2000). Ab initio molecular dynamics studies on HIV-1 reverse transcriptase triphosphate binding site: Implications for nucleoside-analog drug resistance. *Protein Science, 9*, 2535–2546.

Altona, C., & Sundaralingam, M. (1972). Conformational analysis of the sugar ring in nucleosides and nucleotides. New description using the concept of pseudorotation. *Journal of the American Chemical Society, 94*, 8205–8212.

Bader, R. F. W. (1990). *Atoms in molecules: A quantum theory*. Oxford: Clarendon.

Bakhmutov, V. I. (2008). *Dihydrogen bond: Principles, experiments and applications*. Hoboken, NJ: John Wiley & Sons.

Balzarini, J., Kruining, J., Wedgwood, O., Pannecoque, C., Aguaro, S., Perno, C. F., ... McGuigan, C. (1997). Conversion of 2′,3′-dideoxyadenosine (dDA) and 2′,3′-dideoxy-2′,3′-dideoxyadenosine (d4A) to their corresponding aryloxyporphoramide derivatives markedly potentiates their activity against human immunodeficiency virus and hepatitis B virus. *FEBS Letters, 410*, 324–328.

Becke, A. D. (1993). Density functional thermochemistry. 3. The role of exact exchange. *Journal of Chemical Physics, 98*, 5648–5652.

Bondi, A. (1964). Van der Waals volumes and radii. *Journal of Physical Chemistry, 68*, 441–451.

Brandhorst, K., & Grunenberg, J. (2008). How strong is it? The interpretation of force and compliance constants as bond strength descriptors. *Chemical Society Reviews, 37*, 1558–1567.

Brandhorst, K., & Grunenberg, J. (2010). Efficient computation of compliance matrices in redundant internal coordinates from Cartesian Hessians for nonstationary points. *Journal of Chemical Physics, 132*, 184101–184107.

Chu, C. K., Schinazi, R. F., Arnold, B. H., Cannon, D. L., Doboszewski, B., Bhadti, V. B., & Gu, Z. P. (1988). Comparative activity of 2′,3′-saturated and unsaturated pyrimidine and purine nucleosides against human immunodeficiency virus type-1 in peripheral-blood mononuclear cells. *Biochemical Pharmacology, 37*, 3543–3548.

el Kouni, M. H. (2002). Trends in the design of nucleoside analogues as anti-HIV drugs. *Current Pharmaceutical Design, 8*, 581–593.

Everaert, D. H., Peeters, O. M., Deranter, C. J., Blaton, N. M., Vanaerschot, A., & Herdewijn, P. (1993). Conformational analysis of substituent effects on the sugar puckering mode of cytosine in the equilibrium among the A, B, and Z conformation. *Journal of the American Chemical Society, 94*, 8205–8212.

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., ... Pople, J. A. (2003). *Gaussian 03*. Pittsburgh, PA: Gaussian.

Gelbin, A., Schneider, B., Clooney, L., Hsieh, S.-H., Olson, W. K., & Berman, H. M. (1996). Geometric parameters in nucleic acids: Sugar and phosphate constituents. *Journal of the American Chemical Society, 118*, 519–529.
Properties of the potential HIV-1 RT inhibitors d4A and d4G 739

Grunenberg, J. (2004). Direct assessment of interresidue forces in Watson-Crick base pairs using theoretical compliance constants. Journal of the American Chemical Society, 126, 16310–16311.

Hamamoto, Y., Nakashima, H., Matsui, T., Matsuda, A., Ueda, T., & Yamamoto, N. (1987). Inhibitory effect of 2',3'-didehydro-2',3'-dideoxynucleosides on infectivity, cytopathic effects, and replication of human-immunodeficiency-virus. Antimicrobial Agents and Chemotherapy, 31, 907–910.

Hartmann, B., & Lavery, R. (1996). DNA structural forms. Quarterly Reviews of Biophysics, 29, 309–368.

Herdewijn, P. (Ed.). (2008). Characterization of C.

Koch, U., & Popelier, P. L. A. (1995). Stereoelectronic effects in nucleosides and nucleotides and their structural implications. Uppsala: Uppsala University Press.

Matsuya, H., Yarchoan, R., Kageyama, S., & Broder, S. (1991). Targeted therapy of human immunodeficiency virus-related disease. Faseb Journal, 5, 2369–2381.

Nikolaenko, T. Y., Bulavin, L. A., & Hovorun, D. M. (2011). How flexible are DNA constituents? The quantum-mechanical study. Journal of Biomolecular Structure & Dynamics, 29, 563–575.

Palafox, M. A., Iza, N., de la Fuente, M., & Navarro, R. (2009). Simulation of the first hydration shell of nucleosides d4T and thymidine: Structures obtained using MP2 and DFT methods. Journal of Physical Chemistry B, 113, 2458–2476.

Poltev, V. I., Anisimov, V. M., Danilov, V. I., van Mourik, T., Deriabina, A., Gonzalez, E., … Polteva, N. (2010). DFT study of polymorphism of the DNA double helix at the level of dinucleoside monophosphates. International Journal of Quantum Chemistry, 110, 2548–2559.

Polomareva, A. G., Yureno, Y. P., Zhuravkivsky, R. O., van Mourik, T., & Hovorun, D. M. (2012). Complete conformational space of the potential HIV-1 reverse transcriptase inhibitors d4U and d4C. A quantum chemical study. Physical Chemistry Chemical Physics: PCCP, 14, 6787–6795.

Rao, S. N. (1998a). Conformational studies on nucleosides with furanose ring modifications. 1. Nucleosides and Nucleotides, 17, 791–814.

Rao, S. N. (1998b). Molecular mechanics studies on the conformations of 2',3'-dideoxy-2',3'-dideoxyguanosine nucleoside, d4G. Biophysics Journal, 74, 3131–3139.

Ray, A. S., Yang, Z. J., Chu, C. K., & Anderson, K. S. (2002). Novel use of a guanosine prodrug approach to convert 2',3'-didehydro-2',3'-dideoxyguanosine into a viable antiviral agent. Antimicrobial Agents and Chemotherapy, 46, 887–891.

Russell, J. W., Whitcroft, V. J., Marrero, D., & Klunk, L. J. (1989). Pharmacokinetics of a new anti-HIV agent: 2',3'-dideoxy-2',3'-dideoxythymidine (d4T). Nucleosides and Nucleotides, 8, 845–848.

Saenger, W. (1984). Principles of nucleic acid structure. New York, NY: Springer-Verlag.

Sahlberg, C., & Zhou, X-X. (2008). Development of non-nucleoside reverse transcriptase inhibitors for anti-HIV therapy. Anti- Infective Agents in Medicinal Chemistry, 7, 101–117.

Skowron, G., & Ogden, R. (2006). Reverse transcriptase inhibitors in HIV/AIDS therapy. Totowa, NJ: Humana Press. p. 527.

Slius-Cremer, N., Temiz, N. A., & Bahar, I. (2004). Conformational changes in HIV-1 reverse transcriptase induced by nonnucleoside reverse transcriptase inhibitor binding. Current HIV Research, 2, 323–332.

Thibaudeau, C., Acharya, P., & Chattopadhyaya, J. (2005). Stereoelectronic effects in nucleosides and nucleotides and their structural implications. Uppsala: Uppsala University Press.

van Roey, P., & Chu, C. K. (1992). The crystal and molecular-structure of the complex of 2',3'-didehydro-2',3'-dideoxyguanosine with pyridine. Nucleosides & Nucleotides, 11, 1229–1239.

van Roey, P., & Chu, C. K. (1992). The crystal and molecular-structure of the complex of 2',3'-didehydro-2',3'-dideoxyguanosine with pyridine. Nucleosides & Nucleotides, 11, 1229–1239.
Van Roey, P., Taylor, E. W., Chu, C. K., & Schinazi, R. F. (1993). Conformational analysis of 2′,3′-didehydro-2′, 3′-dideoxypuridine nucleosides. Journal of the American Chemical Society, 115, 5365–5371.

Van Rompay, A. R., Johansson, M., & Karlsson, A. (2003). Substrate specificity and phosphorylation of antiviral and anticancer nucleoside analogues by human deoxyribonucleoside kinases and ribonucleoside kinases. Pharmacology & Therapeutics, 100, 119–139.

Wahl, M. C., & Sundaralingam, M. (1997). C–H...O hydrogen bonding in biology. Trends in Biochemical Sciences, 22, 97–102.

Yurenko, Y. P., Zhurakivsky, R. O., Ghomi, M., Samijlenko, S. P., & Hovorun, D. M. (2007a). Comprehensive conformational analysis of the nucleoside analogue 2′-β-deoxy-6-aza-cytidine by DFT and MP2 calculations. Journal of Physical Chemistry B, 111, 6263–6271.

Yurenko, Y. P., Zhurakivsky, R. O., Ghomi, M., Samijlenko, S. P., & Hovorun, D. M. (2007b). How many conformers determine the thymidine low-temperature matrix infrared spectrum? DFT and MP2 quantum chemical study. Journal of Physical Chemistry B, 111, 9655–9663.

Yurenko, Y. P., Zhurakivsky, R. O., Ghomi, M., Samijlenko, S. P., & Hovorun, D. M. (2008). Ab initio comprehensive conformational analysis of 2′-deoxyuridine, the biologically significant DNA minor nucleoside, and reconstruction of its low-temperature matrix infrared spectrum. Journal of Physical Chemistry B, 112, 1240–1250.

Yurenko, Y. P., Zhurakivsky, R. O., & Hovorun, D. M. (2009). Intermolecular C-H...O hydrogen bonds in biologically significant conformers of canonical 2-deoxyribonucleotides: Ab initio topological analysis of the electron density. Physics Alive (in Ukrainian), 17, 44–53.

Yurenko, Y. P., Zhurakivsky, R. O., Samijlenko, S. P., Ghomi, M., & Hovorun, D. M. (2007). The whole of intramolecular H-bonding in the isolated DNA nucleoside thymidine. AIM electron density topological study. Chemical Physics Letters, 447, 140–146.

Yurenko, Y. P., Zhurakivsky, R. O., Samijlenko, S. P., & Hovorun, D. M. (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. Journal of Biomolecular Structure & Dynamics, 29, 51–65.

Zhou, P.-P., & Qiu, W.-Y. (2009). Red-shifted hydrogen bonds and blue-shifted van der Waals contact in the standard Watson–Crick adenine–thymine base pair. Journal of Physical Chemistry A, 113, 10306–10320.

Zhurakivsky, R. O., & Hovorun, D. M. (2006). Extensive conformational analysis of canonical nucleoside 2′-deoxyxycytidine: A DFT study. Physics Alive (in Ukrainian), 14, 33–46.