

Hybrid MC/QC simulations of water-assisted proton transfer in nucleosides. Guanosine and its analog acyclovir

Nadezhda Markova, Ljupco Pejov, Nina Stoyanova & Venelin Enchev

To cite this article: Nadezhda Markova, Ljupco Pejov, Nina Stoyanova & Venelin Enchev (2017) Hybrid MC/QC simulations of water-assisted proton transfer in nucleosides. Guanosine and its analog acyclovir, Journal of Biomolecular Structure and Dynamics, 35:6, 1168-1188, DOI: [10.1080/07391102.2016.1179594](https://doi.org/10.1080/07391102.2016.1179594)

To link to this article: <http://dx.doi.org/10.1080/07391102.2016.1179594>



Accepted author version posted online: 19 Apr 2016.
Published online: 20 May 2016.



Submit your article to this journal [↗](#)



Article views: 37



View related articles [↗](#)



View Crossmark data [↗](#)

Hybrid MC/QC simulations of water-assisted proton transfer in nucleosides. Guanosine and its analog acyclovir

Nadezhda Markova^a, Ljupco Pejov^{b,c}, Nina Stoyanova^a and Venelin Enchev^{a*}

^aInstitute of Organic Chemistry, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria; ^bFaculty of Natural Sciences and Mathematics, Institute of Chemistry, Skopje, Macedonia; ^cResearch Centre for Environment and Materials, Macedonian Academy of Sciences and Arts, Krste Misirkov 2, 1000 Skopje, Macedonia

Communicated by Ramaswamy H. Sarma

(Received 27 February 2016; accepted 13 April 2016)

To provide an in-depth insight into the molecular basis of spontaneous tautomerism in DNA and RNA base pairs, a hybrid Monte Carlo (MC)–quantum chemical (QC) methodology is implemented to map two-dimensional potential energy surfaces along the reaction coordinates of solvent-assisted proton transfer processes in guanosine and its analog acyclovir in aqueous solution. The solvent effects were simulated by explicit inclusion of water molecules that model the relevant part of the first hydration shell around the solute. The position of these water molecules was estimated by carrying out a classical Metropolis Monte Carlo simulation of dilute water solutions of the guanosine (Gs) and acyclovir (ACV) and subsequently analyzing solute–solvent intermolecular interactions in the statistically-independent MC-generated configurations. The solvent-assisted proton transfer processes were further investigated using two different *ab initio* MP2 quantum chemical approaches. In the first one, potential energy surfaces of the ‘bare’ finite solute–solvent clusters containing Gs/ACV and four water molecules (MP2/6-31+G(d,p) level) were explored, while within the second approach, these clusters were embedded in ‘bulk’ solvent treated as polarizable continuum (C-PCM/MP2/6-31+G(d,p) level of theory). It was found that in the gas phase and in water solution, the most stable tautomer for guanosine and acyclovir is the 1*H*-2-amino-6-oxo form followed by the 2-amino-6-(*sZ*)-hydroxy form. The energy barriers of the water-assisted proton transfer reaction in guanosine and in acyclovir are found to be very similar – 11.74 kcal mol^{−1} for guanosine and 11.16 kcal mol^{−1} for acyclovir, and the respective rate constants ($k = 1.5 \times 10^1$ s^{−1}, guanosine and $k = 4.09 \times 10^1$ s^{−1}, acyclovir), are sufficiently large to generate the 2-amino-6-(*sZ*)-hydroxy tautomer. The analysis of the reaction profiles in both compounds shows that the proton transfer processes occur through the asynchronous concerted mechanism.

Keywords: nucleosides; guanosine; acyclovir; proton transfer; tautomerism; *ab initio*; Monte Carlo simulation

1. Introduction

Studies of tautomeric equilibrium in structural components of DNA and RNA are of great interest today (Bebenek, Pedersen, & Kunkel, 2011; Brovarets & Hovorun, 2015a, 2015b, 2015c, 2015d; Brovarets, Zhurakivsky, & Hovorun, 2014; Harris et al., 2003; Shukla & Leszczynski, 2013; Topal & Fresco, 1976; Wang, Hellinga, & Beese, 2011). The biological importance of guanines as part of the genetic information system has been acknowledged ever since the description of the DNA helical structure by Watson and Crick (Watson & Crick, 1953). The fidelity of the replication and translation of the genetic material can be opposed by the occurrence of wobble base pairs or of rare tautomeric forms of the nucleic acid bases, and these phenomena can induce point mutations. It must be emphasized that in all the nucleoside and nucleotide crystal structures published thus far, only the canonic tautomeric forms of bases have been observed. Calcula-

tions and spectroscopic data demonstrated that very small amounts of other tautomeric forms can exist in solution (Saenger, 1984).

However, several tautomeric forms of the purine ring can be theoretically formulated for both guanine (Gorb & Leszczynski, 1997, 1998) and guanosine (Guo, Xue, Ke, Phillips, & Zhao, 2009; Thewalt, Bugg, & Marsh, 1970). To the best of our knowledge, the possibility of tautomeric conversion in 2'-deoxyguanosine has been investigated (Kochina, Zhurakivsky, & Hovorun, 2008) but not in guanosine.

Guanosine (Gs), the guanine nucleoside is a major component of most types of RNA. From the IR spectra of the mono- and dihydrated clusters of Gs (Saigusa, Mizuno, Asami, Takahashi, & Tachikawa, 2008), it has been found that multiple structural isomers exist in both mono- and dihydrates of Gs and that the internal hydrogen-bonding structure of the Gs monomer is retained in all hydrates. In addition, the amino-keto forms of Gs,

*Corresponding author. Email: venelin@orgchm.bas.bg

which are missing in the R2PI spectrum, can be observed upon hydration (Saigusa, Urashima, & Asami, 2009). The dynamics of crystal water molecules of guanosine dihydrate and guanosine monophosphate system are investigated in detail by molecular dynamics simulation (Yoneda, Sugawara, & Urabe, 2005).

2-Amino-1,9-dihydro-9-((2-hydroxyethoxyl)methyl)-6H-purin-6-one or acyclovir, (**ACV**), is a guanosine analog antiviral drug (Avcioglu & Golcu, 2015; Jiang, Feng, Lin, & Guo, 2016). **ACV** has been extensively studied from the pharmaceutical and medical point of view. However, only a few studies related to its molecular structure have been published employing lower levels of theory. Gavira et al. have reported the normal coordinate analysis of acycloguanosine (Gavira, De La Fuente, Navarro, & Hernanz, 1997). Conformations of some derivatives of **ACV** with biological activities have been reported (Birnbbaum, Cygler, & Shugar, 1984), as well as studies of complexes of **ACV** with several metals (Barceló-Oliver et al., 2004; Brandi-Blanco et al., 2011). Using a range of techniques (X-ray diffraction, solid-state NMR, vibrational, Raman and IR spectroscopy) Lutker et al. (Lutker, Quiñones, Xu, Ramamoorthy, & Matzger, 2011) have shown that **ACV** has several forms ranging from hydrates to anhydrates.

The tautomeric equilibrium in **ACV** involving the keto and enol forms has been observed from the UV/Vis spectra (Plass, Kristl, & Abraham, 1999). This equilibrium depends on the polarity of the solvent, and therefore, in water solution, the keto form prevails, while in methylene chlorides it is the enol one (Alvarez-Ros & Alcolea Palafox, 2014; Plass et al., 1999).

Therefore, the solvent appears to influence the tautomeric conversion to a great extent. This is understandable, since it is the major component of most chemical systems and of all the systems of biochemical interest. On average, each biochemical compound is surrounded by thousands of water molecules; that is, the water is the most abundant biological molecule. Therefore, knowledge of the properties of water and its role as the surrounding medium in which biochemical processes take place is essential for understanding the biology of living beings. As a solvent, water has two different effects on a chemical system. First, water has a relevant influence on the dynamics of chemical structures. The second effect of water concerns its influence on the potential energy of the chemical species in solution. Water can modify the rate of a chemical reaction by specific stabilization or destabilization of the transition state (Orozco, Cubero, Barril, Colominas, & Luque, 1999).

On the other hand, water solubility is an important molecular property for successful drug development as it is a key factor governing drug access to biological membranes. According to Faller and Ertl (Faller & Ertl, 2007), the main reason for inaccurate calculation of

hydrophobicity parameter (logP) is the use of incorrect tautomeric forms of structures in the calculations. Because of the ‘tautomeric problem,’ some of drug-like structures as **ACV** may exist in dozen of tautomeric forms, energetically quite close. In some cases, the form present in the database (and therefore used for calculation of properties) does not correspond to the form that is actually present in the test tube.

The main aim of the present study is to model the water-assisted proton transfer in the title systems. In realistic media, especially relevant to biomedical sciences, such process occurs in water solutions, with a substantial unavoidable influence of the solvent. Modeling of in-liquid processes of such complexity, under realistic conditions, is far from a trivial task if one wants to account properly for all essential features of the solvent influence. In the absence of specific solvent–solute intermolecular interactions, it is possible to build up a realistic model by treating the solvent as a polarizable continuum, at different levels of sophistication of this methodology (Cramer & Truhlar, 2008; Tomasi, Mennucci, & Cammi, 2005). When specific noncovalent solute–solvent interactions (such as hydrogen bonding) are present, however, purely electrostatic representation of the solvent does not account for all essential specificities of the system and the process. One way to solve this problem is to in a sense ‘cut out’ a relevant part of the solvent, i.e., represent a solute–solvent subsystem as a finite cluster consisting of one solute and several solvent molecules, and account for the remaining part of the solvent by the PCM approach. Still, however, the way in which the ‘relevant’ part of the solvent is chosen within the finite cluster representation of the solute–solvent complex system is to some extent ambiguous, though ‘chemical intuition’ and other argumentations may certainly be applied to build up a realistic cluster description of the studied system. The previous argument is of particular importance when one wants to model the system under realistic conditions at which it actually occurs in biologically important media (i.e. at finite temperature, usually quite far from absolute zero, at which PES-based *ab initio* calculations are usually carried out). To provide a physically justifiable basis to augment the ‘chemical-intuition’-based choice of solute–solvent clusters to study the water-assisted proton transfer processes in the current study, we implement a hybrid statistical physics–quantum mechanical approach (Markova, Pejov, & Enchev, 2015). This methodology consists of the following. The studied solvent molecule is modeled by a statistical physics approach in explicitly described solvent under realistic physical conditions. Subsequently, the network of specific solute–solvent intermolecular interactions is analyzed and the statistics of involvement of each atomic center within the solute molecule in such interactions is quantitatively described. Subsequently, with the aid of this

statistics, chemical intuition, and the knowledge concerning the relevant part of the solute molecule in which the proton-transfer occurs, a realistic finite-cluster model is being built up.

2. Computational details

2.1. Statistical physics simulations

In the present study, we use Metropolis Monte Carlo (MC) technique to carry out the statistical physics simulation of the systems of interest. MC simulations were carried out in *NPT* ensemble at temperature of 298 K and pressure of 1 atmosphere. Four MC simulations have actually been carried out. In each simulation, the particular tautomer was surrounded by 500 water molecules in a cubic box, imposing periodic boundary conditions (PBC). The side length of the cubic box was approximately 25 Å, as determined from the experimental value of the water density at these particular experimental conditions. Since PBC have been imposed, the long-range corrections to the interaction energy were calculated in the case of all interacting pairs of atoms between which the distance is larger than the cut-off radius, which was defined as half of the unit cell length. Intermolecular interaction potentials were of the Lennard-Jones 12-6 + Coulomb type, i.e.:

$$U_{ab} = \sum_i^a \sum_j^b 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (1)$$

In the previous equation, i and j are sites in interacting molecular systems a and b , r_{ij} is the interatomic distance between sites i and j , and all other symbols have the usual meanings. The ‘geometric mean’ combination rules were used to generate two-site Lennard-Jones parameters ϵ_{ij} and σ_{ij} from the single-site ones:

$$\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j} \quad (2)$$

$$\sigma_{ij} = \sqrt{\sigma_i \sigma_j} \quad (3)$$

To describe the solvent (water) molecules, the SPC model potential parameters were used. For organic solute molecules, the Lennard-Jones parameters were taken from the OPLS force field database (Jorgensen, Maxwell, & Tirado-Rives, 1996), while the ‘atomic charges’ used subsequently in the ‘Coulomb’-part of the potential were calculated by fitting the sets of point charges, placed at the atomic positions, to the molecular electrostatic potential calculated from the MP2/6-31+G(d,p) wave function at series of points generated by the CHelpG algorithm (Breneman & Wiberg, 1990); additional constraint was, however, imposed in the course of fitting procedure, that the molecular dipole moment is correctly reproduced by the chosen set of point charges representing the continual

molecular charge distribution. Note that in these calculations, the MP2 electronic density has been used as the reference in all cases.

Lennard-Jones contribution to the interaction energy for distances larger than the previously defined cutoff radius was estimated under the assumption of a uniform density distribution in the liquid ($g(r) \approx 1$). Electrostatic reaction field method involving the dipolar interactions was used to estimate the electrostatic contribution.

All MC simulations in the present paper have been carried out with the DICE code (Coutinho & Canuto, 1997). The thermalization phase in all cases consisted of at least 5×10^7 MC steps; it was followed by averaging (simulation) phase of at least 45×10^7 MC steps.

2.2. Ab initio calculations

The geometries and normal mode vibrational frequencies of the tautomers, and of the transition states were computed at the Møller–Plesset level using 6-31+G(d,p) basis set (Franci et al., 1982; Hariharan & Pople, 1973). Vibration frequency calculations were performed numerically to obtain vibrational zero point and thermal energies and to validate that the found structures corresponded to the energy minima or transition states. For the optimized transition state, one single imaginary frequency was found in the diagonalized mass-weighted Hessian matrix, and the corresponding vibrational mode was confirmed to determine the reaction coordinate. All transition structures were checked by intrinsic reaction coordinate (IRC) calculations. Starting from the transition state, the reaction path was generated as the steepest descent path in mass-scaled coordinates (intrinsic reaction coordinate, IRC) using the Gonzalez–Schlegel algorithm (Gonzales & Schlegel, 1989), employing a step size of 0.05 Bohr (1 Bohr corresponds to 0.53 Å). On both branches of the reaction coordinate, 50 steps were performed.

Solvent effect was accounted by using the self-consistent reaction field method with the conductor polarizable continuum model (C-PCM) formalism (Cossi, Rega, Scalmani, & Barone, 2003). In this model, the molecule is embedded in a cavity surrounded by an infinite dielectric which approximates the solvent as a structureless polarizable continuum characterized by its macroscopic dielectric permittivity, ϵ . All of the stationary-point geometries were fully optimized again in the reaction field of the implicit solvent. Solution-phase vibrational frequencies were computed for each stationary point on the basis of numerical Hessians. The minima and transition states were confirmed to have zero and one imaginary frequency, respectively.

The Gibbs free energy, G , for all structures was obtained as sum:

$$G = E_t + E_{\text{corr}} \quad (4)$$

where E_t is the total (electronic + nuclear) energy and E_{corr} – thermal correction. The values of Gibbs free energies (ΔG) and activation barriers (ΔG^\ddagger) were calculated at temperature 298.15 K.

The values of the populations (p_i) of the tautomeric forms were calculated by:

$$p_i = e^{-\Delta G_i/RT} / \sum_i e^{-\Delta G_i/RT} \quad (5)$$

The classical rate constant of the forward (k_f) and the reverse (k_r) tautomerization reaction was obtained using the Eyring equation:

$$k_{f,r} = \frac{k_B T}{h} e^{-\frac{\Delta G_{f,r}^\ddagger}{RT}} \quad (6)$$

where k_B , h and R are the Boltzmann, Planck and universal gas constants, respectively; $\Delta G_{f,r}^\ddagger$ is Gibbs free energy of activation for the tautomerization reaction in the forward (f) and reverse (r) directions and $T = 298.15$ K. The equilibrium constant K_T was calculated by:

$$K_T = \frac{k_f}{k_r} \quad (7)$$

The time $\tau_{99.9\%}$ necessary to reach 99.9% of the equilibrium concentration of the both tautomers in the system of reversible forward (k_f) and reverse (k_r) reactions can be calculated by the formula:

$$\tau_{99.9\%} = \frac{\ln 10^3}{k_f + k_r} \quad (8)$$

The program package GAMESS (Gordon & Schmidt, 2005; Schmidt et al., 1993) was used to perform the present *ab initio* calculations.

3. Results and discussions

The possible tautomeric forms and their rotamers for guanosine are shown in Figure 1, while Figure 2 shows those for acyclovir. Both compounds can exist in six tautomeric forms: 2 amino-oxo (**A** and **E**), 1 amino-hydroxy (**B**), 2 imino-hydroxy (**C** and **D**) and 1 imino-oxo (**F**). The amino-oxo tautomers, 1*H*-2-amino-6-oxo (**A**) and 3*H*-2-amino-6-oxo (**E**) do not have rotameric forms. The rotamers of the all rest tautomers differentiate between *E/Z* (imino groups) or between *sE/sZ* (hydroxy groups). Two rotamers are possible for the amino-hydroxy tautomer: 2-amino-6-(*sZ*)-hydroxy (**B1**) and 2-amino-6-(*sE*)-hydroxy (**B2**) as well as for the imino-oxo form: 1,3*H*-2-(*Z*)imino-6-oxo (**F1**) and 1,3*H*-2-(*E*)imino-6-oxo (**F2**). The imino-hydroxy tautomers can exist in four

rotameric forms as follows: 1*H*-2-(*Z*)imino-6-(*sE*)-hydroxy (**C1**), 1*H*-2-(*Z*)imino-6-(*sZ*)-hydroxy (**C2**), 1*H*-2-(*E*)imino-6-(*sE*)-hydroxy (**C3**) and 1*H*-2-(*Z*)imino-6-(*sE*)-hydroxy (**C4**) for tautomer **C**, and 3*H*-2-(*Z*)imino-6-(*sE*)-hydroxy (**D1**), 3*H*-2-(*Z*)imino-6-(*sZ*)-hydroxy (**D2**), 3*H*-2-(*E*)imino-6-(*sE*)-hydroxy (**D3**) and 3*H*-2-(*Z*)imino-6-(*sE*)-hydroxy (**D4**) for tautomer **D**.

All species, presented in Figures 1 and 2, were optimized at MP2/6-31+G(d,p) level. The relative Gibbs free energies of the different tautomers and rotamers of the **Gs** and **ACV** in the gas phase and in water solution where the solvent is considered as a structureless polarizable continuum characterized by its macroscopic dielectric permittivity (C-PCM) are presented in Table 1.

4. Guanosine

The computed energies of the guanosine tautomers in the gas phase reveal that the 2-amino-6-(*sZ*)-hydroxy tautomer **Gs-B1** and the 1*H*-2-amino-6-oxo tautomer **Gs-A** are almost isoenergetic but the **Gs-B1** tautomer is more stable by 0.26 kcal mol⁻¹. The energy difference between the rotamers 2-amino-6-(*sZ*)-hydroxy (**Gs-B1**) and 2-amino-6-(*sE*)-hydroxy (**Gs-B2**) is the same as the difference between **Gs-B1** and **Gs-A**. The quantities of these three species, according to their calculated relative stabilities (Equation 5), amount to 43.26% for **Gs-B1** and 27.89 and 28.85% for **Gs-A** and **Gs-B2**, respectively (Table 1).

However, in water solution, where the solvent is represented as a structureless polarizable continuum, the tautomer **Gs-A** is predicted to become most stable and the energy difference between **Gs-A** and **Gs-B1** is calculated to be 4.42 kcal mol⁻¹. The calculated fraction of **Gs-A** increases to 99.93%, while the fraction of **Gs-B1** decreases to 0.06%.

A different conformations of the ribose ring in guanosine tautomer **Gs-A** have been investigated in (Zhurakivsky & Hovorun, 2007). The conformation of the sugar residue considered in our study is closer to the experimental observed in native RNA. In the solid state, guanosine exists in 1*H*-2-amino-6-oxo tautomeric form **Gs-A** (Thewalt et al., 1970). The calculated and available experimental bond distances for **Gs-A** are shown in Table 2. The MP2/6-31+G(d,p) calculations predict C–C bond lengths longer by 0.021 Å than the experimental ones in the guanine moiety. The C3=O10 double bond is 1.237 Å, slightly longer (0.009 Å) than the calculated one. Larger deviation is observed for C5–N7 bond, where the calculated distance is by 0.046 Å longer. The same deviation is found for C3–N4 bond in the six-membered ring of guanine.

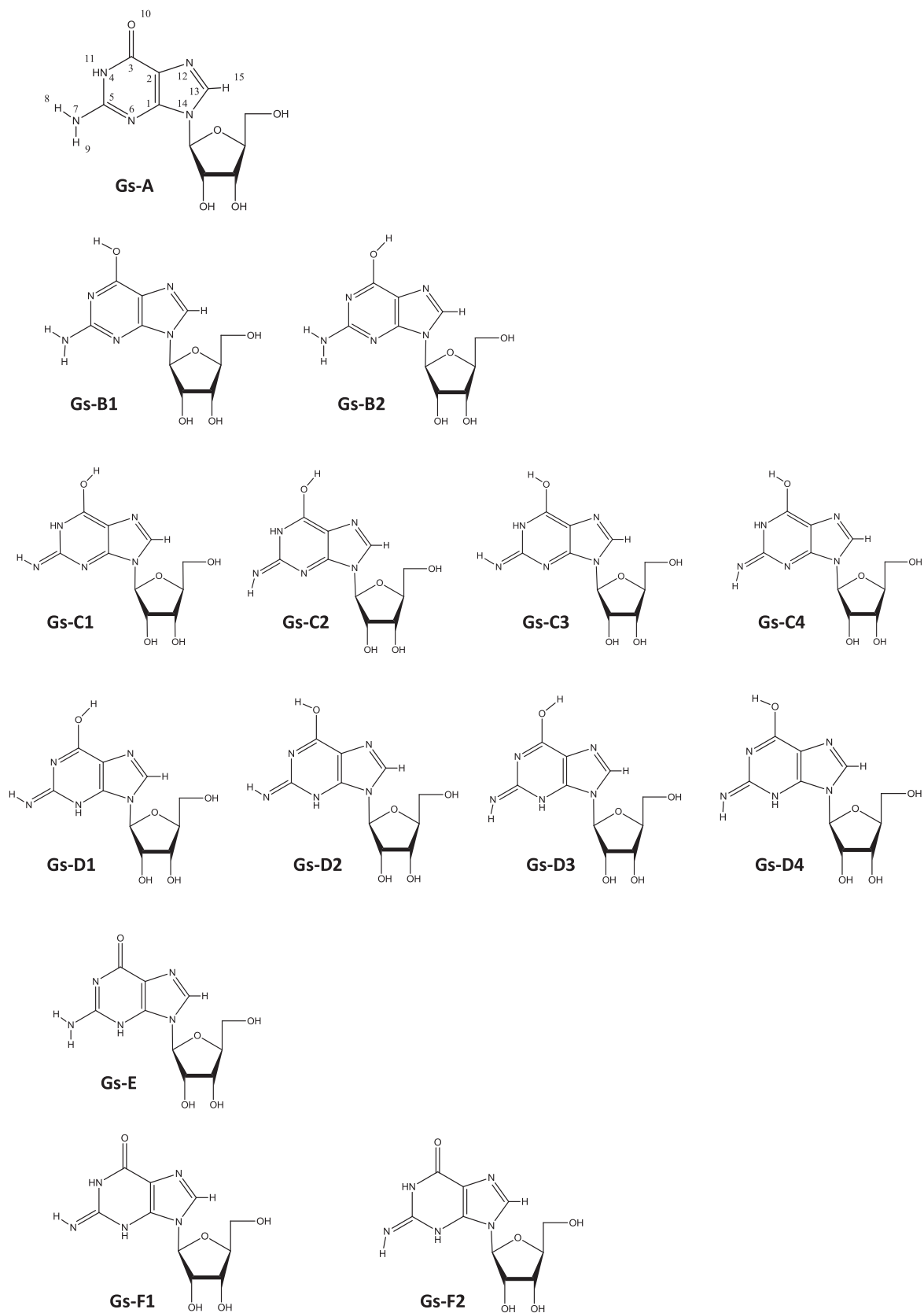


Figure 1. Guanine tautomers and rotamers, and the label of the guanine fragment.

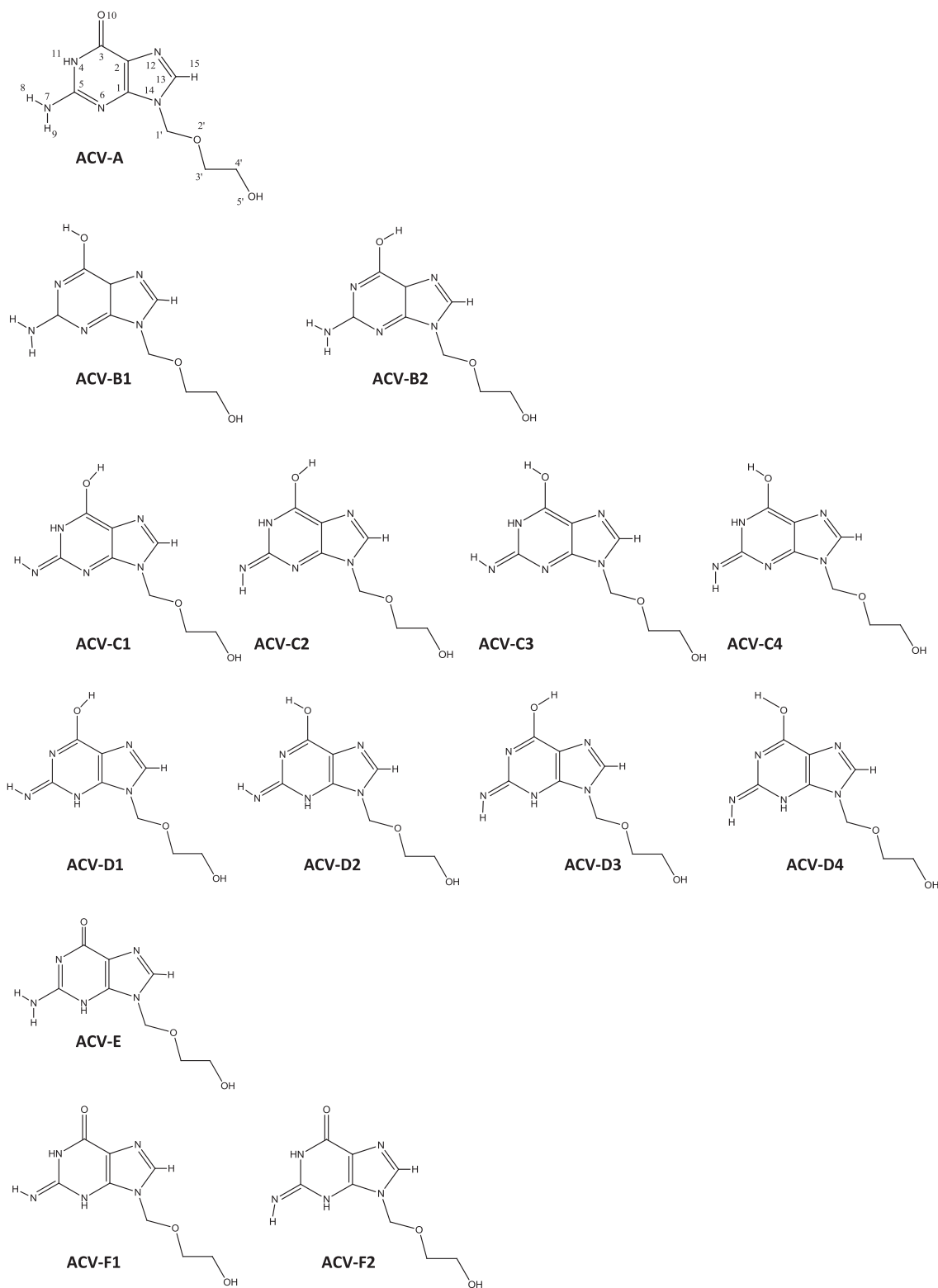


Figure 2. Acyclovir tautomers and rotamers, and the label of the guanine fragment and the side chain.

Table 1. Relative free energies, ΔG_{298} (kcal mol⁻¹), and populations (in %) for the tautomers and rotamers of guanosine (Figure 1) and acyclovir (Figure 2) in the gas phase and in water solution (C-PCM), calculated at MP2/6-31+G(d,p) level.

Molecule	Gas	Pop.	C-PCM	Pop.	Molecule	Gas	Pop.	C-PCM	Pop.
Gs-A	0.26	27.89	0.00	99.93	ACV-A	0.00	55.83	0.00	99.98
Gs-B1	0.00	43.26	4.42	0.06	ACV-B1	0.36	30.41	5.43	0.01
Gs-B2	0.24	28.85	5.97	4×10^{-3}	ACV-B2	0.83	13.76	5.47	9.8×10^{-3}
Gs-F1	17.73		16.95		ACV-F1	14.16		13.73	
Gs-F2	18.87		18.35		ACV-F2	15.51		13.98	
Gs-E	19.85		14.46		ACV-E	18.00		9.89	
Gs-D2	25.45		28.08		ACV-D2	23.01		24.12	
Gs-D1	27.07		27.41		ACV-D1	24.54		24.24	
Gs-D3	34.25		30.48		ACV-D3	31.75		27.29	
Gs-D4	31.98		27.58		ACV-D4	29.30		27.32	
Gs-C4	28.80		30.20		ACV-C4	29.41		30.27	
Gs-C3	35.53		33.45		ACV-C3	37.89		34.40	
Gs-C2	22.80		26.99		ACV-C2	23.87		26.76	
Gs-C1	28.42		29.61		ACV-C1	30.55		31.03	

Table 2. MP2/6-31+G(d,p) calculated and experimental bond lengths (in Å) for guanosine tautomer **Gs-A**. For numbering of the atoms see Figure 1.

Bonds	Exptl.*	Isolated	Gs + 4 H ₂ O	C-PCM	C-PCM Gs + 4 H ₂ O
C1–C2	1.374	1.397	1.396	1.395	1.392
C2–C3	1.419	1.440	1.435	1.432	1.431
C3–N4	1.390	1.435	1.415	1.413	1.401
C3–O10	1.237	1.228	1.242	1.241	1.251
N4–C5	1.370	1.370	1.368	1.374	1.372
C5–N6	1.323	1.315	1.325	1.321	1.328
C5–N7	1.338	1.384	1.361	1.367	1.357
C6–C1	1.354	1.375	1.369	1.367	1.364
C1–N14	1.379	1.370	1.368	1.368	1.366
C2–N12	1.389	1.375	1.378	1.378	1.379
N12–C13	1.300	1.331	1.331	1.334	1.334
C13–N14	1.383	1.377	1.380	1.372	1.375
N14–C16	1.454	1.464	1.464	1.464	1.463
C16–O17	1.412	1.411	1.412	1.409	1.410
O17–C18	1.455	1.452	1.453	1.457	1.459
C18–C19	1.530	1.524	1.526	1.523	1.522
C18–C29	1.513	1.513	1.513	1.511	1.511
C29–O30	1.418	1.431	1.431	1.432	1.431
C19–C20	1.536	1.526	1.522	1.524	1.521
C19–O25	1.421	1.424	1.425	1.430	1.430
C20–C16	1.522	1.541	1.540	1.543	1.539
C20–O21	1.401	1.419	1.424	1.422	1.422
O21–H22		0.978	0.989	0.988	0.986
O25–H26		0.972	0.971	0.972	0.973
O30–H31		0.966	0.966	0.968	0.968
N4–H11		1.014	1.022	1.015	1.026
N7–H8		1.011	1.012	1.012	1.013
N7–H9		1.011	1.014	1.011	1.051
O34–H11			1.903		1.833
O10–H37			1.814		1.827
O40–H9			2.171		2.139
N6–H44			2.148		2.162

*Reference (Thewalt et al., 1970).

4.1. Solute–solvent hydrogen bonding network analysis and the choice of clusters

In solution, the hydrogen-bonding network changes dynamically. One has to, therefore, necessarily adopt a statistical description of the fluctuating network. The statistical description is, however, based on analysis of numerous configurations (snapshots) from the equilibrated MC run. The hydrogen-bonding interaction itself is a rather complex phenomenon, and even for a single configuration extracted from numerous phase space samples it is not an easy task to define whether, e.g., a given solvent molecule takes part in a hydrogen-bonding interaction with a solute molecule. In a similar manner as in

our previous study, following the approach initially implemented in Ref. (Markova et al., 2015), we here adopt the following strategy for formulating an unambiguous definition of in-liquid hydrogen-bonding pattern. The strategy will be described on an example of a single hydrogen-bond accepting center within the **Gs-A** molecule. We first construct the $N\cdots O_w$ radial distribution function from the equilibrated MC run (Figure 3(a)). A well-defined peak due to a population of solvent molecules residing in the vicinity of N12 center (Figure 1) is clearly defined starting at about 2.55 Å and ranging up to 3.35 Å (a population of the ‘first-shell’ water molecules around this particular N center). Integration over

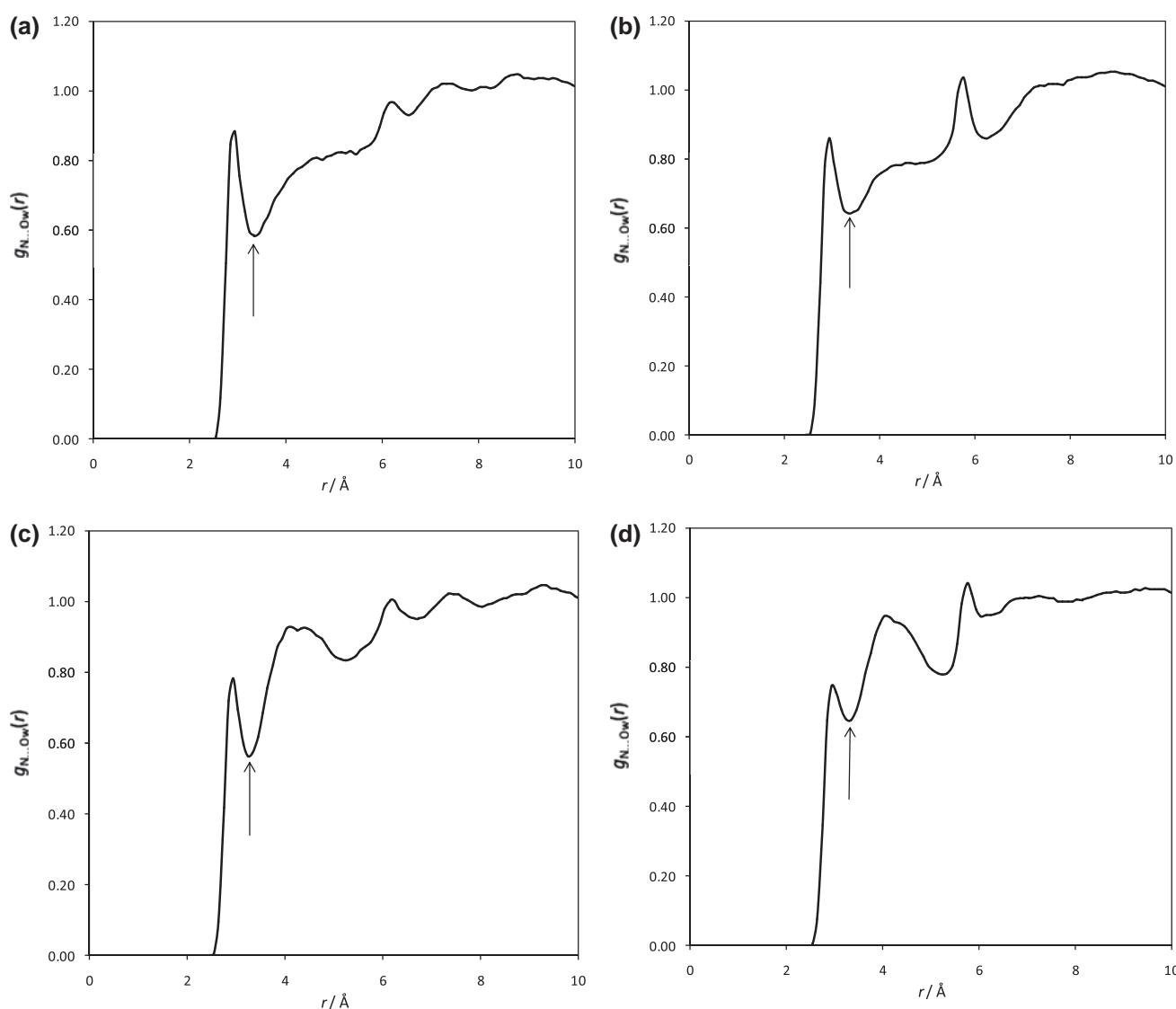


Figure 3. The $N12\cdots O_w$ radial distribution functions computed from the equilibrated MC run for the tautomeric forms of: guanosine – **Gs-A** (a) and **Gs-B1** (b), and acyclovir – **ACV-A** (c) and **ACV-B1** (d). The arrows denote the r_c values, i.e. the upper limit in the integral in Equation (9), used to compute N_c .

the first peak of this particular RDF was used to count the number of water molecules residing in this ‘first hydration shell’ around N12:

$$N_c = 4\pi\rho \int_0^{r_c} g_{N\dots Ow}(r) r^2 dr \quad (9)$$

This particular value of the ‘coordination number’ has been often used as a direct measure of the number of hydrogen bonds between solute and solvent molecules. This, however, is not a completely correct approach, as it corresponds to imposing of only a single criterion in the hydrogen bond definition – the distance criterion. The hydrogen bond is, on the other hand, a noncovalent interaction with a strongly pronounced *directional* character. As a simple analysis of the interaction energy distributions versus intermolecular distances indicates (Figure 4(a) and (b)), even at low intermolecular distances the

interaction energies may be unfavorable for an intermolecular interaction of this type, due to inappropriate orientation of the interacting species, as a consequence of the thermally induced motions within the liquid. A plot of the histogram of interaction energies, besides the singularity at $E = 0$ (which occurs as a result of the very large number of weak ion-dipolar pair interactions at large solute–solvent distances, at which the solvent molecules are irregularly distributed as a consequence of thermal motions within the liquid), clearly exhibits a ‘low-energy tail’, occurring exactly due the existence of a population of solvent molecules taking part in specific, stronger noncovalent interactions with the solute – Figure 5(a) and (b). The second criterion that we therefore impose to exactly define the in-liquid hydrogen bond is the energetic one – the intermolecular interaction energies should not be higher than $-1.0 \text{ kcal mol}^{-1}$ (the estimated ‘high-energy tail’ cut-off of the low-energy peak in the histogram on Figure 5(a) and (b)).

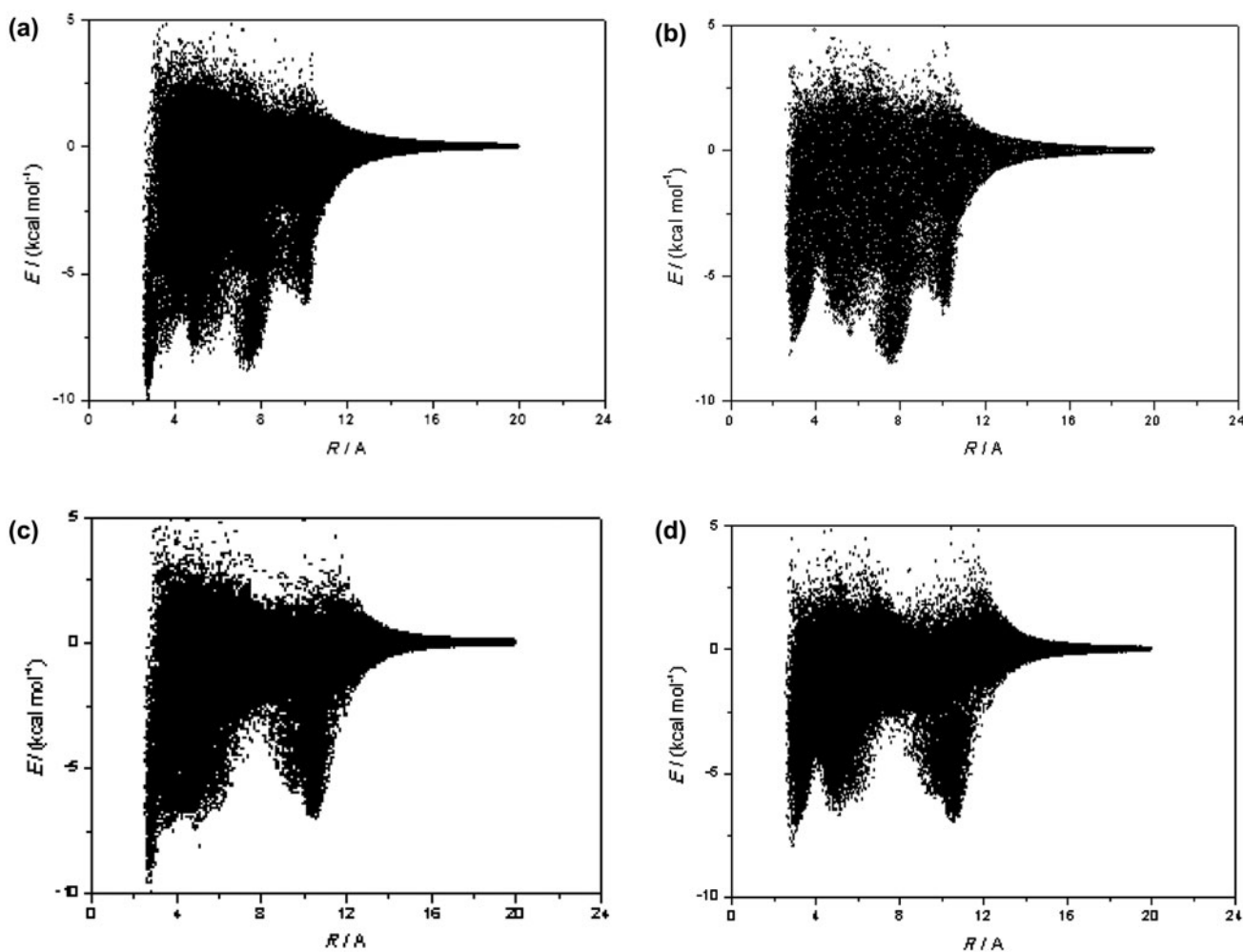


Figure 4. Distribution of pair interaction energies as a function of intermolecular distance R computed from the equilibrated MC run in the case of tautomeric forms of: guanosine – Gs-A (a) and Gs-B1 (b), and acyclovir – ACV-A (c) and ACV-B1 (d).

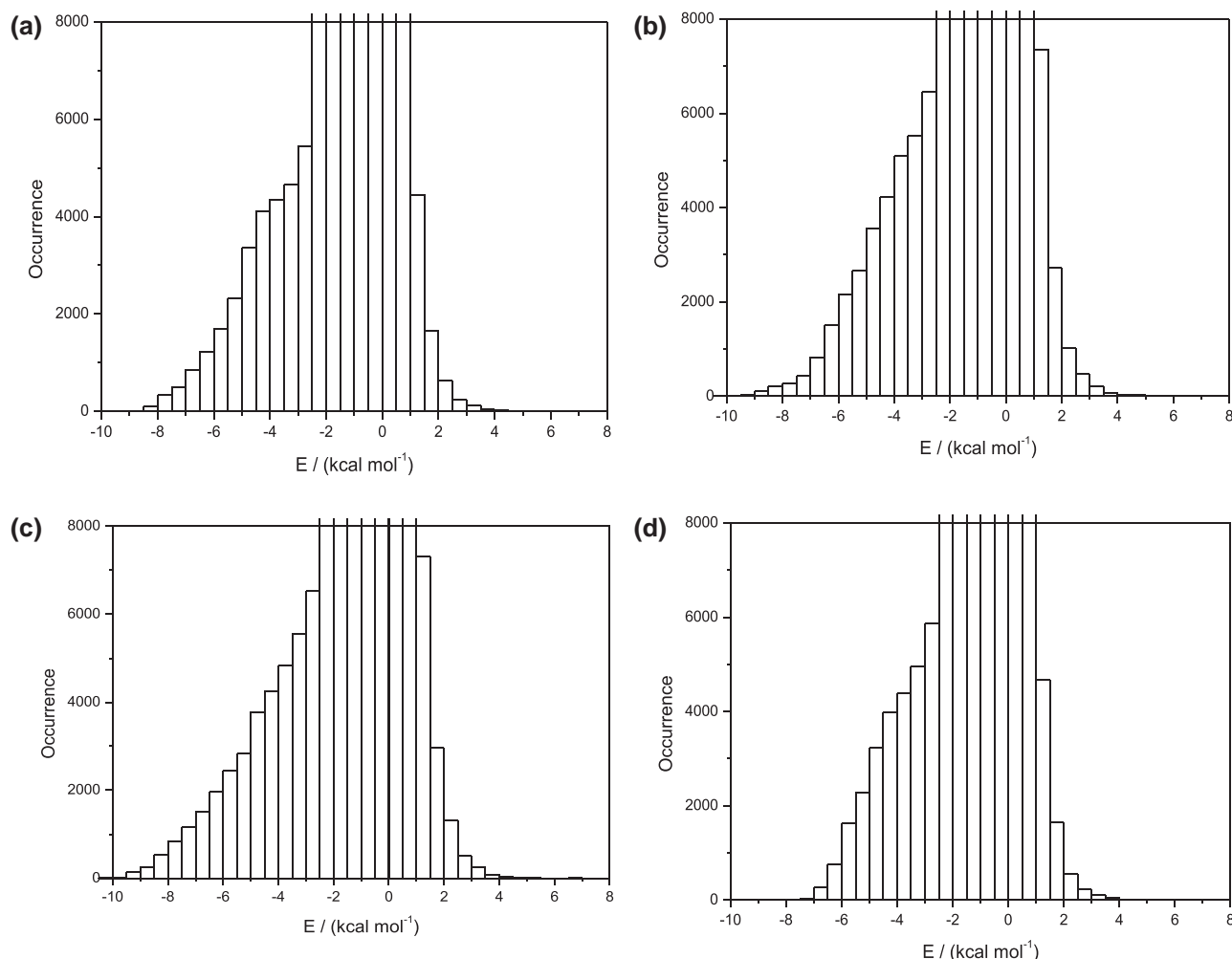


Figure 5. Histogram of distribution of pair interaction energies computed from the equilibrated MC run for tautomeric forms of: guanosine – **Gs-A** (a) and **Gs-B1** (b), and acyclovir – **ACV-A** (c) and **ACV-B1** (d).

Finally, we also impose a criterion related to the correct direction of the NO_wH interaction by plotting the histogram of distributions of the NO_wH angles for those solvent molecules that already fulfill the previously mentioned criteria. This histogram is shown in Figure 5(a) and (b).

It can be inferred from Figure 6(a) and (b) that the low-angle peak obviously corresponds to ‘correct’ orientation of the $\text{N}\dots\text{H}-\text{O}_w$ fragments in the sense of the directional character of hydrogen bond. The high-angle peak corresponds to nonhydrogen-bonded structures, with favorable interaction energy due to dipole-dipole interactions. The minimum between the two peaks ($\sim 60^\circ$) is therefore naturally imposed as an ‘angular’ criterion for in-liquid hydrogen bonding to the considered center in the present case.

On the basis of the defined criteria, we have concluded that the N12 center participates within the in-liquid hydrogen-bonding interactions, serving as proton acceptor, with

1.48 water molecules (on average). Analogous analyses have been performed for all centers within both tautomeric forms of guanosine molecule, playing both the role of proton acceptor and proton donor. The results are summarized in Table 3. A careful inspection of the molecular structures of the considered systems, accounting for the steric hindrance of solute-solvent interactions at particular centers, leads to an unequivocal conclusion that the outlined hydrogen-bonding analysis of the MC results are quite logical and should realistically reflect the relative hydrogen-bonding acceptor or donor abilities of intramolecular centers within the solute molecules.

We have further used the MC results to aid the buildup of finite clusters (supermolecules) that will be subjected to higher-level quantum chemical computations. On the basis of the results shown in Table 3, we have chosen the clusters containing a single solute molecule and four solvent (water) molecules, depicted in Figure 7. The particular choice of clusters was based on

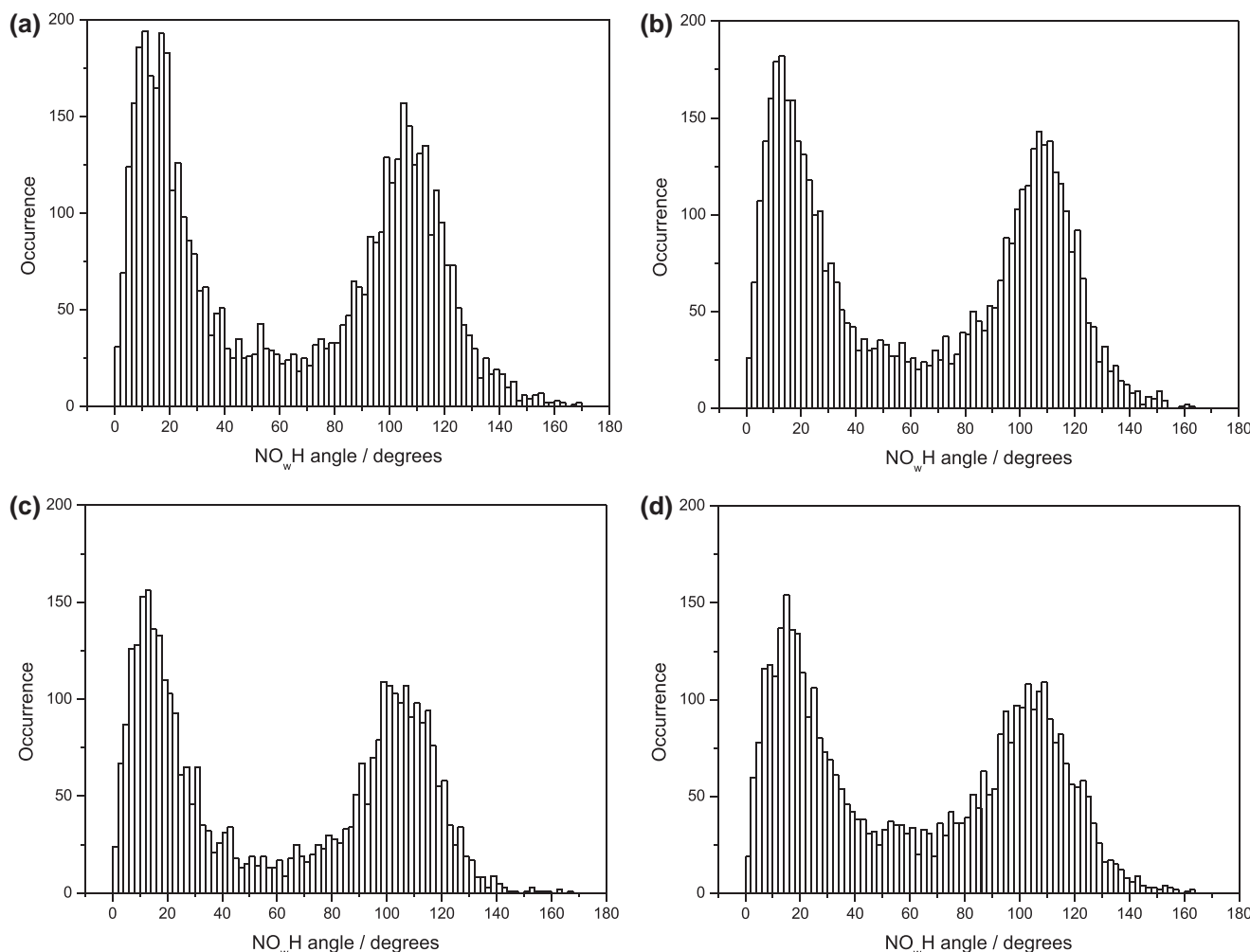


Figure 6. Histogram of distributions of the NO_wH angles for those solvent molecules that already fulfill the previously defined distance and energy criteria computed from the equilibrated MC run for tautomeric forms of: guanosine – **Gs-A** (a) and **Gs-B1** (b), and acyclovir – **ACV-A** (c) and **ACV-B1** (d).

the following reasoning. Since we are primarily interested in the water-assisted intramolecular proton transfer from N4 to O centers (Figure 1), we have accounted for the hydrogen bonding with solvent molecules of these particular centers by adding one water molecule at N4 and O in all considered tautomers of the studied systems. Both NH potential proton donors of the neighboring amino group, however, have a rather significant proton-donating ability. Due to this reason, as well as due to its vicinity to the actual proton-transfer site, we have also included a single solvent water molecule hydrogen bonded to its oxygen atom serving as a proton-acceptor at the N7H9 site (not to the N7H8 one, due to potential steric crowding around the PT center). This particular water molecule was placed such as to potentially serve as a hydrogen bond proton donor with the O21 (Figure 7) proton-donating site of the solvent molecule, which,

according to our MC analysis of the in-liquid hydrogen-bonding pattern, appears to be a close center with maximum hydrogen bond proton-accepting ability. We finally include the fourth, final water molecule within the cluster positioned as depicted in Figure 7, hydrogen bonded to the O40 center of the third solvent molecule that has been explicitly included in the supermolecule.

The complexes containing four water molecules around the tautomeric forms **Gs-A** and **Gs-B1** were used for simulation of the interactions in aqueous medium. We consider the solvent effect on the tautomeric conversion of the **Gs**, shown on Figure 7, through assisted proton transfer mechanism. The TS structure corresponding to the **Gs-A** \rightarrow **Gs-B1** water-assisted proton transfer reaction was found and the energy barrier of the reaction was calculated to be $18.26 \text{ kcal mol}^{-1}$ (Figure 8(a)). 2-Amino-6-(*sZ*)-hydroxy tautomer **Gs-B1** was found to

Table 3. The average number of H-bonds with the solvent water molecules in which each particular center within **Gs-A** and **Gs-B1** tautomers of guanosine (either as proton-donor or as proton-acceptor), computed by the procedure described in the text. For the numbering of the atoms see Figure 1.

Center Gs-A	Average number of accepted hydrogen bonds	Average number of donated hydrogen bonds	Center Gs-B1	Average number of accepted hydrogen bonds	Average number of donated hydrogen bonds
N4	0.01	—	N4	0.62	—
N6	0.12	—	N6	0.16	—
N7	0.32	—	N7	0.38	—
N12	1.48	—	N12	1.53	—
N14	0.00	—	N14	0.01	—
O10	1.21	—	O10	0.36	—
O17	0.68	—	O17	0.67	—
O21	0.86	—	O21	0.98	—
O25	1.06	—	O25	1.16	—
O30	0.44	—	O30	0.39	—
N4H11	—	0.81	O10H11	—	0.86
N7H8	—	0.82	N7H8	—	0.73
N7H9	—	0.79	N7H9	—	0.64
O21H22	—	0.27	O21H22	—	0.13
O25H26	—	0.41	O25H26	—	0.46
O30H31	—	0.94	O30H31	—	0.94

be $5.27 \text{ kcal mol}^{-1}$ higher in energy than the 1*H*-2-amino-6-oxo tautomer **Gs-A**, a value very close to that calculated by C-PCM model (Table 1).

In order to take into account the impact of the surrounding effect on the tautomerization process of the investigated complexes, we have repeated the geometry optimizations at the MP2/6-31+G(d,p) level using the conductor-like polarizable continuum model (CPCM/MP2/6-31+G(d,p)), choosing the continuum with a dielectric constant of $\epsilon = 78.38$ to mimic surrounding media. The correspondence of the stationary points to minimum on the potential energy surface or TS has again been checked by the absence or the presence, respectively, of one and only one imaginary frequency corresponding to the normal mode that identifies the reaction coordinate.

Using this computational model, it was found that the energy difference between 1*H*-2-amino-6-oxo **Gs-A** and 2-amino-6-(*sZ*)-hydroxy **Gs-B1** tautomers slightly increase to $5.60 \text{ kcal mol}^{-1}$, but the tautomeric conversion is characterized with an energy barrier (ΔG_{298}^\ddagger) lower by $11.74 \text{ kcal mol}^{-1}$ (Figure 8(a)). The calculated rate constant of the reaction, $k = 15.37 \text{ s}^{-1}$ (Table 4), is sufficiently large to be able to generate a concentration of the tautomeric form **Gs-B1**.

Since the stationary points were located, the reaction pathway was established by following the IRC in the forward and reverse directions from TS (Figure 9). These calculations ensure that the proper reaction pathway, connecting the reactant and product on each side of the TS, has been found.

The reaction starts as both of the water molecules situated around the N4 and O10 centers (Figure 7) draw

near to guanosine molecule. It can be seen from Figure 9 that the N4...O34 distance decreases from the equilibrium value $2.875\text{--}2.498 \text{ \AA}$. At the same time, the equilibrium value of the O10...O38 distance (2.778 \AA) decreases to 2.434 \AA . The proton H11, involved in the N4–H11...O34 hydrogen bond, exhibits a slight move from 1.022 \AA (equilibrium distance in N4–H11 bond) to 1.076 \AA , while protons H36 (involved in the O34–H36...O38 hydrogen bond) and H37 (involved in the O38–H37...O10 hydrogen bond) move from equilibrium distance of 0.983 \AA in the O–H bonds to 1.035 and 1.019 \AA , respectively. At this point on the reaction coordinate, the protons H11, H36 and H37 suddenly start the transfer to the respective acceptor oxygen atoms O34, O38 and O10. The TS (Figure 7) is formed when the distances N4–H11, O34–H36 and O38–H37 become 1.307 , 1.256 and 1.137 \AA , respectively, the carbonyl bond C3=O10 is elongated from 1.242 to 1.278 \AA and C3–N4 bond in the six-membered ring is shortened from 1.415 to 1.381 \AA . After that the N4...O34 and O10...O38 distances again start to return to their equilibrium values (Figure 9). The protons H11, H36 and H37 continue to transfer to the atoms O34, O38 and O10, respectively, and O34–H11, O38–H36 and O10–H37 bonds are formed. The process is cooperative, and from the mechanistic point of view, this means that the proton transfers happen in a single step without any intermediates.

The reaction profile is different when we consider the second, more sophisticated computational model – optimization of the cluster **Gs** + **4H₂O** using the conductor-like polarizable continuum model, i.e. C-PCM/MP2/6-31+G(d,p) (curve B, Figure 9). The N4...O34 and O10...O38 distances decrease slowly from the

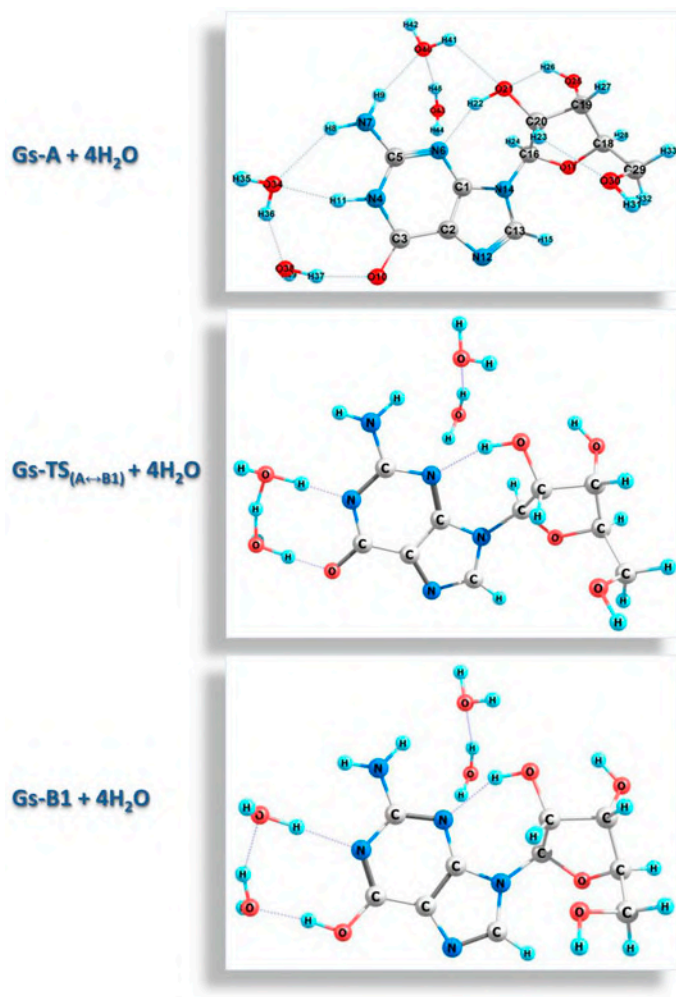


Figure 7. Structures of the hydrated tautomeric forms **Gs-A** and **Gs-B1** and respective transition state calculated at MP2/6-31+G(d, p) level.

equilibrium distance of 2.817 and 2.793 Å to 2.568 and 2.524 Å, resp., while O34...O38 distance exhibit more pronounced variations from 2.739 to 2.452 Å (Figure 9). At this point on the reaction coordinate, the proton H11 rapidly starts to move from nitrogen atom N4 to the oxygen atom O34, and O34–H36 bond slowly begins to elongate. When O34...O38 distance becomes 2.399 Å, proton H36 start a rapid transfer to oxygen atom O38 and O38–H37 bond begins a slow elongation. The TS is formed when the distances N4–H11, O34–H36 and O38–H37 become 1.479, 1.218 and 1.054 Å, respectively; the carbonyl bond C3=O10 is elongated to 1.281 Å and the single C3–N4 bond is shortened to 1.375 Å. After that the N4...O34 and O34...O38 distances begin to elongate slowly to their equilibrium values while O38–O10 continues to shorten. The O34–H11 distance which is 1.074 Å is slowly shortened to formation of O–H bond. Proton H36 continues to transfer rapidly to O38 and O38–H37 bond begins to elongate

strongly. When O38–O10 distance becomes 2.473 Å, the proton H37 starts to transfer more slowly and distance O38–O10 achieves its equilibrium value while bond O10–H37 is formed.

The calculated profiles of the proton transfer along the IRC of the tautomerization (Figure 9) evidence that this tautomerization reaction occurs through the asynchronous concerted mechanism.

The time necessary to reach 99.9% of the equilibrium concentration between the tautomeric forms 1*H*-2-amino-6-oxo **Gs-A** and 2-amino-6-(*sZ*)-hydroxy **Gs-B1** was calculated to be $\tau_{99.9\%} = 3.53 \times 10^{-5}$ s (Table 4).

5. Acyclovir

According to theoretical study of **ACV** by Alvarez-Ros et al. (Alvarez-Ros & Alcolea Palafox, 2014), the most stable structure corresponds to amino-oxo tautomer

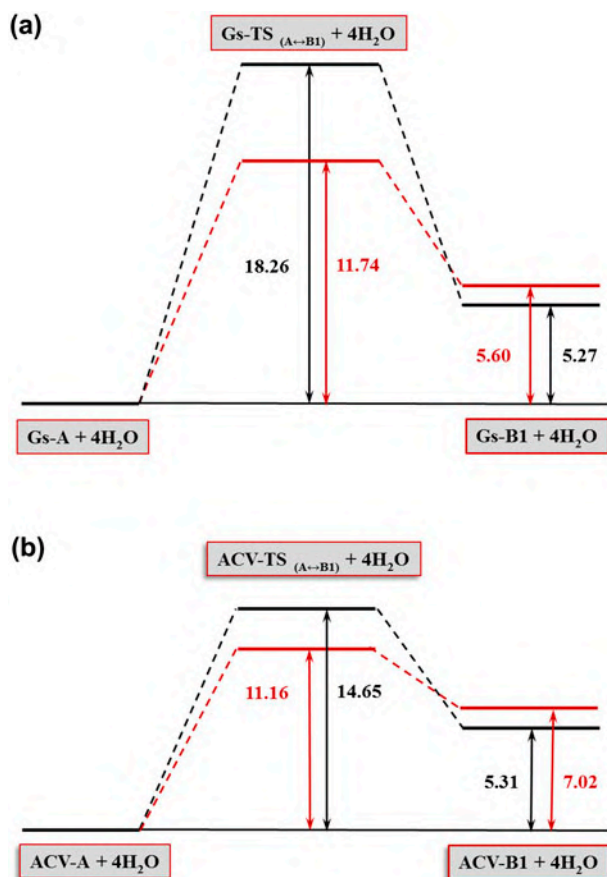


Figure 8. MP2/6-31+G(d,p) and C-PCM/MP2/6-31+G(d,p) (in red) calculated energy difference between four-hydrated complexes and energy barrier of the assisted proton transfer reactions (in kcal mol⁻¹) of the guanosine (a) and acyclovir (b).

ACV-A. The second population is due to rotamer of amino-hydroxy tautomer **ACV-B1** (37.7%) and the third one is to rotamer **ACV-B2** (14.3%). The rest tautomers have very little population, less than 0.05%.

Our results are similar and show that in the gas phase and in water solution, the most stable is the 1*H*-2-amino-6-oxo form **ACV-A** followed by the 2-amino-6-(*sZ*)-hydroxy form, **ACV-B1**. When isolated molecules are considered the energy difference between **ACV-A** and

ACV-B1 is small (0.36 kcal mol⁻¹), similarly to the case of energy difference between the two tautomers of guanosine (Table 1). According to calculated populations, the amount of the amino-oxo tautomer **ACV-A** is 55.83% and the 2-amino-6-(*sZ*)-hydroxy form **ACV-B1** (30.41%) prevails over **ACV-B2** (13.76%).

Taking into account the effect of water only as dielectric medium by the C-PCM model at MP2 level, the energy difference between the 1*H*-2-amino-6-oxo and 2-amino-6-(*sZ*)-hydroxy forms increases substantially. The energy difference between **ACV-A** and **ACV-B1** increases 15 times and become 5.43 kcal mol⁻¹. The amount of the amino-oxo tautomer prevails over both of the amino-hydroxy forms and the ratio become **ACV-A**: **ACV-B1**: **ACV-B2** = 99.98: 0.01: 9.8 × 10⁻³%.

According to crystallographic data (Birnbbaum et al., 1984) and conformational analysis (Alvarez-Ros & Alcolea Palafox, 2014) in solid state only one tautomeric form of acyclovir, **ACV-A** is observed. Three types of structures appear for acyclovir in relation to the side chain that is attached to N14. While in the first two molecules, the bonds are in the preferred gauche conformation in the third one the bonds in the side chain are all in the transconformation, giving rise to an almost planar, zig-zag arrangement. Our results for geometric parameters of tautomeric form **ACV-A** calculated at MP2/6-31+G(d,p) level and presented in Table 5 are closer to the zig-zag molecule conformation (Figure 10). The side chain is characterized by the exocyclic torsional angles χ_1 to χ_5 , by the bond angles θ and α that determine the position of the chain from the base. According to our calculations, the torsional angles in tautomer **ACV-A** are: χ_1 (C1–N14–C1'–O2') = –71.5°, χ_2 (N14–C1'–O2'–C3') = 163.4°, χ_3 (C1'–O2'–C3'–C4') = –175.5°, χ_4 (O2'–C3'–C4'–C5') = 179.1° and χ_5 (C3'–C4'–C5'–H5') = 174.2°. The values of the bond angles θ and α corresponding to the minima on the molecular PES (obtained from full geometry optimization) of **ACV-A** are: θ (C1–N14...O5') = 104.1° and α (N14–C1'...O5') = 137.3°.

The dihedral angles describing the base heterocycle nonplanarity in guanine moiety of all acyclovir tautomers can be defined as: φ_0 (N4–C5=N6–C1), φ_1 (C5=N6–C1=C2), φ_2 (N6–C1=C2–C3), φ_3 (N12=C13–N14–C1)

Table 4. Calculated rate constants for the forward (k_f) and reverse (k_r) (s⁻¹) solvent-assisted proton transfer reaction in guanosine (Figure 8(a)) and acyclovir (Figure 8(b)), equilibrium constant K_T , and time ($\tau_{99.9\%}$) necessary to reach 99.9% of the equilibrium concentration (s).

Compound	k_f	k_r	K_T	$\tau_{99.9\%}$
Guanosine				
Cluster	2.55×10^{-4}	1.86	1.37×10^{-4}	3.71
CPCM/cluster	15.377	1.959×10^5	7.84×10^{-5}	3.53×10^{-5}
Acyclovir				
Cluster	0.113138	883.62127	1.28×10^{-4}	7.82×10^{-3}
CPCM/cluster	4.09×10^1	5.73×10^6	7.14×10^{-6}	1.20×10^{-6}

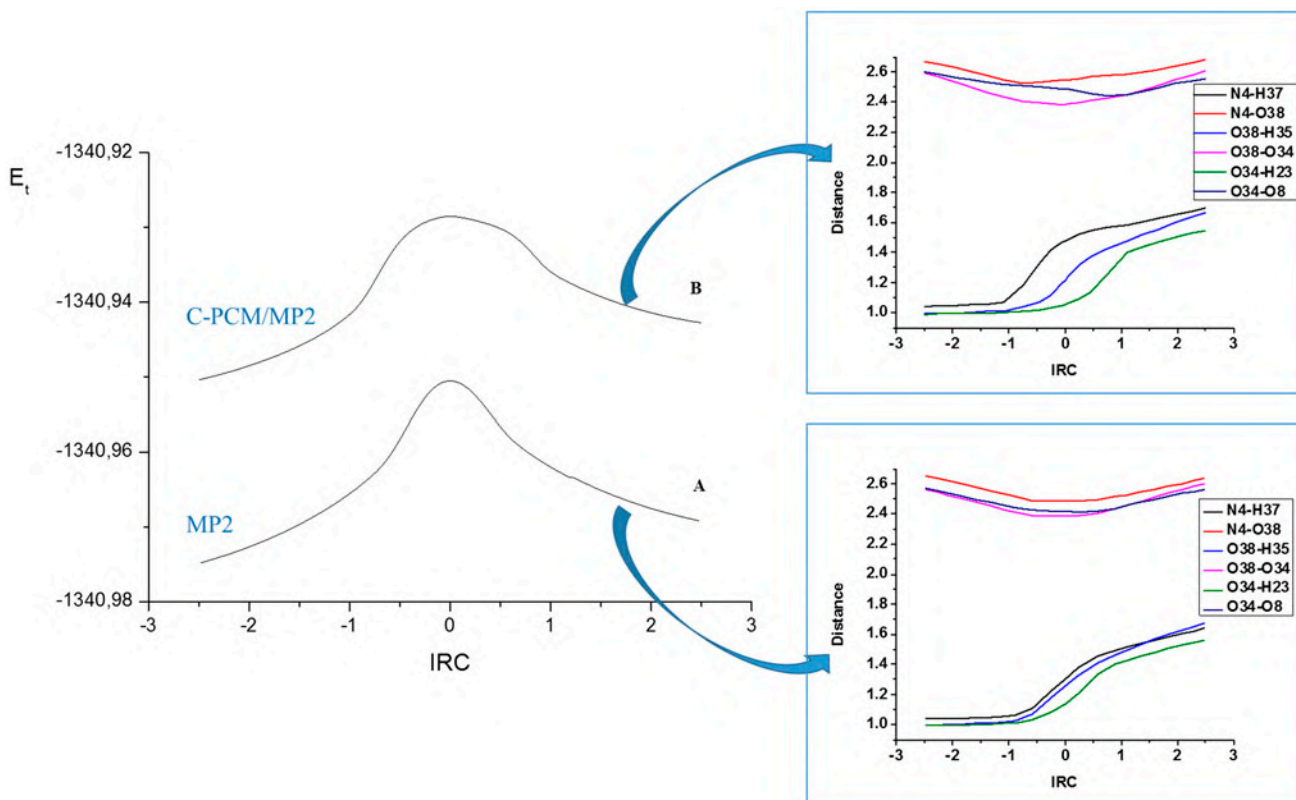


Figure 9. IRC profile of the water-assisted proton transfer reaction of guanosine **Gs-A**→**GS-B1** calculated at MP2/6-31+G(d,p) (profile A) and C-PCM/MP2/6-31+G(d,p) (profile B) levels. Insert: Distances between two selected atoms along the IRC profile of the four-hydrated guanosine as calculated at MP2 and C-PCM/MP2 levels. Total energies are in a.u., distances are in Å and IRC in $\text{amu}^{1/2} \text{ bohr}$.

and φ_4 (C3–N4–C5 = N6). In the analysis of the most stable tautomer of acyclovir, **ACV-A** the base heterocycle has a very small nonplanarity with torsional angles between -0.5° and 1.5° . The highest nonplanarity especially in six-membered ring corresponds to tautomers **ACV-C2**, **ACV-C3**, **ACV-C4**, **ACV-D1**, **ACV-D3**, **ACV-E**, **ACV-F1** and **ACV-F2**. This tautomers and rotamers show deviations higher than 15° in the following torsional angles: tautomer **ACV-C2** in φ_0 (19.6°) and φ_4 (-15.7°), **ACV-C3** in φ_0 (20.8°) and φ_4 (-16.6°), tautomer **ACV-D4** in φ_0 (18.1°) and φ_4 (-14.2°), tautomer **ACV-D1** in φ_0 (22.0°) and φ_1 (-12.8°).

All bond lengths (Table 5) in the guanine residues calculated for isolated tautomer **ACV-A** in gas phase are in good agreement with published values derived from crystal structure analyses. Generally, the MP2/6-31+G(d,p) calculations predict slightly higher values of bond lengths in guanine moiety as well as in the side chain. The deviations are in range from 0.002–0.480 Å. For instance the predicted C3=O10 double bond is 1.229 Å while the corresponding experimental value is 1.241 Å.

5.1. Solute–solvent hydrogen-bonding network analysis and the choice of clusters

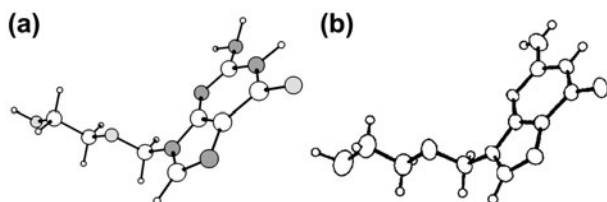
Similar logic has been followed in the course of constructing the supermolecular clusters further to be subjected to quantum-chemical calculations in the case of both considered tautomeric forms of **ACV** as well. We hereby present the relevant RDFs and distribution histograms related to the N12 proton-accepting centers in both tautomeric forms of **ACV** (**ACV-A** and **ACV-B1**). Figure 3(c) and (d) depicts the N...O_w radial distribution functions for the two tautomeric forms of **ACV** constructed from the equilibrated MC run. Analogously as in the case of guanosine, integration up to r_c (denoted by arrows in Figure 3(c) and (d)) gives the total number of solvent molecules residing in the first solvation shell around N12.

Not all of these, however, may be considered as being hydrogen bonded to N12, as discussed above. Existence of unfavorable interaction energy range even at low intermolecular distance is evident from plots given in Figure 4(c) and (d). However, as can be seen from a plot of the histogram of interaction energies (Figure 5(c))

Table 5. MP2/6-31+G(d,p) calculated and experimental bond lengths (in Å) for tautomer **ACV-A** of acyclovir (Figure 2).

Bonds	Exptl.*	Isolated	C-PCM	ACV + 4H ₂ O	C-PCM ACV + 4H ₂ O
C1–C2	1.392	1.394	1.394	1.397	1.395
C2–C3	1.412	1.442	1.434	1.434	1.428
C3–N4	1.383	1.431	1.407	1.408	1.395
C3–O10	1.241	1.229	1.242	1.243	1.251
N4–C5	1.372	1.374	1.376	1.374	1.374
N4–H11		1.014	1.014	1.023	1.028
C5–N6	1.332	1.312	1.317	1.328	1.328
C5–N7	1.336	1.386	1.374	1.361	1.362
N6–C1	1.339	1.366	1.360	1.365	1.361
C1–N14	1.379	1.376	1.373	1.377	1.375
C2–N12	1.385	1.379	1.381	1.380	1.380
N12–C13	1.301	1.325	1.330	1.324	1.328
C13–N14	1.374	1.380	1.376	1.383	1.378
N14–C16	1.443	1.435	1.348	1.457	1.461
C16–O19	1.394	1.416	1.417	1.410	1.409
O19–C20	1.424	1.430	1.435	1.437	1.441
C20–O23	1.460	1.512	1.510	1.518	1.517
H31–O10				1.804	1.787
O28–H11				1.891	1.826
H37–N6				1.859	1.862
H9–O34				1.877	1.898

*Reference (Birnbaum et al., 1984).

Figure 10. Structure of the tautomeric form **A** of acyclovir obtained by optimization at MP2 level (a) and X-ray diffraction (Birnbaum et al., 1984) (b).

and (d)), besides the singularity at $E = 0$, existence of a ‘low-energy tail’ is evident again, corresponding to a population of solute-solvent pairs taking part in specific, stronger noncovalent interactions. Further analysis of this histogram enables us to impose an energetic criterion to more strictly define the in-liquid hydrogen bond.

The first-shell water molecules around the N12 center which satisfy all of the previously described criteria, in order to be considered as hydrogen-bonded to this particular center also have to satisfy the orientation-related criterion, i.e., the directional nature of hydrogen-bonding noncovalent interaction needs to be accounted for. As can be seen from the distribution histograms of the NO_wH angles plotted in Figure 6(c) and (d) the subpopulation of solvent molecules (that already fulfill the previously mentioned criteria) characterized with NO_wH angle value of less than 60° may be actually considered as being hydrogen-bonded to N12. The results from such

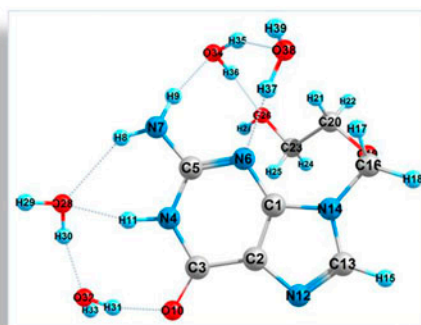
analyses for all centers within both tautomeric forms of acyclovir molecule, playing both the role of proton acceptor and proton donor are summarized in Table 6. On the basis of the results shown in Table 6, for further quantum-mechanical modeling of the proton transfer processes, we have chosen clusters containing a single solute molecule and four solvent (water) molecules, depicted in Figure 11, on the basis of analogous arguments as those presented above, for the case of guanosine.

Hence, we consider the solvent effect on the tautomeric conversion of acyclovir through assisted proton transfer model – cluster from molecule **ACV** and four water molecules. In this model, similarly as in the case of guanosine, two water molecules forming a cluster are situated between N–H and C=O groups and the another two ones – between amino group and nitrogen atom from six-membered ring of **ACV** (Figure 11). In the four-hydrated cluster of **ACV-A**, the orientation of the side chain that is attached to N14 undergoes significant changes. These include changes in torsional angles in **ACV-A**. The angle χ_1 changes from -71.5° to 87.1° , χ_2 from 163.4° became 86.0° and χ_3 from -175.5° to -101.8° . The torsion angles, χ_4 and χ_5 are not altered significantly. The most significant changes were observed in the bond angles θ and α in four-hydrated complex of **ACV-A**: the angles θ and α diminish from 104.1° to 69.4° and from 137.3° to 88.0° , respectively. These alterations in side chain position with respect to the plane of the nucleobase are due to the influence of the water

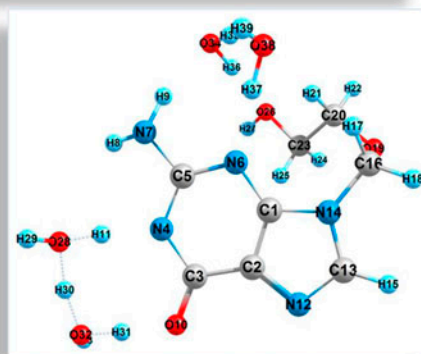
Table 6. The average number of H-bonds with the solvent water molecules in which each particular center within **ACV-A** and **ACV-B1** tautomers of **ACV** (either as proton-donor or as proton-acceptor), computed by the procedure described in the text. For the numbering of the atoms see Figure 2.

Center	Average number of accepted hydrogen bonds ACV-A	Average number of donated hydrogen bonds	Center	Average number of accepted hydrogen bonds ACV-B1	Average number of donated hydrogen bonds
N4	0.01	—	N4	0.68	—
N6	0.60	—	N6	0.68	—
N7	0.35	—	N7	0.40	—
N12	1.12	—	N12	1.40	—
N14	0.08	—	N14	0.07	—
O10	1.22	—	O10	0.37	—
O19	0.79	—	O19	0.89	—
O26	1.28	—	O26	1.30	—
N4H11	—	0.82	N7H8	—	0.63
N7H8	—	0.80	N7H9	—	0.55
N7H9	—	0.66	O26H27	—	0.96
O26H27	—	0.94	O10H11	—	0.86

ACV-A + 4H₂O



ACV-TS_(A↔B1) + 4H₂O



ACV-B1 + 4H₂O

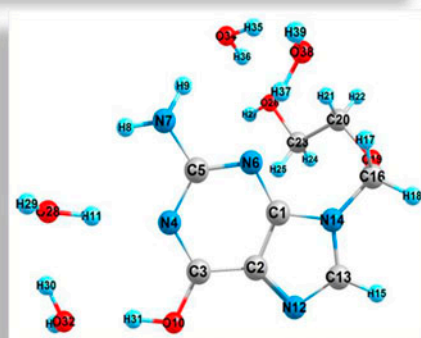


Figure 11. Structures of the hydrated tautomeric forms **ACV-A** and **ACV-B1** and respective transition state calculated at MP2/6-31+G(d,p) level.

molecules. The same side chain position is observed when the four-hydrated complex of **ACV-A** is embedded in water continuum (C-PCM/MP2/6-31+G(d,p)). This orientation of the chain remains in the cases of another tautomer **ACV-B** and consequently in the TS. The hydrogen bonds connecting one of the water molecules (that do not actually assisted the proton transfer) to OH group of the side chain and NH₂ group of the guanine moiety are present in all optimized structures. These hydrogen bonds stabilize the orientation of the chain in such a way that it at least partially resembles a portion of the furanose moiety of **Gs**.

The energy barrier of the proton transfer reaction **ACV-A** → **ACV-B1** was calculated to be 14.65 kcal mol⁻¹ (Figure 8(b)). 2-Amino-6-(*sZ*)-hydroxy tautomer **ACV-B1** was found to be 5.31 kcal mol⁻¹ higher in energy then the 1*H*-2-amino-6-oxo tautomer **ACV-A**, a value practically identical with the one calculated for isolated molecule by C-PCM model (Table 1).

As the next step, we consider more sophisticated computational model which takes into account the impact of the surrounding medium on the tautomerization process. In an analogous way in the case of

guanosine we have repeated the geometry optimizations of the investigated complexes and the transition state (Figure 11) using the conductor-like polarizable continuum model (C-PCM/MP2/6-31+G(d,p)). It was found that the energy difference between 1*H*-2-amino-6-oxo **ACV-A** and 2-amino-6-(*sZ*)-hydroxy **ACV-B1** tautomers increases to 7.02 kcal mol⁻¹ but the energy barrier (ΔG_{298}^\ddagger) decreases to 11.16 kcal mol⁻¹ (Figure 8(b)). The calculated rate constant of the reaction, $k = 4.09 \times 10^1$ s⁻¹ (Table 4), is sufficiently large to generate an appreciable concentration of the tautomeric form **ACV-B1**. The time necessary to reach 99.9% of the equilibrium concentration of the tautomeric forms 1*H*-2-amino-6-oxo **ACV-A** and 2-amino-6-(*sZ*)-hydroxy **ACV-B1** was calculated to be $\tau_{99.9\%} = 1.2 \times 10^{-6}$ s (Table 4).

The reaction pathway was established by following the IRC in the forward and reverse directions from TS for both computational models (Figure 12). First we consider the energy profile of the water-assisted proton transfer reaction of acyclovir (**ACV-A** → **ACV-B1**) calculated at MP2/6-31+G(d,p) level (Figure 12, curve A). Similar to guanosine the reaction starts as the both water molecules situated around the N4-H11 and O10

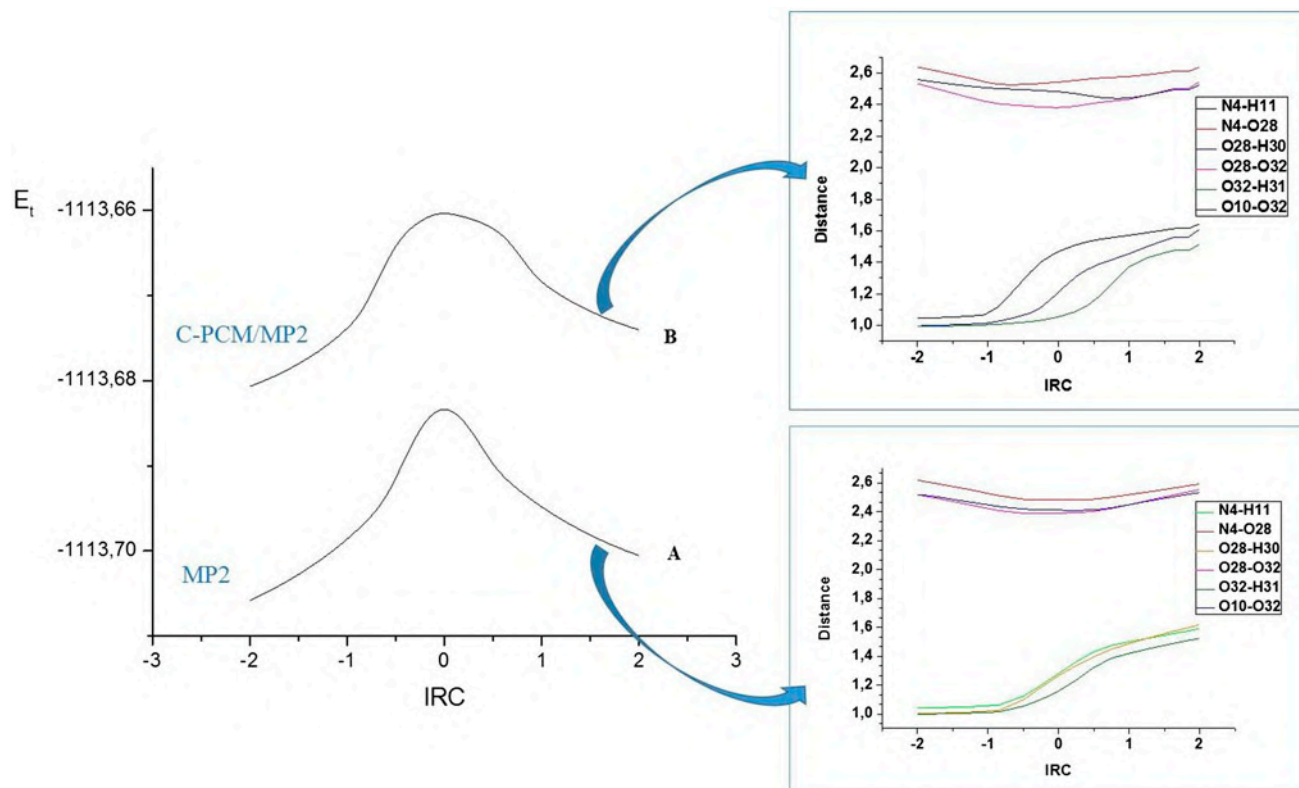


Figure 12. IRC profile of the water-assisted proton transfer reaction of acyclovir **ACV-A** → **ACV-B1** calculated at MP2/6-31+G(d,p) (profile A) and C-PCM/MP2/6-31+G(d,p) (profile B) levels. Insert: Distances between two selected atoms along the IRC profile of the four-hydrated acyclovir as calculated at MP2 and C-PCM/MP2 levels. Total energies are in a.u., distances are in Å and IRC in amu^{1/2} bohr.

centers (Figure 12) draw near to acyclovir molecule, while proton H11 moves from 1.023 Å (the equilibrium distance in N4–H11 bond, Table 5) to 1.074 Å and protons H30 and H31 move from 0.982 Å (equilibrium distance in the O–H bonds in the water molecules) to 1.014 and 1.003 Å, respectively. After that the protons H11, H30, and H31 transfer to the respective acceptor oxygen atoms O28, O32 and O10. The TS (Figure 11) is formed when the distances N4–H11, O28–H30 and O32–H31 become 1.263, 1.269 and 1.172 Å, respectively, the carbonyl bond C3=O10 is elongated from 1.243 Å (Table 5) to 1.283 Å and C3–N4 bond in the six-membered ring is shortened from 1.408 to 1.372 Å. After that the N4...O28 and O10...O32 distances start to return to their equilibrium values (Figure 12) and protons H11, H30 and H31 continue to transfer to the atoms O28, O32 and O10, respectively. The process is cooperative and the proton transfer happens in a single step without any intermediates.

The reaction profile obtained by more sophisticated model, C-PCM/MP2/6-31+G(d,p) (curve B, Figure 12) is different. The N4...O28 and O10...O32 distances decrease slowly to 2.537 and 2.503 Å, resp., and O28...O32 distance to 2.417 Å (Figure 12). After that the proton H11 rapidly moves from nitrogen atom N4 to the oxygen atom O28. When O28...O32 distance becomes 2.397 Å, proton H30 starts to transfer to oxygen atom O32. The TS is formed when the distances N4–H11, O28–H30 and O32–H31 become 1.464, 1.217 and 1.057 Å, respectively; the carbonyl bond C3=O10 is 1.280 Å and the C3–N4 bond is 1.370 Å. After that the N4...O28 and O28...O32 distances begin to elongated to their equilibrium values while O32–O10 continues to shorten. The O28–H11 distance is shortened to the value characteristic formation the O–H bond. When O32–O10 distance becomes 2.457 Å, the proton H31 starts to move from O32 to O10 and the O10–H31 bond is formed. The calculated profile of the water-assisted proton transfer (Figure 12, curve B) indicates that tautomerization reaction occurs through the asynchronous concerted mechanism.

6. Conclusions

The study of tautomerism in acyclovir is very important because of fact that the drug leads to mutagenesis by distinct base-pairing preferences of its different tautomers. ACV tautomers are paired with purine bases (guanine or adenine), by maximizing the number of possible hydrogen bonds between the bases. In aqueous solution, this can induce significant changes in its physical and biological properties.

The presented study consider for a first time water-assisted proton transfer in the guanine nucleoside

guanosine and its analogue acyclovir. The Monte Carlo simulations were used to buildup of finite clusters that were then subjected to higher-level quantum chemical computations. The MP2/6-31+G(d,p) and C-PCM/MP2/6-31+G(d,p) calculations show that in the gas phase and in water solution, the most stable for both compounds is the 1*H*-2-amino-6-oxo form followed by the 2-amino-6-(*sZ*)-hydroxy form.

The energy barriers (ΔG_{298}^\ddagger) of the water-assisted proton transfer reaction in guanosine and in acyclovir are similar, 11.74 kcal mol^{−1} for guanosine and 11.16 kcal mol^{−1} for acyclovir. The rate constant of the reaction, $k = 1.5 \times 10^1$ s^{−1} (guanosine) and $k = 4.09 \times 10^1$ s^{−1} (acyclovir), is sufficiently large to generate 2-amino-6-(*sZ*)-hydroxy tautomer. The calculated time necessary to reach the equilibrium concentration of the 1*H*-2-amino-6-oxo (**A**) and 2-amino-6-(*sZ*)-hydroxy (**B1**) tautomeric forms was 3.5×10^{-5} s for guanosine and 1.2×10^{-6} s for acyclovir.

The analysis of the profile of the water-assisted proton transfer reaction in guanosine and acyclovir show that the reaction occurs through the asynchronous concerted mechanism.

Acknowledgements

We acknowledge financial support by the Project BAS – MANU.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Project Bulgarian Academy of Science - Macedonian Academy of Sciences and Arts [Grant Number Project BAS – MANU].

References

- Alvarez-Ros, M. C., & Alcolea Palafox, M. (2014). Conformational analysis, molecular structure and solid state simulation of the antiviral drug acyclovir (Zovirax) using density functional theory methods. *Pharmaceuticals*, 7, 695–722.
- Avcioglu, M. V., & Golcu, A. (2015). Synthesis of acyclovir metal complexes: Spectral, electrochemical, thermal, and DNA binding studies. *Synthesis and Reactivity in Inorganic, Metal-Organic, and Nano-Metal Chemistry*, 45, 581–590.
- Barceló-Oliver, M., Terrón, A., García-Raso, A., Fiol, J. J., Molins, E., & Miravittles, C. (2004). Ternary complexes metal [Co(II), Ni(II), Cu(II) and Zn(II)] – ortho-iodohippurate (I-hip) – acyclovir. X-ray characterization of isostructural [(Co, Ni or Zn)(I-hip)2(ACV)(H2O)3] with stacking as a recognition factor. *Journal of Inorganic Biochemistry*, 98, 1703–1711.

- Bebenek, K., Pedersen, L. C., & Kunkel, T. A. (2011). Replication infidelity via a mismatch with Watson-Crick geometry. *Proceedings of the National Academy of Sciences*, 108, 1862–1867.
- Birnbaum, G. I., Cygler, M., & Shugar, D. (1984). Conformational features of acycloclonucleosides: Structure of acyclovir, an antitherpes agent. *Canadian Journal of Chemistry*, 62, 2646–2652.
- Brandi-Blanco, M. D. P., Choquesillo-Lazarte, D., Domínguez-Martín, A., González-Pérez, J. M., Castiñeiras, A., & Niclós-Gutiérrez, J. (2011). Metal ion binding patterns of acyclovir: Molecular recognition between this antiviral agent and copper(II) chelates with iminodiacetate or glycylglycinate. *Journal of Inorganic Biochemistry*, 105, 616–623.
- Breneman, C. M., & Wiberg, K. B. (1990). Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *Journal of Computational Chemistry*, 11, 361–373.
- Brovarets, O. O., & Hovorun, D. M. (2015a). How many tautomerization pathways connect Watson-Crick-like G•T DNA base mispair and wobble mismatches? *Journal of Biomolecular Structure and Dynamics*, 33, 2297–2315.
- Brovarets, O. O., & Hovorun, D. M. (2015b). 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. *Journal of Biomolecular Structure and Dynamics*, 33, 925–945.
- Brovarets, O. O., & Hovorun, D. M. (2015c). The physicochemical essence of the purine-pyrimidine transition mismatches with Watson-Crick geometry in DNA: A•C* versa A*•C. A QM and QTAIM atomistic understanding. *Journal of Biomolecular Structure and Dynamics*, 33, 28–55.
- Brovarets, O. O., & Hovorun, D. M. (2015d). Tautomeric transition between wobble A•C DNA base mispair and Watson-Crick-like A•C* mismatch: Microstructural mechanism and biological significance. *Physical Chemistry Chemical Physics: PCCP*, 17, 15103–15110.
- Brovarets, O. O., Zhurakivsky, R. O., & Hovorun, D. M. (2014). Does the tautomeric status of the adenine bases change upon the dissociation of the A*•Asyn Topal-Fresco DNA mismatch? A combined QM and QTAIM atomistic insight. *Physical Chemistry Chemical Physics*, 16, 3715–3725.
- Cossi, M., Rega, N., Scalmani, G., & Barone, V. (2003). Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *Journal of Computational Chemistry*, 24, 669–681.
- Coutinho, K., & Canuto, S. (1997). *A Monte Carlo program for molecular liquid simulation*. In DICE. Brazil: University of São Paulo.
- Cramer, C. J., & Truhlar, D. G. (2008). A Universal Approach to Solvation Modeling. *Accounts of Chemical Research*, 41, 760–768.
- Faller, B., & Ertl, P. (2007). Computational approaches to determine drug solubility. *Advanced Drug Delivery Reviews*, 59, 533–545.
- Francel, M. M., Pietro, W. J., Hehre, W. J., Binkley, J. S., Gordon, M. S., DeFrees, D. J., & Pople, J. A. (1982). Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *The Journal of Chemical Physics*, 77, 3654–3665.
- Gavira, J. M., De La Fuente, M., Navarro, R., & Hernanz, A. (1997). Normal coordinate analysis of acycloguanosine. *Journal of Molecular Structure*, 410–411, 425–429.
- Gonzales, C., & Schlegel, H. B. (1989). An improved algorithm for reaction path following. *The Journal of Chemical Physics*, 90, 2154–2161.
- Gorb, L., & Leszczynski, J. (1997). Ab initio prediction of the geometry and IR frequencies of the mono- and dihydrated complexes of the oxo-amino-tautomers of guanine. *International Journal of Quantum Chemistry*, 65, 759–765.
- Gorb, L., & Leszczynski, J. (1998). Intramolecular proton transfer in mono- and dihydrated tautomers of guanine: An ab initio post Hartree-Fock study. *Journal of the American Chemical Society*, 120, 5024–5032.
- Gordon, M. S., & Schmidt, M. W. (2005). Advances in electronic structure theory: GAMESS a decade later. In C. E. Dykstra, G. Frenking, K. S. Kim, & G. E. Scuseria (Eds.), *Theory and applications of computational chemistry: The first forty years* (pp. 1167–1189). Amsterdam: Elsevier.
- Guo, Z., Xue, J., Ke, Z., Phillips, D. L., & Zhao, C. (2009). Influence of water hydrogen bonding on the reactions of aryl nitrenium ions with guanosine: Hydrogen-bonding effects can favor reaction at the C8 site. *The Journal of Physical Chemistry B*, 113, 6528–6532.
- Hariharan, P. C., & Pople, J. A. (1973). The influence of polarization functions on molecular orbital hydrogenation energies. *Theoretica Chimica Acta*, 28, 213–222.
- Harris, V. H., Smith, C. L., Cummins, W. J., Hamilton, A. L., Adams, H., Dickman, M., ... Williams, D. M. (2003). The effect of tautomeric constant on the specificity of nucleotide incorporation during DNA replication: Support for the rare tautomer hypothesis of substitution mutagenesis. *Journal of Molecular Biology*, 326, 1389–1401.
- Jiang, Y.-C., Feng, H., Lin, Y.-C., & Guo, X.-R. (2016). New strategies against drug resistance to herpes simplex virus. *International Journal of Oral Science*, 8(1), 1–6.
- Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118, 11225–11236.
- Kochina, O. S., Zhurakivsky, R. O., & Hovorun, D. M. (2008). Impact of the tautomerisation of the nucleotide bases on the conformational properties of the nucleosides: Quantum-mechanical study using density functional method. *Representative of National Academy of Science Ukraine*, 181–186.
- Lutker, K. M., Quiñones, R., Xu, J., Ramamoorthy, A., & Matzger, A. J. (2011). Polymorphs and hydrates of acyclovir. *Journal of Pharmaceutical Sciences*, 100, 949–963.
- Markova, N., Pejov, L., & Enchev, V. (2015). A Hybrid statistical mechanics-quantum chemical model for proton transfer in 5-azauracil and 6-azauracil in water solution. *International Journal of Quantum Chemistry*, 115, 477–485.
- Orozco, M., Cubero, E., Barril, X., Colominas, C., & Luque, F. J. (1999). Nucleic acid bases in solution. *Journal of Theoretical and Computational Chemistry*, 8, 119–166.
- Plass, M., Kristl, A., & Abraham, M. H. (1999). Spectroscopic investigation of the tautomeric equilibria in the guanine derivatives of acyclovir. *Journal of the Chemical Society, Perkin Transactions 2*, 2, 2641–2646.
- Saenger, W. (1984). *Principles of Nucleic Acid Structure* (1st ed.). New York, NY: Springer-Verlag.
- Saigusa, H., Mizuno, N., Asami, H., Takahashi, K., & Tachikawa, M. (2008). Ultraviolet spectroscopy and theoretical calculations of mono- and dihydrated clusters of the guanine nucleosides: Possibility of different hydration structures for guanosine and 2'-deoxyguanosine. *Bulletin of the Chemical Society of Japan*, 81, 1274–1281.

- Saigusa, Hiroyuki, Urashima, S. H., & Asami, H. (2009). IR-UV double resonance spectroscopy of the hydrated clusters of guanosine and 9-methylguanine: Evidence for hydration structures involving the sugar group. *The Journal of Physical Chemistry A*, 113, 3455–3462.
- Schmidt, M. W., Baldridge, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., Koseki, S., ... Montgomery, J. A., Jr (1993). General atomic and molecular electronic structure system. *Journal of Computational Chemistry*, 14, 1347–1363.
- Shukla, M. K., & Leszczynski, J. (2013). Tautomerism in nucleic acid bases and base pairs: A brief overview. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 3, 637–649.
- Thewalt, U., Bugg, C. E., & Marsh, R. E. (1970). The crystal structure of guanosine dihydrate and inosine dihydrate. *Acta Crystallographica Section B Structural Crystallography and Crystal Chemistry*, 26, 1089–1101.
- Tomasi, J., Mennucci, B., & Cammi, R. (2005). Quantum mechanical continuum solvation models. *Chemical Reviews*, 105, 2999–3094.
- Topal, M. D., & Fresco, J. R. (1976). Complementary base pairing and the origin of substitution mutations. *Nature*, 263, 285.
- Wang, W., Hellinga, H. W., & Beese, L. S. (2011). Structural evidence for the rare tautomer hypothesis of spontaneous mutagenesis. *Proceedings of the National Academy of Sciences*, 108, 17644–17648.
- Watson, J. D., & Crick, F. H. C. (1953). Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature*, 171, 737–738.
- Yoneda, S., Sugawara, Y., & Urabe, H. (2005). Crystal water dynamics of guanosine dihydrate: Analysis of atomic displacement parameters, time profile of hydrogen-bonding probability, and translocation of water by MD simulation. *The Journal of Physical Chemistry B*, 109, 1304–1312.
- Zhurakivsky, R. O., & Hovorun, D. M. (2007). Guanosine conformational possibilities: The DFT quantum chemical investigation. *Physics of the Alife*, 15, 24–34.