

Correlated *ab Initio* Study of Nucleic Acid Bases and Their Tautomers in the Gas Phase, in a Microhydrated Environment and in Aqueous Solution. Guanine: Surprising Stabilization of Rare Tautomers in Aqueous Solution

Michal Hanus,[†] Filip Ryjáček,[†] Martin Kabeláč,[†] Tomáš Kubař,^{†,||}
Tetyana V. Bogdan,^{‡,§} Semen A. Trygubenko,^{‡,§} and Pavel Hobza^{*,†}

Contribution from the J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic and Center for Complex Molecular Systems and Biomolecules, 182 23, Prague 8, Czech Republic and Department for Physical and Mathematical Sciences, The National University of Kyiv–Mohyla Academy, Kyiv 04070, Ukraine

Received January 20, 2003; E-mail: hobza@indy.jh-inst.cas.cz.

Abstract: Altogether eight keto and enol tautomers of guanine were studied theoretically in the gas phase, in a microhydrated environment (1 and 2 water molecules) and in bulk water. The structures of isolated, as well as mono- and dihydrated tautomers were determined by means of the RI-MP2 method using the extended TZVPP (5s3p2d1f/3s2p1d) basis set. The relative energies of isolated tautomers included the correction to higher correlation energy terms evaluated at the CCSD(T)/aug-cc-pVDZ level. The relative enthalpies at 0 K and relative free energies at 298 K were based on the above-mentioned relative energies and zero-point vibration energies, temperature-dependent enthalpy terms and entropies evaluated at the MP2/6-31G** level. The keto form having hydrogen atom at N7 is the global minimum while the canonical form having hydrogen atom at N9 represents the first local minimum at all theoretical levels in vacuo and in the presence of 1 and 2 water molecules. All three unusual rare tautomers having hydrogens at N3 and N7, at N3 and N9, and also at N9 and N7 are systematically considerably less stable and can be hardly detected in the gas phase. The theoretical predictions fully agree with existing theoretical as well as experimental results. The effect of bulk solvent on the relative stability of guanine tautomers was studied by self-consistent reaction field and molecular dynamics free energy calculations using the thermodynamic integration method. Bulk solvent, surprisingly, strongly favored these three rare tautomers over all remaining low-energy tautomers and probably only these forms can exist in water phase. The global minimum (tautomer with hydrogens at N3 and N7) is by 13 kcal/mol more stable than the canonical form (3rd local minimum). Addition of one or two water molecules does not change the relative stability order of isolated guanine tautomers but the respective trend clearly supports the surprising stabilization of three rare forms.

1. Introduction

Different tautomers of nucleic acid (NA) bases are obtained when considering different positions of hydrogen around the base. Rare tautomers may be involved in various biochemical processes including the point mutations.¹ Nevertheless, direct evidence of their presence in biomolecules is rather rare and NA bases are dominantly present in a most stable canonical form. The situation is different in gas-phase experiments where various tautomers coexist. Extrapolation of (rather academical) gas-phase results to more biologically relevant water-phase results requires description of a solvent effect. Water might play

a dominant role in these highly polar systems and can dramatically change the relative stability of various tautomers. Many theoretical studies on the tautomeric equilibria of NA bases exist^{2–13} and attention was particularly paid to cytosine. For

[†] J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic and Center for Complex Molecular Systems and Biomolecules.

[‡] The National University of Kyiv–Mohyla Academy.

[§] Current address: University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, United Kingdom.

^{||} Undergraduate—Charles University, Faculty of Science, Prague.

(1) Florian, J.; Leszczynski, J. *J. Am. Chem. Soc.* **1996**, *118*, 3010.

(2) (a) Hobza, P.; Šponer, J. *Chem. Rev.* **1999**, *99*, 3247. (b) Šponer, J.; Leszczynski, J.; Hobza, P. *Biopolymers* **2002**, *61*, 3.

(3) (a) Fogarasi, G. *J. Mol. Struct.* **1997**, *413*, 271. (b) Aleman, C. *Chem. Phys.* **2000**, *253*, 13. (c) Les, A.; Adamowicz, L.; Bartlett, R. J. *J. Phys. Chem.* **1989**, *93*, 4001. (d) Estrin, D. A.; Paglieri, L.; Corongiu, G. *J. Phys. Chem.* **1994**, *98*, 5653. (e) Ha, T. K.; Keller, H. J.; Gunde, R.; Gunthard, H. H. *J. Phys. Chem. A* **1999**, *103*, 6612. (f) Russo, N.; Toscano, M.; Grand, A. *J. Phys. Chem. B* **2001**, *105*, 4735. (g) Russo, N.; Toscano, M.; Grand, A. *J. Am. Chem. Soc.* **2001**, *123*, 10 272. (h) Nowak, M. J.; Lapinski, L.; Fulara, J. *Spectrochim. Acta Part A* **1989**, *45*, 229. (i) Fogarasi, G. *J. Phys. Chem. A* **2002**, *106*, 1381.

(4) van Mourik, T.; Benoit, D. M.; Price, S. L.; Clary, D. C. *Phys. Chem. Chem. Phys.* **2000**, *2*, 1281.

(5) Clary, D. C.; Benoit, D. M.; van Mourik, T. *Acc. Chem. Res.* **2000**, *33*, 441.

(6) Kobayashi, R. *J. Phys. Chem. A* **1998**, *102*, 10 813.

(7) Sambrano, J. R.; Souza, A. R.; Queralt, J. J.; Andrés, J. *Chem. Phys. Lett.* **2000**, *317*, 437.

(8) Kryachko, E. S.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **2001**, *105*, 1288.

cytosine, an agreement¹³ exists that besides the canonical form, two enol and two imino forms are energetically similar and, therefore, they should coexist in the gas phase. Bulk water changes the relative stabilities significantly and the canonical form is now unambiguously favored over other structures.¹³ The situation with guanine is different since it is the nucleobase with the largest number of tautomeric forms (oxo-hydroxy, amino-imino, N1H–N3H, N7H–N9H). Many theoretical studies were devoted to tautomeric equilibria of guanine^{14–20} and calculations mostly show the existence of four low-energy tautomers (amino-oxo and amino-hydroxy) with the amino-oxo form having hydrogen at N7 as their most stable gas-phase tautomer. The imino tautomers were shown to be energetically less stable.¹⁹ Experimental results fully agree with the theoretical data and four tautomeric forms of guanine were recently detected in the gas phase.²¹ Besides these low-energy tautomers, other ones also exist but their energies and, also, free energies are considerably higher. Similarly, as in the case of cytosine, the water phase significantly destabilizes the enol forms and favors the keto forms.¹⁸ However, the authors of the cited study considered only the five lowest energy tautomers and evaluated the hydration free energy only for them. It is, however, possible to expect that hydration free energy will be for higher-energy tautomers much larger than for other ones that can result in favorable free energy. This is not speculation. Recently, an interesting study was presented²² showing that the most stable tautomer among gas-phase Al-guanine complexes is the one in which both of the nitrogen atoms in the five-membered ring are bound to hydrogen. This tautomer is energetically clearly unfavorable but its interaction with aluminum atom overcomes the energy (free energy) barrier. It must be mentioned that there was no prior theoretical or experimental observation of such isomer.

In this paper, we present a study of the tautomerism of guanine. Gas-phase calculations of the relative stability of different tautomers are accompanied by the microhydration studies as well as studies taking the solvent effects of the bulk water fully into consideration. These calculations, performed systematically for all guanine tautomers, yield a final insight into tautomeric equilibrium of guanine forms in the gas phase, microhydrated environment as well as in bulk water.

Although in our previous paper,¹³ the thermodynamics parameters were calculated using the BOSS²³ and AMBER²⁴

packages with standard potential energy functions, in the present study, we adopted GROMACS molecular modeling package^{25,26} modified by our own code with thermodynamic integration (TI)²⁷ method and soft core potential energy scaling.²⁸

2. Methods

Computational Strategy. The gas-phase geometry of guanine tautomers was determined using the resolution of identity MP2 (RI-MP2) procedure²⁹ using the double-polarized triple- ζ (TZVPP) basis set [5s3p2d1f/3s2p1d] and the default auxiliary basis set. The relative energies of tautomers were evaluated at the same level as well as at higher CCSD(T) level covering higher portion of correlation energy. Relative free energies were obtained by including the zero-point vibration energies (ZPVE), temperature-dependent enthalpy terms and entropies. A molecular dynamics/quenching technique with the Cornell et al. empirical potential³⁰ was utilized to explore the potential energy surface (PES) of the guanine...water and guanine...(water)₂ clusters. Four energetically most stable structures were later recalculated at the RI-MP2/TZVPP level. Relative stability of various guanine tautomers in a water environment was deduced from the free energies determined as a sum of gas-phase free energies and free energies of hydration. The latter energies were estimated using a self-consistent reaction field or molecular dynamics free energy calculations.

Quantum Chemical Calculations. Energetical and geometrical characteristics of eight guanine tautomers in the gas phase were investigated using the RI-MP2 procedure. Efficient handling of four-center integrals used in the RI-MP2 method allowed decreasing the computational cost by approximately the order of magnitude in comparison with the exact MP2 with almost no loss in accuracy, as it was shown for hydrogen-bonded and stacked DNA base pairs.³¹ Besides the RI-MP2 method, also the exact MP2 and CCSD(T) methods were applied using the aug-cc-pVDZ basis set. The respective calculations were performed for the RI-MP2/TZVPP geometries. Harmonic vibration analysis providing the ZPVE, thermodynamic characteristics (via partition functions) as well as a character of the stationary point found was performed at the MP2/6-31G** level.

Relative gas-phase free energies were determined as a sum of relative energy changes, ZPVE, temperature-dependent enthalpies, and entropies. Relative energy changes were calculated at the RI-MP2/TZVPP level. Thermodynamic functions were taken from partition functions computed with MP2/6-31G** characteristics (geometry, vibrations) according to rigid rotor-harmonic oscillator-ideal gas approximation. We believe that the use of Hartree–Fock (HF) characteristics is not sufficient since the HF level does not describe correctly the nonplanarity of the guanine amino group and the respective vibration modes (which can play an important role in evaluation of vibration entropy) can be thus far from the accurate values.

Interaction energy of guanine...water and guanine...(water)₂ complexes was determined at the RI-MP2/TZVPP level with inclusion of basis set superposition error³² and deformation energy.

- (9) Kryachko, E. S.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **2001**, *105*, 1934.
 (10) Chandra A. K.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Mol. Struct.* **2000**, *519*, 1.
 (11) Carles, S.; Lecomte, F.; Schermann, J. P.; Desfrancois, C. *J. Phys. Chem. A* **2000**, *104*, 10 622.
 (12) Gorb, L.; Leszczynski, J. *Int. J. Quantum Chem.* **1998**, *70*, 4.
 (13) Trygubenko, S. A.; Bogdan, T. V.; Rueda, M.; Orozco, M.; Luque, F. J.; Spomer, J.; Slavíček, P.; Hobza, P. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4192.
 (14) Leszczynski, J. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Ed.; John Wiley: Chichester, 1998; 2951.
 (15) Szczepaniak, K.; Szczesniak, M. *J. Mol. Struct.* **1987**, *156*, 29.
 (16) Dolgounitcheva, O.; Zakrzewski, V. G.; Ortiz, J. V. *J. Am. Chem. Soc.* **2000**, *122*, 12 304.
 (17) Leszczynski, J. *J. Phys. Chem. A* **1998**, *102*, 2357.
 (18) Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811.
 (19) Sabio, M.; Topiol, S.; Lumma, W. C. *J. Phys. Chem.* **1990**, *94*, 1366.
 (20) Szczepaniak, K.; Szczepaniak, M.; Szaida, W.; Person, W. B.; Leszczynski, J. *Can. J. Chem.* **1991**, *69*, 1718.
 (21) (a) Piuze, F.; Mons, M.; Dimicoli, I.; Tardivel, B.; Zhao, Q. *Chem. Phys.* **2001**, *270*, 205. (b) Mons, M.; Dimicoli, I.; Piuze, F.; Tardivel, B.; Elhamine, M. *J. Phys. Chem. A* **2002**, *106*, 5088.
 (22) Pedersen, D. B.; Simard, B.; Martinez, A.; Moussatova, A. *J. Am. Chem. Soc.*, submitted.
 (23) Jorgensen, W. L. *BOSS v. 4.2*; Yale University: New Haven, 1999.

- (24) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K.; Stanton, R.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Ferguson, D. M.; Radmer, R. J.; Siebel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 6*; University of California: San Francisco, CA, 1997.
 (25) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Comm.* **1995**, *91*, 43.
 (26) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Mod.* **2001**, *7*, 306.
 (27) Pearlman, D. A.; Rao, B. G. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Ed.; John Wiley: Chichester, 1998; 1036.
 (28) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529.
 (29) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. *Chem. Phys. Lett.* **1993**, *208*, 359.
 (30) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. E.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
 (31) Jurečka, P.; Nachtigall, P.; Hobza, P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 4578.
 (32) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1989**, *214*, 15.

All calculations were performed with GAUSSIAN98,³³ MOLPRO³⁴ and TURBOMOLE³⁵ suites of programs.

Molecular Dynamics/Quenching (MD/Q) Technique. Interaction of guanine with microhydrated environment was investigated using molecular dynamics/quenching technique with the modified Cornell et al. potential.³⁰ MD/Q simulations were performed in the NVE microcanonical ensemble (N, V, E mean number of particles, volume, and energy) within quaternion formalism. The respective code uses a fifth-order predictor-corrector algorithm with a 0.5 fs integration step. The MD simulations were performed at a constant total energy corresponding to the average temperature of 298 K which is high enough to allow crossing over relatively high energy barriers and thus to sample the whole potential energy surface. Every 1 ps, the MD run was interrupted, the kinetic energy was removed and the structure of the cluster of guanine with one or two water molecules was optimized using the conjugate gradient method, the geometry, and the energy stored, whereas the MD run was restarted from the point where it had been interrupted. The total simulation time was 250 ns.

Employment of the molecular dynamics techniques involved modification of canonical guanosine libraries in the Cornell et al. force field³⁰ to include new atomic types and their parametrization of all guanine tautomers. Constants for geometrical parameters most strongly affected by tautomerization (i.e., bond lengths, valence and dihedral angles of the enol group) were derived from quantum chemical calculations. The torsion constants, being the most important for precise description of the hydrogen bond, were parametrized using Hopfinger and Pearlstein differential fitting procedure.³⁶ The atomic charges of the tautomers were generated with an electrostatic potential fitting procedure³⁷ (RESP) at the HF/6-31G* level. For the dihedral angles of hydroxy group a set of thirteen constrained geometries was calculated at the HF/6-31G* level. Then, we calculated the molecular mechanical torsion curves for the same values of dihedrals as used in the quantum chemical calculations, with zero dihedral force constants. The desired differential curve was obtained as a difference between these two curves. The differential curve was fitted with the least-squares method (using software Mathematica³⁸) and the coefficients obtained were built into the AMBER libraries.

Free Energy Calculations. Molecular Dynamics–Thermodynamic Integration (MD-TI) technique was utilized to calculate relative free energy of hydration (ΔG^{HYD}) between guanine tautomers. All calculations were performed using GROMACS molecular modeling package^{25,26} with our own code implementing the TI method (taken from ref 27). van der Waals parameters for guanine originate from the Cornell et al. force field³⁰ and were converted from AMBER to GROMACS using our own code AMBCONV.³⁹ To avoid instability of simulations and incorrect accumulation of free energy during the TI perturbations arising from singularities in van der Waals and Coulombic potential functions, soft core potential scaling²⁸ was systematically used.

The rectangular periodic box of approximate size $25 \times 26 \times 20 \text{ \AA}$ was filled with about 470 TIP3P⁴⁰ water molecules and all perturbations were performed in the NPT ensemble at 1 atm and 298.15 K. Time step was set to 1 fs and nonbonded cutoff of 9 \AA was used for van der Waals and Coulombic interactions. Before perturbations a standard equilibration run was performed: 100 ps run; 1 fs time step; 9 \AA cutoff. Berendsen temperature and pressure coupling schemes⁴¹ were systematically used with weak coupling constant and the center of mass motion and rotation were periodically removed.⁴² Shorter perturbations (2 ns) were carried out in 251 and longer perturbations (6 ns, 11 ns) in 501 sampling windows. The number of configurations used in equilibration part and in data collection part is shown in Table 1. Calculations in vacuo (reference states) were performed also according to a standard protocol: 1 fs time step; 100 ps simulation time. SETTLE⁴³ algorithm was used for water molecules, no constraint algorithms were applied to solute.

To ensure real independence of free energies on soft core parameter α (α_{SC}), several calculations with different α_{SC} values were performed. Parameter α_{SC} affects the minimum effective distance⁴⁴ (r_{eff}) and height⁴⁵ at $r = 0$ (V_x) according to the following equations

$$r_{\text{eff}} = \sigma_x \sqrt{\lambda^2 \alpha} \quad (1)$$

$$V_x(r=0) = 4\epsilon_x \left(\frac{1 - \lambda^2 \alpha}{\lambda^4 \alpha^2} \right) \quad (2)$$

where σ_x is scaled⁴⁶ as an arithmetical and ϵ_x as a geometrical average of atomic vdW parameters σ_{ii} and σ_{ij} , respectively ϵ_{ii} and ϵ_{ij} , and λ is standard coupling parameter. Critical value of V_x in aqueous solution is 7.5 kcal/mol, which is comparable with the mean kinetic energy of water molecules at 298 K. At that point of perturbation water molecules can effectively penetrate volume delimited by van der Waals radius of perturbed atom, which could affect the configuration space sampling and also, indirectly, the total free energy estimation.⁴⁷ The following values of α_{SC} were used:

- (i) $\alpha_{\text{SC}} = 0.6$; water can penetrate at the very end of atom annihilation ($\lambda = 0.99$, “harder” soft core);
- (ii) $\alpha_{\text{SC}} = 1.51$; effective penetration of water even in the middle of simulation ($\lambda = 0.50$, “softer” soft core);
- (iii) $\alpha_{\text{SC}} = 1$ seems to be a good compromise between previously mentioned values and will be used in the present study. It also fully corresponds with recommendations listed in the GROMACS manual.⁴⁶

Another parameter, which is critical for correct TI calculation, is a correct sampling of configuration space.⁴⁸ 2, 6, and 11 ns forward and backward runs were used to ensure convergence of free energy and to estimate error limits.

Relative free energy of hydration of guanine (1,7) tautomer (related to canonical (1,9) tautomer) can be determined from thermodynamic cycle shown below (Scheme 1): From the cycle it is apparent that

$$\Delta G_{(1,9)}^{\text{HYD}} + \Delta G_{(1,9 \rightarrow 1,7)}^{\text{BULK}} = \Delta G_{(1,9 \rightarrow 1,7)}^{\text{VAC}} + \Delta G_{(1,7)}^{\text{HYD}} \quad (3)$$

It means that relative free energy of hydration can be determined as follows

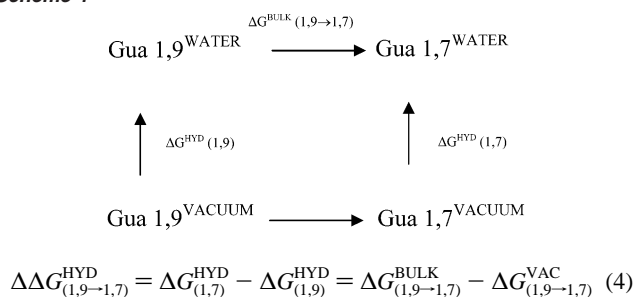
- (33) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.7; Gaussian, Inc.: Pittsburgh, PA, 1998.
- (34) *Molpro, a package of ab initio programs designed by M.-J. Werner*, P. J. Knowles, version 2002.1, R. D. Amos et al.
- (35) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165.
- (36) Hopfinger, A. J.; Pearlstein, R. A. *J. Comput. Chem.* **1984**, *5*, 486.
- (37) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.
- (38) Mathematica, Wolfram Research, Inc.
- (39) http://www.gromacs.org/contributions/uploaded_contributions/ambconv.tgz.

- (40) Jorgensen, W. L.; Chandrasekhar, J. *J. Phys. Chem.* **1983**, *79*, 926.
- (41) Berendsen, J. *Chem. Phys.* **1984**, *34*, 3684.
- (42) Harvey, S. C.; Tan, R. K. Z.; Cheatham, T. E. *J. Comput. Chem.* **1998**, *19*, 726.
- (43) Miyamoto S.; Kollman P. A. *J. Comput. Chem.* **1992**, *13*, 952.
- (44) Pitera, J. W.; Van Gunsteren, W. F. *Molecular Simulation* **2002**, *28*, 45.
- (45) Schäfer, H.; Van Gunsteren, W. F.; Mark, E. A. *J. Comput. Chem.* **1999**, *20*, 1604.
- (46) GROMACS User Manual, version 3.1.1, www.gromacs.org/documentation/paper_manuals.php.
- (47) Villa, A.; Mark, E. A. *J. Comput. Chem.* **2002**, *23*, 548.
- (48) Van Gunsteren, W. F. *J. Chem. Phys.* **1998**, *108*, 6109.

Table 1. Thermodynamic Integration: Parameters of Simulation, Estimated Free Energy (ΔG) and Its Deviation (dev)

| system | env ^a | SC α ^b | NWIN ^c | TIME ^d | NEQUIL ^e | NCOLL ^f | ΔG ^g | dev ^h |
|----------------------------|------------------|--------------------------|-------------------|-------------------|---------------------|--------------------|-------------------------|------------------|
| (1,9) \rightarrow (1,7) | vacuum | 1.00 | 251 | 0.1 | 200 | 200 | -77.38 | ± 0.00 |
| | water | 0.60 | 251 | 2.0 | 4000 | 4000 | -78.24 | ± 0.00 |
| | water | 1.51 | 251 | 2.0 | 4000 | 4000 | -78.97 | ± 0.08 |
| | water | 1.00 | 251 | 2.0 | 4000 | 4000 | -78.06 | ± 0.40 |
| | water | 1.00 | 501 | 6.0 | 2000 | 10000 | -78.20 | ± 0.23 |
| | water | 1.00 | 501 | 11.0 | 2000 | 20000 | -78.29 | ± 0.17 |
| (1,9) \rightarrow (7,O2) | vacuum | 1.00 | 251 | 0.1 | 200 | 200 | -95.27 | ± 0.00 |
| | water | 0.60 | 251 | 2.0 | 4000 | 4000 | -91.90 | ± 0.04 |
| | water | 1.51 | 251 | 2.0 | 4000 | 4000 | -91.46 | ± 0.41 |
| | water | 1.00 | 251 | 2.0 | 4000 | 4000 | -92.29 | ± 0.17 |
| | water | 1.00 | 501 | 6.0 | 2000 | 10000 | -92.03 | ± 0.34 |
| | water | 1.00 | 501 | 11.0 | 2000 | 20000 | -92.17 | ± 0.01 |
| (1,9) \rightarrow (9,O2) | vacuum | 1.00 | 251 | 0.1 | 200 | 200 | -49.76 | ± 0.00 |
| | water | 0.60 | 251 | 2.0 | 4000 | 4000 | -43.39 | ± 0.25 |
| | water | 1.51 | 251 | 2.0 | 4000 | 4000 | -44.07 | ± 0.78 |
| | water | 1.00 | 251 | 2.0 | 4000 | 4000 | -43.12 | ± 0.57 |
| | water | 1.00 | 501 | 6.0 | 2000 | 10000 | -43.66 | ± 0.07 |
| | water | 1.00 | 501 | 11.0 | 2000 | 20000 | -44.18 | ± 0.05 |
| (1,9) \rightarrow (3,7) | vacuum | 1.00 | 251 | 0.1 | 200 | 200 | -54.78 | ± 0.02 |
| | water | 0.60 | 251 | 2.0 | 4000 | 4000 | -73.40 | ± 0.08 |
| | water | 1.51 | 251 | 2.0 | 4000 | 4000 | -73.49 | ± 0.16 |
| | water | 1.00 | 251 | 2.0 | 4000 | 4000 | -73.17 | ± 0.47 |
| | water | 1.00 | 501 | 6.0 | 2000 | 10000 | -73.68 | ± 0.44 |
| | water | 1.00 | 501 | 11.0 | 2000 | 20000 | -73.26 | ± 0.10 |
| (1,9) \rightarrow (3,9) | vacuum | 1.00 | 251 | 0.1 | 200 | 200 | -8.50 | ± 0.26 |
| | water | 0.60 | 251 | 2.0 | 4000 | 4000 | -32.76 | ± 0.35 |
| | water | 1.51 | 251 | 2.0 | 4000 | 4000 | -31.18 | ± 1.60 |
| | water | 1.00 | 251 | 2.0 | 4000 | 4000 | -32.78 | ± 0.40 |
| | water | 1.00 | 501 | 6.0 | 2000 | 10000 | -32.58 | ± 0.01 |
| | water | 1.00 | 501 | 11.0 | 2000 | 20000 | -32.90 | ± 0.00 |
| (1,9) \rightarrow (7,9) | vacuum | 1.00 | 251 | 0.1 | 200 | 200 | -84.85 | ± 0.02 |
| | water | 0.60 | 251 | 2.0 | 4000 | 4000 | -115.53 | ± 0.00 |
| | water | 1.51 | 251 | 2.0 | 4000 | 4000 | -115.50 | ± 0.00 |
| | water | 1.00 | 251 | 2.0 | 4000 | 4000 | -115.18 | ± 0.33 |
| | water | 1.00 | 501 | 6.0 | 2000 | 10000 | -115.96 | ± 0.36 |
| | water | 1.00 | 501 | 11.0 | 2000 | 20000 | -115.53 | ± 0.27 |

^a Environment according to thermodynamic cycle in Scheme 1. ^b Soft core parameter [cf. eqs 1 and 2]. ^c Number of sampling windows ($\Delta\lambda^{-1}$). ^d Simulation time in ns. ^e Number of equilibration steps in each window. ^f Number of data collection steps in each window. ^g Estimated average free energy (in kcal/mol) ΔG^{VAC} or ΔG^{BULK} according to Scheme 1 and eq 4. ^h Deviation of forward and reverse runs (defined as $1/2|\Delta G_{\text{fwd}} - \Delta G_{\text{rev}}|$).

Scheme 1

The total relative free energy of guanine tautomer was finally determined as a sum of relative gas-phase free energy and relative free energy of hydration (according to Scheme 1); in fact, the MD-TI relative gas-phase free energy from Scheme 1 should be replaced by more accurate ab initio one (see later).

The hydration free energies were besides the MD-TI method determined also by using the self-consistent reaction field using the COSMO formalism with 6-31G* basis set.⁴⁹

3. Results and Discussion

Gas-Phase Tautomers. We have examined the relative stability of two keto forms having hydrogen either at nitrogen N9 (1,9) or at nitrogen N7 (1,7), three enol forms having hydrogen at N9 (9,O1), (9,O2) or at N7 (7,O2), and three rare

tautomers having either hydrogens on both nitrogens (N7, N9) at five membered ring (7,9) or at N3 and N9 (3,9) or at N3 and N7 (3,7) (see Figure 1).

Energetically less favorable imino- forms¹⁹ were not considered. The geometries, rotational constants and dipole moments of all these tautomers are depicted in Table 2 while Table 3 shows their relative energies, enthalpies, and free energies.⁵⁰ From Table 2, it is evident that various tautomers have very different values of dipole moment. The smallest one belongs to (1,7) keto form, whereas the other keto form (1,9) possesses more than three times larger dipole moment. Let us remember here that the dipole moment of (1,9) keto form is the largest among all canonical NA bases. Surprisingly, however, (7,9) and (3,9) rare tautomers have even larger dipole moments, around 10D! These extreme values indicate considerable stabilization when these tautomers are exposed to a polar solvent like water. Furthermore, their interaction with several water molecules should be accompanied by significant stabilization. Concerning the relative energy it is evident that the (1,7) keto form is the global minimum, whereas the (1,9) keto form is only slightly less stable. Two isomers of enol form (9,O2), (9,O1) are energetically close and are again slightly less stable than both

(50) The RI-MP2 dipole moments presented in Table 1 differ only slightly (in tenths of D) from the HF/6-31G* ones used in RESP procedure, and, consequently, also from the dipole moments yielded by the empirical potential. This means that the electrostatic term of the empirical force field (dominant among nonbonded terms) is described adequately.

(49) Barone, V.; Cossi, M.; Tomasi, J. *J. Comput. Chem.* **1998**, *19*, 404.

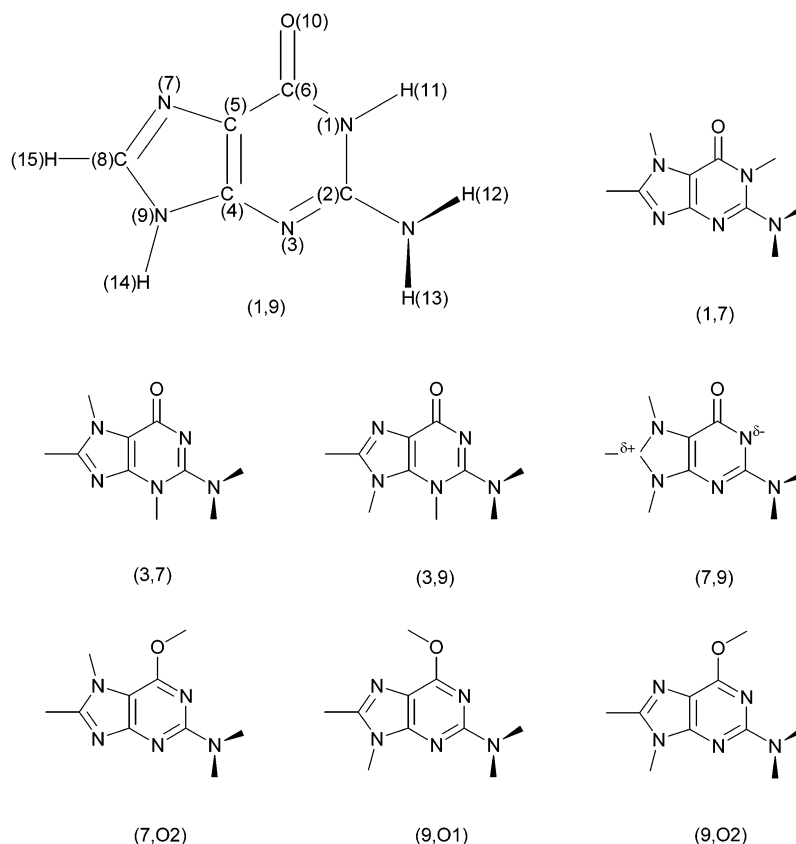


Figure 1. The eight stable guanine tautomers. Standard numbering and adopted nomenclature are presented.

Table 2. Geometries, Rotational Constants and Dipole Moments of Guanine Tautomers (RI-MP2/TZVPP level of theory)

| atom ^a taut. ^b | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | I ^c | μ^d |
|---|--------------|--------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|----------------|---------|
| (1,7) | x^e 0.0000 | 1.0654 | 2.2308 | 2.2970 | 1.1380 | 3.6231 | 4.2945 | 3.4890 | -1.2545 | 1.0609 | 4.0131 | 5.3701 | -1.1886 | -1.9504 | -0.8788 | 1.9209 | 1.87 |
| | y 0.0000 | 0.0000 | 0.0213 | 0.0090 | 0.0005 | 0.0282 | 0.0496 | 0.0449 | 0.0713 | 0.0053 | 0.0291 | 0.0680 | -0.0603 | -0.5291 | 0.0816 | 1.1279 | |
| | z 1.3776 | 2.1174 | 1.4019 | 0.0111 | -0.8222 | -0.3002 | 0.8883 | 1.9369 | 1.9585 | -2.0433 | -1.2285 | 0.9326 | 2.9565 | 1.5441 | -0.4921 | 0.7114 | |
| (9,O2) | 0.0000 | 1.0601 | 2.2024 | 2.3611 | 1.1571 | 3.6904 | 4.3222 | 3.4812 | -1.2404 | 1.1580 | 5.3928 | 3.7390 | -1.2648 | -1.9975 | 0.2264 | 1.9208 | 3.02 |
| | 0.0000 | 0.0000 | 0.0233 | 0.0378 | 0.0182 | 0.0624 | 0.0619 | 0.0369 | 0.0526 | 0.0147 | 0.0793 | 0.0375 | -0.2426 | -0.2658 | 0.0076 | 1.1361 | |
| | 1.3547 | 2.1603 | 1.4703 | 0.0795 | -0.6352 | -0.2809 | 0.8751 | 1.9603 | 1.9425 | -1.9739 | 0.9915 | 2.9327 | 2.9030 | 1.3638 | -2.2352 | 0.7142 | |
| (9,O1) | 0.0000 | 1.0526 | 2.1994 | 2.3517 | 1.1597 | 3.6755 | 4.3196 | 3.4846 | -1.2403 | 1.1502 | 2.0738 | 5.3903 | 3.7543 | -1.2768 | -2.0029 | 1.9277 | 3.75 |
| | 0.0000 | 0.0000 | 0.0316 | 0.0501 | 0.0229 | 0.0832 | 0.0862 | 0.0538 | 0.0511 | 0.0165 | 0.0366 | 0.1121 | 0.0559 | -0.2395 | -0.2308 | 1.1405 | |
| | 1.3563 | 2.1792 | 1.5009 | 0.1154 | -0.6189 | -0.2607 | 0.8915 | 1.9801 | 1.9370 | -1.9613 | -2.2475 | 1.0025 | 2.9490 | 2.8979 | 1.3476 | 0.7169 | |
| (7,O2) | 0.0000 | 1.0505 | 2.2273 | 2.3376 | 1.1586 | 3.6768 | 4.3056 | 3.4797 | -1.2475 | 1.2042 | 4.1113 | 5.3800 | -1.2688 | -1.9991 | 0.2824 | 1.9080 | 4.21 |
| | 0.0000 | 0.0000 | 0.0293 | 0.0504 | 0.0229 | 0.0772 | 0.0712 | 0.0411 | 0.0537 | 0.0158 | 0.1020 | 0.0911 | -0.2350 | -0.2803 | 0.0031 | 1.1389 | |
| | 1.3636 | 2.1702 | 1.5248 | 0.1268 | -0.6148 | -0.1448 | 1.0713 | 2.0959 | 1.9409 | -1.9630 | -1.0514 | 1.1433 | 2.9039 | 1.3643 | -2.2567 | 0.7136 | |
| (1,9) | 0.0000 | 1.0821 | 2.2050 | 2.3320 | 1.1372 | 3.6505 | 4.3127 | 3.4877 | -1.2365 | 0.9851 | 5.3847 | 3.7569 | -1.1764 | -1.9727 | -0.8805 | 1.9223 | 6.29 |
| | 0.0000 | 0.0000 | 0.0207 | 0.0156 | 0.0063 | 0.0328 | 0.0456 | 0.0380 | 0.0663 | 0.0118 | 0.0621 | 0.0406 | -0.1109 | -0.4556 | 0.0634 | 1.1218 | |
| | 1.3657 | 2.0935 | 1.3276 | -0.0582 | -0.8593 | -0.4395 | 0.7010 | 1.7957 | 1.9718 | -2.0657 | 0.7978 | 2.7653 | 2.9620 | 1.5247 | -0.4924 | 0.7090 | |
| (3,7) | 0.0000 | 0.6794 | 2.0428 | 2.7517 | 2.0679 | 0.6267 | 3.0518 | 4.2416 | 4.0955 | 0.0121 | 2.4980 | -0.9768 | 0.4238 | 2.8965 | 5.1847 | 1.9481 | 4.34 |
| | 0.0000 | 0.0000 | -0.0449 | -0.1292 | -0.1431 | -0.0944 | -0.2274 | -0.2607 | -0.2009 | -0.1229 | -0.1927 | 0.2095 | 0.5233 | -0.2584 | -0.3276 | 1.1196 | |
| | 2.3109 | 1.1997 | 1.0916 | 2.2646 | 3.4603 | 3.5571 | 4.4057 | 3.7471 | 2.4249 | 4.6153 | 0.2057 | 0.1321 | -0.7485 | 5.4006 | 4.2602 | 0.7116 | |
| (3,9) | 0.0000 | 1.1217 | 2.3102 | 2.4070 | 1.1941 | 3.7316 | 4.4037 | 3.5817 | -1.1889 | 1.1591 | 5.4744 | 3.8747 | 1.0225 | -1.2283 | -1.9763 | 1.9444 | 10.54 |
| | 0.0000 | 0.0000 | 0.1044 | 0.1321 | 0.1014 | 0.1861 | 0.1876 | 0.1327 | 0.0411 | 0.1463 | 0.2336 | 0.0975 | 0.3750 | -0.5643 | -0.0853 | 1.1072 | |
| | 1.2883 | 2.1048 | 1.4252 | 0.0512 | -0.7605 | -0.3248 | 0.8036 | 1.9117 | 1.9943 | -1.9741 | 0.9027 | 2.8732 | 3.0348 | 2.7992 | 1.3763 | 0.7065 | |
| (7,9) | 0.0000 | 1.0414 | 2.1733 | 2.3242 | 1.1598 | 3.6818 | 4.3356 | 3.4609 | -1.2222 | 1.2630 | 4.0772 | 5.3978 | 3.6901 | -1.2645 | -1.9932 | 1.9893 | 9.12 |
| | 0.0000 | 0.0000 | 0.0241 | 0.0423 | 0.0230 | 0.0620 | 0.0593 | 0.0353 | 0.0238 | 0.0234 | 0.0770 | 0.0740 | 0.0312 | -0.2520 | -0.2117 | 1.1385 | |
| | 1.3183 | 2.1766 | 1.5084 | 0.1568 | -0.7096 | -0.1048 | 1.0214 | 2.0268 | 1.9133 | -1.9152 | -1.0211 | 1.1299 | 2.9965 | 2.8670 | 1.3315 | 0.7243 | |

^a Atom 1 is placed at the origin. ^b For atom numbering and tautomer classification cf. Figure 1. ^c Rotational constants A, B, C in GHz. ^d Dipole moments in Debye. ^e Coordinates are in Å.

keto forms; the energy gap is larger for the second enol form (7,O2). Rare tautomer (3,7) is only by about 6 kcal/mol less stable than the canonical form. Very large energy difference of about 19 and 20 kcal/mol was found for remaining rare tautomers (7,9) and (3,9).

Let us remember that the above-mentioned results were obtained from the RI-MP2/TZVPP calculations. In our previous

paper on cytosine tautomers,¹³ we have shown that these relative energies were accurate and very close to the MP2/aug-cc-pVTZ values. To estimate the role of higher correlation energy contributions we performed the CCSD(T) calculations using the aug-cc-pVDZ basis set. From Table 3, it is apparent that the MP2 and CCSD(T) relative energies evaluated with this basis set are very close. The largest difference was found for (3,9)

Table 3. Relative Energies (ΔE), Zero Point Vibration Energies ($\Delta ZPVE$), Enthalpies (ΔH) and Free Energies (ΔG) (in kcal/mol) of Guanine Tautomers in the Gas Phase

| method structure ^a | ΔE (RI-MP2) ^a | ΔE (MP2) ^b | ΔE [CCSD(T)] ^c | $\Delta ZPVE$ ^d | $\Delta(G_0^{298} - E)$ ^d | ΔE ^e | ΔH_0^{0f} | ΔG_0^{298g} |
|----------------------------------|----------------------------------|-------------------------------|-----------------------------------|----------------------------|--------------------------------------|-------------------------|--------------------|---------------------|
| (1,7) | -0.50 | -0.71 | -0.69 | 0.07 | 0.04 | -0.48 | -0.41 | -0.44 |
| (9,O2) | 0.09 | 0.30 | 0.22 | -0.05 | 0.11 | 0.01 | 0.04 | 0.12 |
| (9,O1) | 0.29 | 0.58 | 0.62 | -0.05 | 0.09 | 0.34 | 0.29 | 0.43 |
| (7,O2) | 3.10 | 3.10 | 3.01 | -0.25 | -0.16 | 3.01 | 2.76 | 2.85 |
| (1,9) | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| (3,7) | 5.77 | — | — | -0.13 | -0.29 | 5.77 ⁱ | 5.64 ⁱ | 5.48 ⁱ |
| (3,9) | 18.95 | 18.66 | 18.10 | -0.50 | -0.71 | 18.39 | 17.89 | 17.78 |
| (7,9) | 20.08 | — | — | -0.23 | -0.27 | 20.08 ⁱ | 19.85 ⁱ | 19.81 ⁱ |

^a RI-MP2/TZVPP/RI-MP2/TZVPP. ^b MP2/aug-cc-pVDZ/RI-MP2/TZVPP. ^c CCSD(T)/aug-cc-pVDZ/RI-MP2/TZVPP. ^d MP2/6-31G***. ^e $\Delta E^{MP2} + (\Delta E^{CCSD(T)} - \Delta E^{RI-MP2})$. ^f $\Delta E + \Delta ZPVE$. ^g $\Delta E + \Delta(G^{298} - E)$. ^h Cf. Figure 1. ⁱ Not corrected for higher correlation energy terms.

tautomer (0.56 kcal/mol, i.e., less than 3% of the respective energy difference). It is thus possible to conclude that relative gas-phase energies are only negligibly affected by the higher correlation energy contributions. To compare these values with experimental data it is necessary to include the ZPVE. It means to pass from relative energies to relative enthalpies at 0 K. From the fifth column of Table 3, it is evident that the ZPVE corrections are systematically small and only for the (3,9) tautomer is it larger than 0.3 kcal/mol. The relative stability of guanine tautomers is thus not affected when the ZPVE corrections were considered. The same is true when the temperature-dependent enthalpy terms and entropy are taken into account. The sixth column in Table 3 collects these changes and it is seen that most affected is the (3,9) isomer but the respective change is smaller than 0.8 kcal/mol.

Final relative energies (ΔE), relative enthalpies at 0 K (ΔH^0) and relative free energies at 298 K (ΔG^{298}) are summarized in the last three columns of Table 3.⁵¹ The relative energies were constructed from RI-MP2 values corrected for the [CCSD(T)-MP2] terms, relative enthalpies at 0 K were based on these energies and were corrected by the $\Delta ZPVE$ terms, and, finally, relative free energies were determined from relative energies and $\Delta(G^{298} - E)$ terms.

Nonplanarity of the Guanine Amino Group. It is well established that amino groups of isolated nucleic acid bases are nonplanar due to a partial sp^3 pyramidalization of their amino group nitrogen atoms and, further, that nonplanarity of guanine amino group is considerably larger than that in other NA bases.²

Amino group nonplanarity in all guanine tautomers, evaluated at the RIMP2/TZVPP level, is collected in Table 4 and can be now considered as new reference data. Amino group in all tautomers is strongly nonplanar and the largest nonplanarity (even larger than in canonical guanine) was detected for (1,7)

Table 4. Nonplanarity of the Guanine Tautomers' Amino and Enol Groups^a

| tautomer | N1C2NH | N3C2NH | ΣXNH | N1C6OH | C2N1C6O |
|----------|--------|--------|--------------|--------|---------|
| (1,7) | 44.4 | -10.9 | 337.4 | | |
| (9,O2) | 22.6 | -21.1 | 345.9 | 0.6 | 179.8 |
| (9,O1) | 20.2 | 20.8 | 347.6 | 179.9 | 179.7 |
| (7,O2) | 23.7 | -20.6 | 345.3 | 0.6 | 179.6 |
| (1,9) | 38.2 | -13.6 | 340.4 | | |
| (3,7) | 12.9 | -37.4 | 340.8 | | |
| (3,9) | 9.3 | -43.3 | 338.3 | | |
| (7,9) | 16.2 | -19.1 | 350.4 | | |

^a ΣXNH is the sum of the three amino group valence angles. Angles in degrees.

and (3,9) tautomers. The sum of XNH valence angles in the amino group is in these two cases the smallest one (and differs considerably from 360°) which reflects the fact the amino group is not planar. For these two tautomers the two amino hydrogen dihedral angles are also significantly nonequivalent which is explained by electrostatic repulsion between amino group hydrogen and H1 (H3) hydrogen. On the other hand, smallest nonplanarity was detected in (7,9) tautomer.

Microhydrated Tautomers. The MD/Q simulations on monohydrated and dihydrated tautomers yielded about 5 stable structures with one water and about 20 stable structures with two waters. The four energetically most stable structures were then studied using ab initio methods and their structures are presented in Figures 2 and 3. The stability of these structures decreases from left to right. The relative energies of these structures are presented in Table 5 while interaction energies and relative energies of global minima can be found in Table 6.

Binding Patterns. The water-binding motifs of monohydrated structures (cf. Figure 2) agree well with patterns founded by Shishkin et al.⁵² Our work extends an interest also to less stable tautomers (3,9), (7,9), and (3,7) and rotamer of enol-form of guanine (9,O1). Opposite to monohydrated cytosine¹³ a structure of a global minimum depends on type of guanine tautomer. For (1,7), (9,O1), (7,9), and (3,7) tautomers a water molecule forms a bridge between O6 oxygen and N7 nitrogen atoms. For (7,O2) and (1,9) tautomers a water molecule interacts with O6 and N1 nitrogen atoms, whereas (9,O2) and (3,9) tautomers prefer an interaction of a water molecule with nitrogen atoms only.

The four most stable structures of dihydrated guanine tautomers are depicted in Figure 3. The most stable structure in all tautomers [with the exception of (9,O1) and (3,9) ones]

(51) Comparing the ab initio relative free energies with the empirical potential ones (cf Table 1) we found a dramatic difference and the latter values are much too large. We were aware of this problem and we presented the gas-phase empirical potential relative free energies only to show the convergence of free energy and to estimate the errors bars. Let us further mention that large values of these relative free energies are not an artifact of the code used; the same values were obtained when using the Amber/Gibbs program. The unrealistically large gas-phase free energies are due to two main reasons: first, empirical force field is by definition unable to describe energy changes during a bond creation, annihilation or modification (Leach, A. R. *Molecular Modelling, Principles And Applications*; Pearson Education: Harlow, 1996; p 194); second, bonded hydrogen atom cannot be described by classical harmonic potential (GROMACS User Manual, version 3.1.1.). The fact that intermolecular free energy changes are often large is known. The solution of the problem mentioned is relatively easy. The values mentioned will cancel when combining different parts of thermodynamic cycle used (Harris, D.; Loew, G. J. *Comput. Chem.* **1996**, *17*, 273). Specifically, the intramolecular contribution to the free energy is the same in vacuo and bulk water and is therefore compensated while the remaining contribution (intermolecular one) forms the hydration free energy

(52) Shishkin, O. V.; Sukhanov, O. S.; Gorb, L.; Leszczynski, J.; *Phys. Chem. Chem. Phys.* **2002**, *4*, 5359.

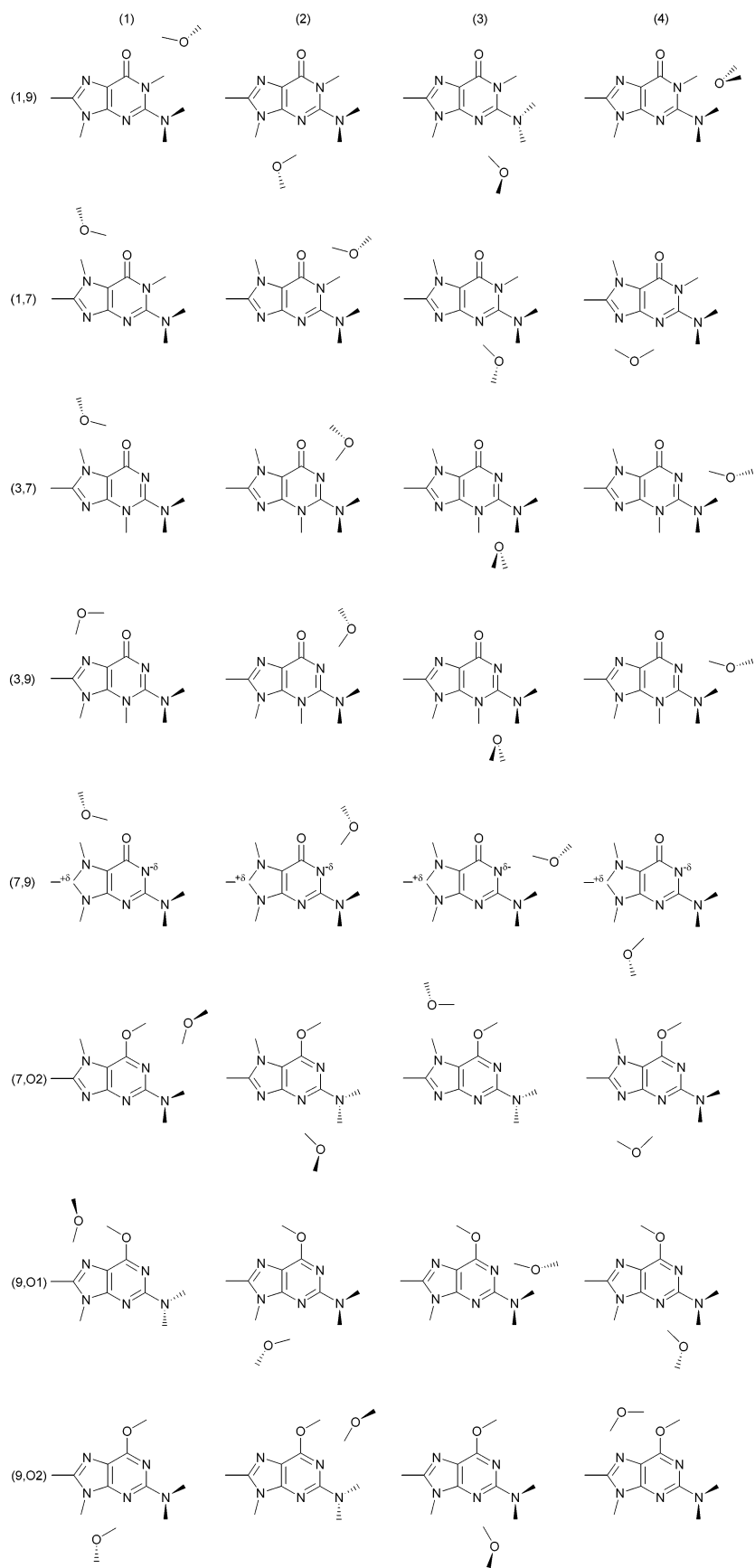


Figure 2. The four most stable structures of each guanine tautomer with one water molecule optimized at the RIMP2/TZVPP level of theory. Stability is decreasing from left to right.

contains in the first hydration shell a water dimer motif. This motif is also present in other dihydrated tautomers.

Relative Stability and Interaction Energies. The relative energies and interaction energies of structures presented are

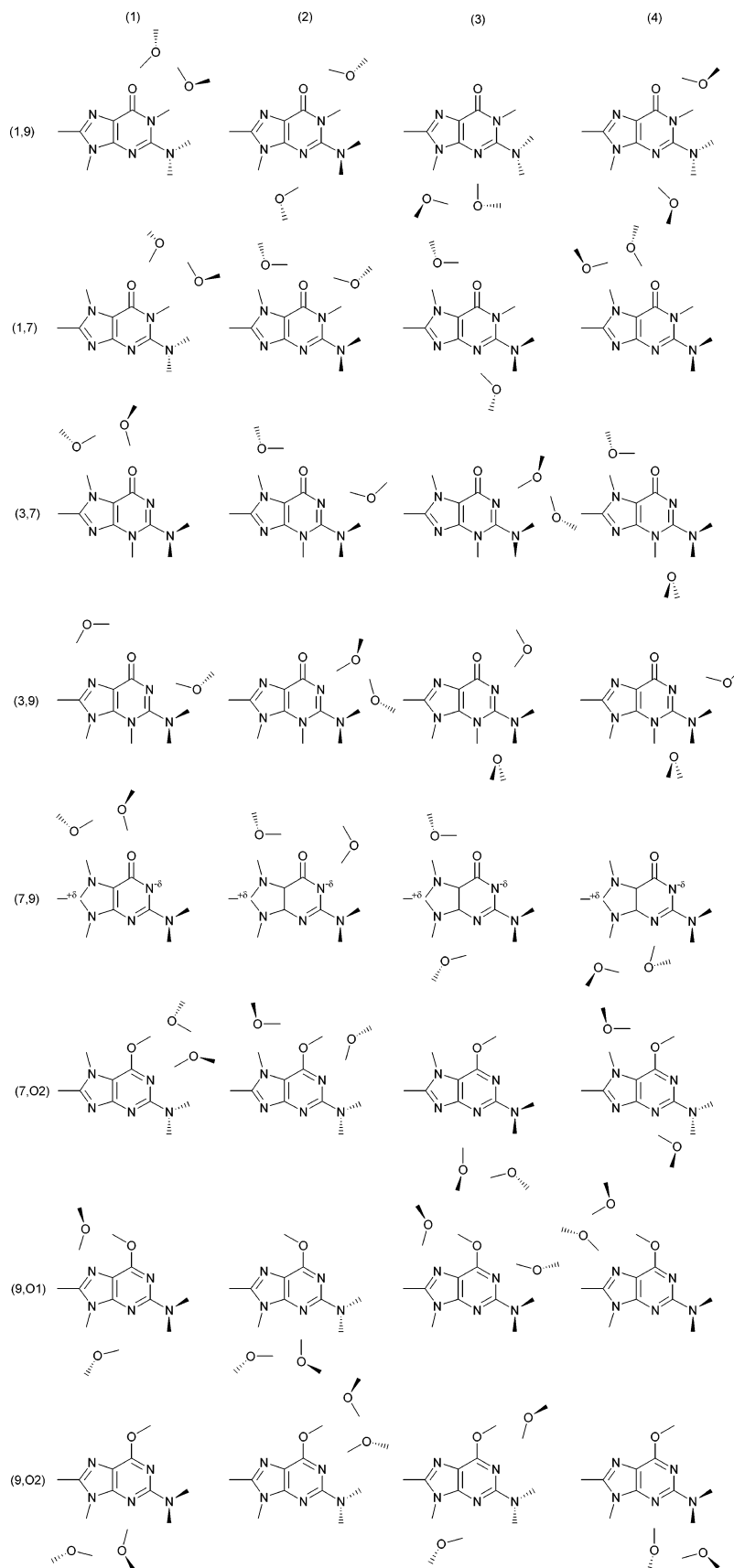


Figure 3. The four most stable structures of each guanine tautomer with two water molecule optimized at the RIMP2/TZVPP level of theory. Stability is decreasing from left to right.

summarized in Tables 5 and 6. Table 5 shows that the relative stability order of various structures is almost independent of

the computational level (empirical force field, RI-MP2). Similarly as for cytosine¹³ the differences between results obtained

Table 5. Relative Energies (in kcal/mol) for Tautomers of Guanine with One and Two Water Molecules Evaluated at Different Levels of Theory^a

| structure | MD/Q ^b | RI-MP2 ^c | structure | MD/Q ^b | RI-MP2 ^c |
|------------------------------|-------------------|---------------------|-----------|-------------------|---------------------|
| (1,7)-H₂O | | | | | |
| 1 | 0.00 | 0.00 | 1 | 0.00 | 0.00 |
| (9,02)-H₂O | | | | | |
| 1 | 0.00 | 0.00 | 1 | 0.00 | 0.00 |
| 2 | 0.35 | 0.32 | 2 | 1.02 | 1.29 |
| 3 | 2.11 | 2.29 | 3 | 2.12 | 2.27 |
| 4 | 1.51 | 2.14 | 4 | 2.73 | 3.91 |
| (9,01)-H₂O | | | | | |
| 1 | 0.00 | 0.00 | 1 | 0.00 | 0.00 |
| 2 | 1.16 | 1.99 | 2 | 0.20 | 0.92 |
| 3 | 1.26 | 3.73 | 3 | 0.64 | 1.71 |
| 4 | 3.85 | 4.56 | 4 | 2.63 | 3.55 |
| (7,02)-H₂O | | | | | |
| 1 | 0.00 | 0.00 | 1 | 0.00 | 0.00 |
| 2 | 0.10 | 1.35 | 2 | 1.20 | 1.44 |
| 3 | 0.33 | 1.91 | 3 | 0.74 | 2.50 |
| 4 | 0.72 | 3.93 | 4 | 2.23 | 3.55 |
| (1,9)-H₂O | | | | | |
| 1 | 0.00 | 0.00 | 1 | 0.00 | 0.00 |
| 2 | 3.67 | 1.98 | 2 | 4.08 | 3.09 |
| 3 | 4.34 | 3.00 | 3 | 5.31 | 3.45 |
| 4 | 1.80 | 3.42 | 4 | 4.69 | 4.34 |
| (3,7)-H₂O | | | | | |
| 1 | 0.00 | 0.00 | 1 | 0.00 | 0.00 |
| 2 | 3.18 | 2.71 | 2 | 1.98 | 0.75 |
| 3 | 4.02 | 3.23 | 3 | 4.25 | 2.38 |
| 4 | 5.13 | 5.40 | 4 | 4.20 | 3.11 |
| (3,9)-H₂O | | | | | |
| 1 | 0.00 | 0.00 | 1 | 0.00 | 0.00 |
| 2 | 0.24 | 0.76 | 2 | -1.32 | 0.11 |
| 3 | 0.96 | 2.22 | 3 | 3.02 | 1.54 |
| 4 | 2.37 | 3.29 | 4 | 2.32 | 1.66 |
| (7,9)-H₂O | | | | | |
| 1 | 0.00 | 0.00 | 1 | 0.00 | 0.00 |
| 2 | 5.34 | 4.63 | 2 | 1.56 | 0.16 |
| 3 | 5.79 | 5.86 | 3 | 0.93 | 0.98 |
| 4 | 6.30 | 7.26 | 4 | 3.97 | 1.34 |

^a Order of relative stability was referred to the most energetically preferable hydrated structure within complexes of one tautomer. ^b Modified Cornell et al. force field. ^c RI-MP2/TZVPP.

from Cornell et al. force field and RI-MP2/TZVPP calculations are very small, usually about 1 kcal/mol.

The interaction energies (Table 6) are generally in the range of 9–12 kcal/mol and 18–22 kcal/mol for mono and dihydrated tautomers, respectively and these values are similar to those obtained previously for cytosine.¹³ Larger values of stabilization energy were only obtained for (3,7) and (7,9) rare tautomers. The largest stabilization energy was determined for the latter tautomer, which is in full agreement with its large dipole moment. The global minimum of monohydrated structure is surprisingly stable and its stabilization energy is higher than 19 kcal/mol. This means that this structure is by about 6 kcal/mol more stable than the other most stable structures of remaining monohydrated tautomers. The situation is even more pronounced in the case of dihydrated (7,9) tautomer. Stabilization energy of the global minimum is more than 29 kcal/mol and also the first and the second local minima exhibit stability higher than 28 kcal/mol. The corresponding structures of other tautomers are by more than 9 kcal/mol less stable. An exception forms dihydrated (3,7) and (3,9) rare tautomers that are more stable than other (classical) tautomers. We are aware of the fact that relative stability of each tautomer should be taken into account. This disfavors mono- and dihydrated forms of the (7,9) tautomer. It is, however, possible to expect that addition of more

Table 6. Relative and Interaction Energies (in kcal/mol) of Guanine Tautomers in the Gas Phase and Mono- and Dihydrated Environment (global minima for each tautomer-water complex are presented)

| structure | relative energies ^{a,b} | | interaction energies ^{c,d} | |
|--|----------------------------------|------------------------|-------------------------------------|-----------------------|
| | RI-MP2 | RI-MP2 _{ZPVE} | RI-MP2 | RI-MP2 _{TOT} |
| (1,7) | -0.50 | -0.43 | | |
| (9,02) | 0.09 | 0.04 | | |
| (9,01) | 0.29 | 0.24 | | |
| (7,02) | 3.10 | 2.85 | | |
| (1,9) | 0.00 | 0.00 | | |
| (3,7) | 5.77 | 5.64 | | |
| (3,9) | 18.95 | 18.45 | | |
| (7,9) | 20.08 | 19.85 | | |
| (1,7)-H ₂ O | -1.76 | | -12.21 | -11.62 |
| (9,02)-H ₂ O | 1.44 | | -10.59 | -9.87 |
| (9,01)-H ₂ O | -0.66 | | -9.44 | -8.74 |
| (7,02)-H ₂ O | 3.84 | | -10.58 | -10.20 |
| (1,9)-H ₂ O | 0.00 | | -10.83 | -10.43 |
| (3,7)-H ₂ O | 2.91 | | -14.63 | -13.72 |
| (3,9)-H ₂ O | 19.31 | | -12.17 | -11.76 |
| (7,9)-H ₂ O | 15.85 | | -19.24 | -17.95 |
| (1,7)-(H ₂ O) ₂ | 0.59 | | -23.42 | -21.71 |
| (9,02)-(H ₂ O) ₂ | 2.38 | | -20.14 | -18.66 |
| (9,01)-(H ₂ O) ₂ | 1.22 | | -20.01 | -17.87 |
| (7,02)-(H ₂ O) ₂ | 5.47 | | -21.54 | -20.18 |
| (1,9)-(H ₂ O) ₂ | 0.00 | | -21.80 | -20.36 |
| (3,7)-(H ₂ O) ₂ | 3.76 | | -25.18 | -23.99 |
| (3,9)-(H ₂ O) ₂ | 18.24 | | -24.76 | -23.11 |
| (7,9)-(H ₂ O) ₂ | 13.62 | | -31.58 | -29.02 |

^a Order of relative stability of each tautomer is given with respect to canonical tautomer. Total energies of tautomers, mono- and dihydrated tautomers are considered. RI-MP2_{ZPVE} is defined as a sum of relative RI-MP2 energy and ΔZPVE; the former energy is evaluated with TZVPP basis set while the latter are at the MP2/6-31G** level. ^b For description of abbreviations used for methods cf. notes to Table 5. ^c Interaction energies were evaluated with TZVPP basis set. ^d Total complexation energy RI-MP2_{TOT} is defined as a sum of the RI-MP2 interaction energy and deformation energies of the monomers.¹¹

water molecules will finally compensate relative instability of the tautomer and the microhydrated form will become dominant (and thus spectroscopically visible) in the gas phase. The respective calculations are in progress in our laboratory.⁵³ The other rare tautomer (3,9) possessing also very large dipole moment is connected with smaller increase of stabilization energy with 1 or 2 water molecules. Interesting case represents the (3,7) tautomer, which possesses rather large stabilization energies with 1 and 2 water molecules. This tautomer is, however, only moderately less stable than the canonical form (cf. Table 3).

Hydrated Tautomers. Test of the Computational Procedure: As mentioned before, we used in the present study the MD-TI method with several newly introduced options and the method should be thus carefully tested. We systematically used the soft core potential scaling and it was necessary to test the dependence of relative free energies on α_{SC}. Table 1 shows that relative free energies of guanine tautomers are practically independent of the value of α_{SC} and in the following study we used α_{SC} equal to 1.0. This conclusion is in full accord with previously derived results.²⁷ Further, it was necessary to test the convergence of the procedure. From Table 1, it becomes clear that reasonable relative energies have been already obtained with rather short simulation time of 2 ns. Increasing the simulation length to 6 ns the convergence was practically reached and this length of simulation can be recommended for

(53) Kabeláč, M.; Hanus, M.; Hobza, P. In preparation.

Table 7. Relative Gas-Phase Free Energies (ΔG^{298}), Relative Free Energies of Hydration (ΔG^{HYD}), Evaluated with MD-TI and COSMO Methods and Relative Free Energies in Aqueous Solution [$\Delta G^{298}(\text{TI})$ and $\Delta G^{298}(\text{COSMO})$] (in kcal/mol) of Guanine Tautomers

| structure ^a | ΔG^{298} ^b | ΔG^{HYD} (TI) | ΔG^{HYD} (COSMO) | ΔG^{298} (TI) | ΔG^{298} (COSMO) |
|------------------------|-------------------------------|---------------------------------|------------------------------------|--------------------------|-----------------------------|
| (1,7) | -0.44 | -0.90 | 0.91 | -1.34 | 0.45 |
| (9,O2) | 0.12 | 5.57 | 4.96 | 5.69 | 5.08 |
| (9,O1) | 0.43 | | | | |
| (7,O2) | 2.85 | 3.10 | 3.43 | 5.95 | 6.28 |
| (1,9) | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| (3,7) | 5.48 | -18.48 | -1.75 | -13.00 | 3.73 |
| (3,9) | 17.78 | -24.84 | -8.37 | -7.06 | 9.41 |
| (7,9) | 19.81 | -30.68 | -11.30 | -10.87 | 8.51 |

^a Cf. Figure 1. ^b Cf. Table 1

further studies. The error (defined as $1/2|\Delta G_{\text{fwd}} - \Delta G_{\text{rev}}|$) becomes in most cases smaller when time was increased from 2 to 6 ns and, when longest simulations were performed, it reached negligible values. Finally, we successively reproduced the relative hydration free energies for cytosine tautomers calculated by using the MC-FEP method.¹³ Our results for (canonical form \rightarrow imino form 3a) and (canonical form \rightarrow imino form 3b) mutations fully agree with published values (4.8 ± 0.2 kcal/mol and 5.3 ± 0.1 kcal/mol, respectively). To further prove the applicability of the procedure, we tested it for a series of small systems for which experimental as well as theoretical hydration free energies are known. Specifically, we investigated the annihilation of methane (methane \rightarrow nothing) and mutation of methane to methanol, methane to methylamine, benzene to aniline and benzene to phenol. In all these cases the agreement between present theoretical results and theoretical results taken from literature⁵⁴ as well as experimental data was perfect.⁵⁴ For application details and reliability of the TI method see also ref 55.

Guanine Tautomers: The relative hydration free energies for guanine tautomers are shown in Table 7, which also gives gas-phase free energy and free energy of tautomerization in aqueous solution. In the case of five low-energy tautomers, there is qualitative agreement between the relative free energy of hydration determined using the MD-TI and COSMO⁴⁹ approaches. For the rare (3,7), (7,9), and (3,9) tautomers, the COSMO values are considerable smaller than the more reliable TI values. It must, however, be mentioned that the relative order of hydration free energy determined by both procedures is the same, and clearly all three rare tautomers are strongly stabilized in water. The MD-TI relative free energy values for (1,7) and (9,O2) tautomers (-0.9 and 5.6 kcal/mol) agree reasonably well with the previous calculations¹⁸ (1.7 and 6.4 kcal/mol, respectively) which again supports reliability of the present calculations.⁵⁶

4. Final Assessment of the Relative Stabilities of Guanine Tautomers in the Gas Phase, in a Microhydrated Environment and in Water

Theoretical relative enthalpies for isolated tautomers support the coexistence of five low-energy structures while three rare tautomers [(3,7), (7,9), and (3,9)] will be populated only negligibly and cannot be thus detected experimentally. The relative stabilities will not be changed when the temperature of the experiment increases and even at 298 K, only the five mentioned low-energy tautomers will be observed. Recent experiments from the Mons laboratory fully agree with these conclusions.²¹ The bulk solvent brings, however, dramatic changes in the relative stabilities and the three highest-energy tautomers [(3,7), (7,9), and (3,9)] become the clearly dominant ones. Following the hydration free energies obtained from the more reliable TI method, we must conclude that only these tautomers can exist in a water phase and (3,7) and (7,9) tautomers will be populated dominantly. What is the source of stabilization of these two tautomers in water phase? Favorable relative free energies of these tautomers in aqueous solution are realized in different ways. The first local minimum in aqueous solution (7,9) is highly unstable in a gas phase and favorable total relative free energy in aqueous solution is due to its very large hydration free energy. It is to be mentioned that absolute values of relative gas-phase free energy and relative hydration free energy are for this tautomer the highest among all tautomers. Let us repeat that large relative hydration free energy of (7,9) tautomer was expected on the basis of its large dipole moment (cf. Table 2). On the other hand, the global minimum (3,7) does not possess a large dipole moment and, consequently, its relative hydration free energy is moderate. By combination of moderately unfavorable ΔG^{298} and moderately favorable ΔG^{HYD} (TI) terms one obtains the most favorable ΔG^{298} (TI) term. Further, it is instructive to compare the net hydration contribution to the relative tautomer stabilities caused by microhydration. From Table 6, it is evident that (7,9) tautomer is by about 20 kcal/mol less stable than the canonical form. This very large energy difference is, however, dramatically reduced upon complexation with one (~ 13 kcal/mol) and two (~ 9 kcal/mol) water molecules. Single water or two water molecules were not able to change the tautomeric equilibrium of an isolated guanine but we expect that this can happen when a higher number of water molecules is added. The situation with (3,7) tautomer is slightly different. The isolated tautomer is by 5.8 kcal/mol less stable than the canonical form and addition of one and two water molecules reduces this energy gap (to 2.3 and 2.0 kcal/mol, respectively).⁵⁷ Evidently, the microsolvated data are between the gas-phase and bulk solvent predictions and supports the unexpected bulk solvent results.

(54) McDonald, N. A.; Heather, C. A.; Jorgensen, W. L. *J. Phys. Org. Chem.* **1997**, *10*, 563.

(55) (a) Pearlman, D. A.; Rao, B. G. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Ed.; John Wiley: Chichester, 1998; 1054. (b) Kollman, P. A. *Chem. Rev.* **1993**, *7*, 2395. (c) Miller, J. L.; Kollman, P. A. *J. Phys. Chem.* **1996**, *100*, 8587.

(56) We are aware of the fact that the direct comparison of the thermodynamic results from Monte Carlo-Free Energy Perturbation approach (used in ref 18) and our present results based on Molecular Dynamics-Thermodynamic Integration is not straightforward. Because the convergence limit was not determined in ref 18 we believe that our data are therefore more reliable. Higher number of sampling windows as well as larger box used in the present study give additional evidence about reliability of our results.

(57) Favorable relative hydration free energy of (7,9) tautomer can be explained on the basis of large stabilization energies with water environment. This is confirmed by large stabilization energies of this tautomer with one and two water molecules. Situation with (3,7) tautomer is different since this form possess only modest dipole moment (contrary to (7,9) having very large dipole moment) and, therefore, its stabilization energy with one and two water molecules is also modest. To explain large hydration free energy we evaluated the temperature dependence of hydration free energy. From the slope of this dependence we calculated the entropy term. At 298.15 K the following estimate was obtained: $\Delta H(-4$ kcal/mol); $T\Delta S(15.5$ kcal/mol). Despite the fact that the error in determining the enthalpy and entropy contributions can be significant we must admit that the leading stabilization term is entropic one. For the sake of comparison, we performed the same calculation also for (7,9) tautomer and here the entropy term was found considerably smaller: $\Delta H(-24$ kcal/mol); $T\Delta S(6.5$ kcal/mol).

Finally, it should be mentioned that surprising stabilization of (3,7) and (7,9) tautomers in a water phase is understandable in the light of unexpected stabilization of these tautomers by their interaction with aluminum atom;²² Al-(7,9) and Al-(3,7) tautomers represent the global and first local minimum.

Acknowledgment. This project, LN00A032 (Center for Complex Molecular Systems and Biomolecules), was supported by the Ministry of Education of the Czech Republic.

Note Added in Proof: In a recent paper,⁵⁸ the relative stabilities of guanine tautomers in various environments were studied. Authors utilized the DFT method for the gas-phase predictions and Poisson-Boltzmann continuum-solvation model for estimation of hydration free energies. The relative DFT gas phase free energies agree well with our results, and the largest

difference (1.2 kcal/mol) was found for 7,9 tautomer. The Poisson-Boltzmann relative free hydration energies agree reasonably well with our COSMO values and describe correctly the relative destabilization (1,7; 9,O2; 7,O2) or stabilization (3,7; 3,9; 7,9) of various tautomers. Absolute values of hydration free energies are larger than the respective COSMO data (largest difference of 5 kcal/mol was found for 3.9 tautomer), which can be explained by the fact that we did not perform the optimization of the tautomer's geometry in a continuum solvent. An important finding is that both Poisson-Boltzmann continuum-solvation model and COSMO model yielded the same order of relative hydration free energies. However, the present MD-TI results (which we believe are more reliable) yield considerably larger (more negative) hydration energies, and therefore, the global and the first local minima in the bulk water determined in ref 58 differ from the present ones.

(58) Jang, Y. H.; Goddard III, W. A.; Noyes, K. T.; Sowers, L. C.; Hwang, S.; Chung, D. S. *J. Phys. Chem. B* **2003**, *107*, 344.

JA034245Y