RESEARCH PAPER

PCCP
www.rsc.org/pccp

Correlated *ab initio* study of nucleic acid bases and their tautomers in the gas phase, in a microhydrated environment and in aqueous solution. Part 4. Uracil and thymine†‡

Jaroslav Rejnek, Michal Hanus, Martin Kabeláč, Filip Ryjáček and Pavel Hobza*

Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic and Center for Biomolecules and Complex Molecular Systems, 166 10 Prague 6 Czech Republic. E-mail: hobza@uochb.cas.cz

Received 28th January 2005, Accepted 30th March 2005
First published as an Advance Article on the web 11th April 2005

Altogether 13 *keto* and *enol* tautomers of uracil and 13 *keto* and *enol* tautomers of thymine were studied theoretically in the gas phase, in a microhydrated environment (1 and 2 water molecules) and in a water environment. Bulk water was described using the thermodynamic integration method, Conductor-like polarizable continuum model (C-PCM, COSMO) and hybrid model (C-PCM + 1–2 explicit water molecules). The structures of various tautomers were determined at the RI-MP2 level using the TZVPP basis set while relative energies were determined at the CCSD(T) level. The relative free energies at 298 K were based on the relative energies mentioned above and zero-point vibration energies, and temperature dependent enthalpy terms and entropies evaluated at the MP2/6-31G** level. The effect of bulk solvent on the relative stability of uracil and thymine tautomers was studied using molecular dynamics free energy calculations by means of the thermodynamic integration method and self-consistent reaction field. Despite the completely different nature of these methods they provide comparable solvation free energies. Besides theoretical investigation, experimental detection of uracil and thymine tautomers was performed by means of steady-state fluorescence. We conclude that it is impossible to utilize the method used by Suwaiyan and Morsy (M. A. Morsy, A. M. Al-Somali and A. Suwaiyan, *J. Phys. Chem. B*, 1999, **103**(50), 11205) for tautomer detection, even if a very sensitive fluorimeter is used. Theoretical relative energies and free energies for isolated uracil and thymine tautomers support the existence of the canonical form only. The microhydrated environment and bulk solvent stabilize *enol* forms more than the canonical *keto* one, but gas phase destabilization of these *enol* forms is too high. Population of rare *enol* forms of uracil and thymine in bulk water will thus be very low and canonical structure will also be dominant in this phase.

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 processes[1] including point mutations.[2,3] Their presence in biomolecules is, nevertheless, rather rare and NA bases are dominantly present in the most stable canonical form (up to now no direct evidence has been found of the existence of tautomers in DNA). Many studies on the tautomeric equilibria of NA bases exist which focus particularly on cytosine[4] and guanine,[5] but also on uracil and thymine.[6,7] Uracil and thymine are structural elements for a broad range of biologically relevant molecules (pyrimidine, purine, pteridine, folic acid, flavine, and antitumor agents) and a large amount of various experimental data (IR spectroscopy, Raman spectra, rare gas matrices, NMR, electron diffraction)[8–15] has been collected, but (from an experimental point of view) the picture concerning tautomeric equilibrium still remains unclear.

It is generally believed that uracil and thymine exist in the canonical diketo form in the gas phase as well as in the aqueous solution[16] but there is experimental evidence of small amounts of rare tautomeric forms in the gas phase[17,18] and in the solution.[19–22] When studying the literature carefully, one can find that the situation is more involved. Specifically, Brady[23] found that fluorescence behavior in the Fujii experiments[17] was caused by an impurity produced in the oven. Similarly, Katritzky[24] reported great uncertainty about his results "due to the fact that uracil does not follow the Hammett acidity function". Finally, Suwaiyan and Morsy[20,21] used rare tautomers relative quantum yields from the controversial experiments by Fujii[17] and Tsuchiya[18] in their calculations of the concentration of thymine. It is thus evident that the ratio of uracil rare tautomers in the aqueous solution is still unclear and requires further theoretical and experimental investigation.

In this paper we present a detailed theoretical study of the tautomerism of uracil and thymine. To support our theoretical calculations, we also present experimental data on steady state emission spectra of uracil and thymine. Gas-phase calculations of the relative stability of all 13 diketo, keto–enol, and dienol tautomers of uracil and thymine are accompanied by microhydration studies as well as studies which take the solvent effects of the bulk water fully into consideration. Experimental findings on the population of rare uracil and thymine tautomers made by Suwaiyan[20] and Morsy[21] contradict our data and literature data and this prompted us to repeat their experiments and evaluate their reliability. We attempted to reproduce Morsy's results[20,21] for thymine and to prove the validity of this technique in light of theoretical investigation of thymine and uracil. Our work improved overall insight into the tautomeric preference of uracil in the gas phase, microhydrated environment, and bulk water.

---

† For Part 3, see ref. 25.

‡ Electronic supplementary information (ESI) available: Geometries, rotational constants and dipole moments of uracyl and thymine tautomers; sample fluorescence spectra (Tables A and B and figures). See http://www.rsc.org/suppdata/cp/b5/b501499a/

DOI: 10.1039/b501499a

Our recent studies on cytosine,[4] guanine[5] and adenine[25] tautomers and other literature data[26] have shown that bulk water plays an important role and can significantly change the relative stabilities of various tautomers. The canonical form of cytosine (being in the gas phase the first local minimum destabilized over the global minimum by *ca.* 2 kcal mol^{-1}) becomes clearly favored over other structures in water environment.[4] The canonical form of guanine in the gas phase was energetically comparable with three other tautomers. Surprisingly, bulk water stabilized unusual rare tautomers (with very large dipole moments) which were extremely disfavored energetically (by *ca.* 20 kcal mol^{-1}).[5] In the case of adenine, polar solvent also reduced the gap between the canonical form (global minimum) and the first two local minima which resulted in the coexistence of these three forms which we predicted. This conclusion was fully confirmed by NMR experiments in DMSO.[25]

Similar to our previous studies, the effect of bulk water was estimated by two entirely different methods which considered the solvent as a continuum or a discrete medium. It must be mentioned that neither of the methods used are parameter free and both methods have advantages and disadvantages. The main problem is that the effect of specific hydration cannot be described properly by either approach. Molecular dynamic/empirical potential simulations (in the present versions) that are adopted to describe discrete solvents cannot account for important quantum mechanical (QM) effects like electron and charge transfers or polarization. On the other hand, quantum chemical methods used in the self-consistent reaction field calculations of isolated solute do not consider the specific electrostatic effects of most strongly-bound water molecules. Combining both approaches seems to be a natural solution. The so-called hybrid approach where solute molecules with a small number of the most strongly interacting waters is placed in the continuum solvent, attracted a lot of attention.[27–31] There are some uncertainties, the critical one concerns how to define criterion when specific hydration should be considered and how many water molecules will be taken into account. Our aim was to answer these questions and find a method which would be easy to apply to much larger systems (*e.g.*, fragments of DNA) than the present isolated bases, and which would provide reliable estimates of hydration free energies.

2. Methods

(A) Experiment

Steady state emission spectra were recorded on Jobin-Yvon Horiba spectrometer Fluorolog 3. The responses of the excitation monochromator and the emission monochromator–detector were corrected using reference spectra. Fluctuations in lamp intensity were compensated for with a reference diode detector. The excitation and emission slits were set to 10 nm. The samples were excited at 260 and 295 nm. All the chemicals used were of spectroscopic grade and water was deionized and tested to make sure they were free of any fluorescence impurities.

We used the same experimental conditions (pH, excitation wavelength, and absorption spectra range) as in Morsy and Suwaiyan's experiment[21] and a broad concentration range from 10^{-7} M up to 10^{-3} M to prove the presence of tautomeric forms.

(B) Theory

Structures. 13 structures of uracil tautomers (*cf.* Fig. 1) and 13 structures of thymine tautomers (*cf.* Fig. 2) were first presented by Kryachko[32] and van Mourik.[33] They represent all existing structures. From the figures it is evident that for all structures the classical mesomeric structure can be ascribed. Contrary to purine tautomers studied by us previously,[5,34] no "zwitterion" like structures can be derived.

Computation strategy. The gas phase geometry of uracil and thymine tautomers was determined using the resolution of identity MP2 (RI-MP2) procedure[35] with a rather large TZVPP [5s3p2d1f/3s2p1d] basis set. Relative energies were determined at the RI-MP2 level and also at the CCSD(T) level, which give higher contributions to correlation energy. Relative enthalpies and free energies were obtained by including zero-point vibration energies (ZPVE), temperature-dependent enthalpy terms, and entropies. A molecular dynamics/quenching technique with the Cornell *et al.*[36] empirical potential was utilized to explore the potential energy surface (PES) of uracil···water and thymine···water (monohydrations) and uracil···$(water)_2$ and thymine···$(water)_2$ clusters (dihydrations). The energy of the most stable structures was later recalculated at the RI-MP2/TZVPP level. The relative stability

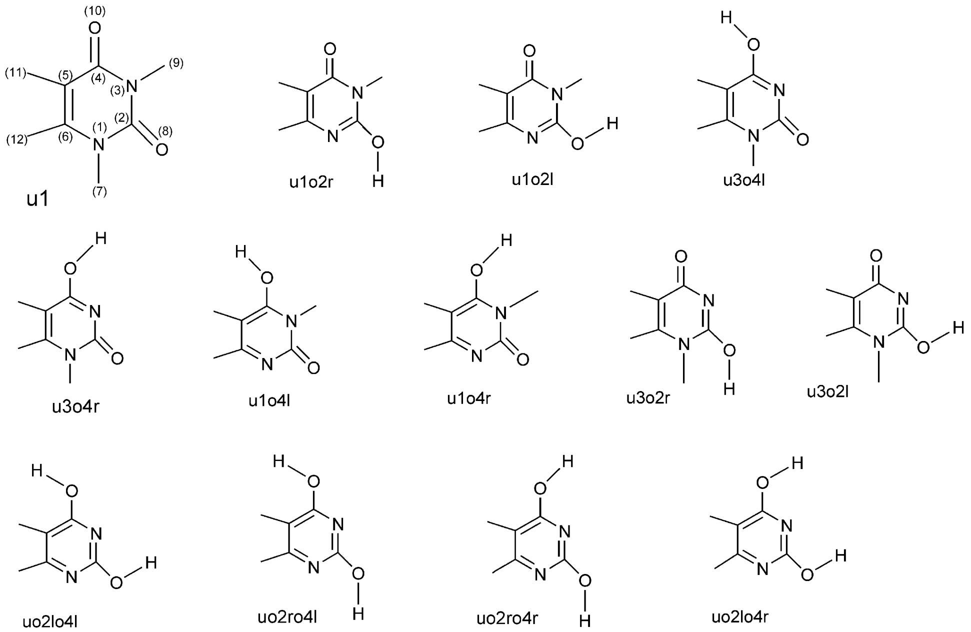

Fig. 1 The thirteen uracil tautomers. Standard numbering and adopted nomenclature are presented. In the case of bases single lines represent hydrogens, hydroxy group hydrogens of bases are depicted explicitly.

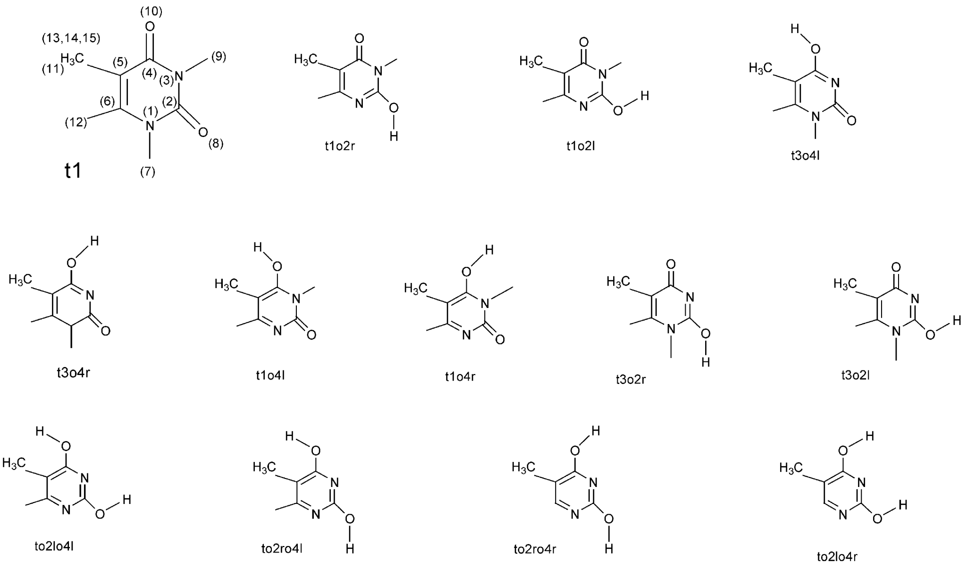

Fig. 2 The thirteen thymine tautomers. Standard numbering and adopted nomenclature are presented. In the case of bases single lines represent hydrogens, hydroxy group hydrogens of bases are depicted explicitly.

of various uracil tautomers in water environment was deduced from the free energies determined as the sum of gas-phase free energies and free energies of hydration. The latter energies were determined using the self-consistent reaction field (SCRF),[37,38] a hybrid model based on the SCRF method which considers the water molecules most strongly bound to the solute molecule, and molecular dynamics (MD) free energy calculations. In the case of *enol* tautomers, where two different structures (rotamers) with the opposite orientation of OH group exist (and one is usually energetically considerably less stable), the molecular dynamics free energy calculation was performed only for the more stable rotamer.

In the case of polar systems (like NA bases), water might introduce some changes of geometry and these changes were expected to be larger for more polar systems. Therefore, the effect of geometry optimization was studied for C-PCM (COSMO)[31,38,39] and hybrid approaches.

Quantum chemical calculations. The energetical and geometrical characteristics of 13 uracil tautomers and 13 thymine tautomers in the gas phase were investigated using the RI-MP2 procedure[35] with TZVPP basis set and default auxiliary basis sets. Correction for higher correlation terms was performed at the CCSD(T)/6-31G** level of theory. Harmonic vibration analysis providing ZPVE and thermodynamic characteristics and the character of the stationary point found was performed at the MP2/6-31G** level. Thermodynamic functions were taken from partition functions computed from MP2/6-31G** characteristics (geometry, vibrations) according to a rigid rotor-harmonic oscillator-ideal gas approximation. Relative gas phase free energies were determined as the sum of relative energy changes, zero point vibration energies, temperature dependent enthalpies, and entropies. Relative energy changes were calculated at the RI-MP2/TZVPP level with inclusion of the [CCSD(T)-MP2]/6-31G** correction term. The interaction energy of uracil$\cdots$(water)$_N$ and thymine$\cdots$(water)$_N$ complexes ($N = 1$, 2) was determined at the RI-MP2/TZVPP level with the inclusion of a basis set superposition error[40] and deformation energy. All calculations were performed with GAUSSIAN,[41,42] MOLPRO[43] and TURBOMOLE programs.[44]

Molecular dynamics/quenching (MD/Q) technique. MD/Q simulations were performed in the NVE microcanonical ensemble within quaternion formalism using the modified Cornell *et al.*[36] potential. The 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 an average temperature of 298 K. Every 1 ps the MD run was interrupted, the kinetic energy was removed, and the structure of the cluster of uracil with one or two water molecules was optimized using the conjugate gradient method. The geometry and energy in the minimum was stored, then the MD run was restarted from the point where it was interrupted.

Constants for geometrical parameters of non-canonical tautomers (not parameterized in the standard Cornell *et al.*[36] force field) were derived from quantum chemical calculations using the recommended standard procedure.[45] The atomic charges of the tautomers were generated using the electrostatic potential fitting procedure[46,47] (RESP) at the HF/6-31G* level.

Self-consistent reaction field – continuum and hybrid approaches. Bulk water was modeled with a continuum model based on the C-PCM (COSMO)[31,38,39] methodology implemented in GAUSSIAN 03.[42] The cavity was described by UAHF radii[48] (United Atoms radii optimized for HF/6-31G* level of theory). In the case of the hybrid method, the solvation-first-shell effect was approximated considering 1–2 explicit water molecules. In all cases, we adopted the following strategy concerning the choice of geometry for C-PCM calculations:

(A) *Gas-phase geometry*. Single point (SP) calculations considering gas-phase optimized geometries were carried out at HF/6-31G*/UAHF level using the GAUSSIAN 03 standard parameters (UAHF, scaling 1.2, solvent excluded surface (SES)) as recommended.[31,37] The C-PCM model was applied since it is known that it adequately describes the polar solvent.[25]

(B) *C-PCM optimized geometry*. The geometry of isolated uracil and thymine tautomers (*i.e.*, without considering any explicit water molecule) and also their clusters with one and two water molecules was optimized in the continuum solvent at the B3LYP/6-31G* level using standard GAUSSIAN 03 parameters and the recommended optimization strategy[31] (iterative method, preconditioning, tesserae 0.4 Å^2). Additional nonelectrostatic first derivatives (related to cavitation, dispersion and repulsion energies) were included in the SCRF procedure. According to ref. 31 reliable solvation free energies were obtained when the geometry optimization was followed by a single point calculation at the HF/6-31G*/UAHF level. The optimization procedure should include correlation energy, therefore we adopted the B3LYP/6-31G* level. On the other side, solvation free energies should be estimated at the HF level, because the size of the solute cavities has been optimized for HF calculations.

Table 1 Relative energies (ΔE), zero-point vibration energies (ΔZPVE), and free energies (ΔG) (all in kcal mol^{-1}) of uracil tautomers in the gas phase

Tautomer	ΔE(MP2)^a ΔE(MP2 + ZPVE)^b	ΔE(RI-MP2)^a	CCSD(T)^c	ΔE (RI-MP2 + CCSD(T))	ΔG^b/kcal mol^{-1}	μ/D
u1	0.00	0.00	0.00	0.00	0.00	5.0
u1o2l	19.88 (19.09)	17.71	−0.41	17.30	18.99	2.5
u1o2r	11.07 (10.88)	9.86	−0.32	9.54	11.06	3.7
u1o4l	22.28 (21.59)	20.56	−0.36	20.20	21.62	7.8
u1o4r	25.71 (24.66)	23.61	−0.35	23.26	24.49	6.2
u3o2l	19.78 (19.20)	17.73	−0.45	17.29	19.20	7.0
u3o2r	31.31 (29.75)	27.98	−0.72	27.26	29.71	10.0
u3o4l	20.32 (19.73)	17.46	−0.37	17.10	19.75	8.6
u3o4r	12.54 (12.33)	10.81	−0.21	10.60	12.46	5.5
uo2lo4l	17.43 (16.9)	14.54	−0.15	14.38	17.25	4.0
uo2lo4r	12.82 (12.55)	10.56	−0.07	10.49	12.85	2.5
uo2ro4l	17.44 (17.00)	14.45	−0.14	14.31	17.26	4.6
uo2ro4r	11.61 (11.42)	9.44	−0.03	9.41	11.73	1.6

[a] RI-MP2/TZVPP//RI-MP2/TZVPP. [b] MP2/6-31G**. [c] (CCSD(T)/6-31G**-MP2/6-31G**).

(C) *Hybrid model.* Four options were examined: (i) one water molecule was systematically used for all tautomeric structures with water molecules in different positions found by the MD/Q technique. The energetically most stable one (largest stabilization energy, ΔE) was used; (ii) the same as in (i) but the sum of the interaction energy (ΔE) and solvation energy (ΔG^{CPCM}) was examined; ΔE and ΔG^{CPCM} values were calculated for gas phase complexes (in light of recent work of Gaigeot and Sprik,[49,50] specific hydration seems to be more important than continuous hydration); (iii) the same as (i) but geometries were optimized in a solvent; (iv) the same as (ii) but geometries were optimized in a solvent.

Free energy calculations. The molecular dynamics–thermodynamic integration (MD-TI) method was applied to calculate the relative free energy of hydration (ΔG^{HYD}) between uracil and thymine tautomers. All calculations were performed using the GROMACS molecular modeling package[51,52] with our own code implementing the TI method. Soft core[53] potential scaling was used systematically.

The rectangular periodic box was filled with 477 TIP3P[54] water molecules and all perturbations were performed in the NPT canonical ensemble at 1.0 atm and 298.15 K. The time step was set to 1.0 fs and non-bonded cut-off of 9.0 Å was used for van der Waals and Coulombic interactions. Before perturbations, the typical 100 ps equilibration run was performed with a 1.0 fs time step, 9.0 Å cut-off, and Berendsen temperature and pressure coupling schemes.[55] Perturbations were divided into 250 (shorter perturbations) or 500 (longer perturbations) sampling windows. Vacuum calculations (representing the intramolecular terms—intrinsic terms of uracil and thymine molecule – in ΔG) were also performed according to this typical protocol: 250 windows, 1.0 fs time step, 100 ps of simulation time. The soft core parameter (α) was set to 1.00 and three simulation times (2 ns, 6 ns and 11 ns) for both forward and backward runs were used to ensure the convergence of the free energy and to estimate the error limits. All details of the simulations performed were presented in our previous studies.[4,5,25]

The relative free energy of uracil and thymine tautomers was finally determined as the sum of relative gas phase free energies and the free energy of hydration obtained from a self-consistent reaction field, hybrid method, or MD-TI treatment. As in our previous studies, the incorrect MD-TI relative gas-phase free energies were replaced by accurate *ab initio* ones[5] and the solvation free energies (ΔG^{SOL}) of specific water molecules were not taken into account.

4. Results and discussion

(A) Experiment

The steady state emission spectra of uracil and thymine were measured following the experimental conditions of Morsy[21] and Suwaiyan,[20] in which the authors claimed that fluorescence spectroscopy was a suitable tool for detecting the tautomeres of uracil and thymine. However, when measured on comparable instrumentation our results did not support the conclusion drawn by Morsy and Suwaiyan. First, uracil and thymine exhibited a very low quantum yield of fluorescence. Second, the spectra were strongly disrupted by the close position of an intense Raman band. Thus, within a concentration range of 10^{-3}–10^{-7} M this intense Raman band made it impossible to deduce valuable information on the fluorescence emission spectra of uracil and thymine. For these reasons we are not presenting any spectra inside the paper. Sample fluorescence spectra can be found in the supplementary material Part C.‡ It must be mentioned, that the spectrometer Fluorolog used in our study is characterized by a high sensitivity, which is usually expressed in the signal-to-noise ratio of the Raman band of pure water (*i.e.* 1 : 5000).

B) Theory

Gas-phase tautomers

Uracil. We examined the relative stability of one diketo form having hydrogen both at nitrogen N1 and at nitrogen N3 (canonical form), eight *enol* forms, and four dienol forms having no hydrogens on both N1 and N3 nitrogens (*cf.* Fig. 1). The geometries and rotational constants of all these tautomers are depicted in Table A in the supplementary material,‡ while Table 1 shows their relative energies, enthalpies, free energies and dipole moments. Looking at Table 1 it is evident that various tautomers have very different values of dipole moment. The smallest one belongs to uo2ro4r dienol form (1.6 D) while the u3o2r *enol* form possesses the largest dipole moment (10 D). The canonical diketo form possesses a relatively high dipole (5.0 D). Extreme values of the dipole moment indicate considerable stabilization when these tautomers are exposed to a polar solvent like water. Furthermore, their interaction with several water molecules is accompanied by significant stabilization (see below).

Looking at relative RI-MP2 energies in the second column of Table 1, we can state that all *enol* and dienol forms are energetically much less stable (by 9.4–28 kcal mol^{-1}) than the canonical tautomer. For the detailed analysis, we found it useful to create groups of tautomers according to Kryachko.[32]

The first group of low-energy tautomers (energy range from +9.4 to +10.8 kcal mol^{-1}) includes enols u1o2r and u3o4r, and dienols uo2ro4r and uo2lo4r. The second group (+14.4 and +14.5 kcal mol^{-1}) contains two dienol tautomers uo2ro4l and uo2lo4l with moderate steric repulsion. The third group (+17.5 to +17.7 kcal mol^{-1}) contains *enol* tautomers u3o4l, u1o2l and u3o2l. The last group represents high energy tautomers u1o4l,

Table 2 Relative energies (ΔE), zero-point vibration energies (ΔZPVE), and free energies (ΔG) (all in kcal mol^{-1}) of thymine tautomers in the gas phase

Tautomer	ΔE(MP2)[a] ΔE(MP2 + ZPVE)[b]	ΔE(RI-MP2)[a]	CCSD(T)[c]	ΔE (RI-MP2 + CCSD(T))	ΔG^b/kcal mol^{-1}	μ/D
t1	0	0.00	0.00	0.00	0.00	4.0
t1o2r	10.71 (10.47)	9.41	−0.13	9.28	10.65	3.2
t1o2l	19.54 (18.66)	17.23	−0.17	17.06	18.47	2.1
t1o4r	27.74 (26.44)	24.18	−0.11	24.07	26.53	6.4
t1o4l	25.18 (24.21)	21.68	−0.21	21.47	24.58	8.1
t3o2r	32.09 (30.40)	27.36	−0.24	27.11	30.79	9.6
t3o2l	19.18 (18.62)	17.25	−0.18	17.07	18.68	6.3
t3o4r	13.35 (13.14)	11.77	−0.34	11.43	13.28	5.8
t3o4l	21.76 (21.17)	19.61	−0.45	19.16	21.26	9.0
to2lo4l	18.49 (17.95)	15.49	−0.14	15.35	18.21	4.4
to2lo4r	14.34 (13.98)	10.55	0.01	10.56	13.13	2.7
to2ro4l	18.47 (17.94)	15.48	−0.14	15.35	18.22	5.0
to2ro4r	13.16 (12.74)	9.45	0.01	9.46	13.46	1.9

[a] RI-MP2/TZVPP//RI-MP2/TZVPP. [b] MP2/6-31G**. [c] (CCSD(T)/6-31G**-MP2/6-31G**).

u1o4r and u3o2r with enormous destabilization (+20.1 to +28.0 kcal mol^{-1}).

RI-MP2 results obtained with extended AO basis set provided accurate relative energy characteristics. Columns 3 and 4 of Table 1 show that the inclusion of the CCSD(T) correction term does not change the relative energy values since these corrections are rather small (the largest one amounts to 0.7 kcal mol^{-1}). Similar conclusions were also found for cytosine[4] and guanine[5] so we can state that reasonable relative energies of NA bases result from the MP2 method and higher order correlation energy contributions are not as important as claimed by Leszczynski in 1992.[56] This finding is important since the CCSD(T) calculations are CPU time demanding. Column 5 of Table 1 shows that ΔG values at $T = 298$ K are systematically larger than any energy values which means that the ΔG difference between canonical form and other tautomers is even higher.

Present gas phase results are in perfect agreement with previous theoretical and experimental results proving that only the canonical form can exist in the gas phase. U1o2r and uo2lo4r are the most stable rare tautomers, both being less stable by more than +11 kcal mol^{-1}.

Thymine. Tautomers investigated are shown in Fig. 2 and, evidently, they are very similar to those of uracil. Geometries and rotational constants are collected in Table B in the supplementary material‡ and Table 2 presents their relative energies, enthalpies, free energies and dipole moments. Again, thymine tautomers possess various dipole moments which are similar to uracil's; the largest value (9.6 D) was again found for the t3o2r form. Substational differences between dipole moments again indicate that microhydration as well as hydration will bring different effects for various tautomers. Investigating various entries in Table 2, we can only state that absolute and relative energy and free energy values for various thymine tautomers are very similar to these found for uracil tautomers (*cf.* Table 1).

Microhydrated tautomers. The four energetically most stable mono- and dihydrated structures of uracil and thymine obtained from the MD/Q simulation were optimized using *ab initio* methods; their structures are presented in Figs. 3–6. The stability of these structures decreases from left to right. The relative and interaction energies of these structures are

Table 3 Relative and interaction energies (in kcal mol^{-1}) of uracil and thymine tautomers in mono- and dihydrated environment (global minima for each tautomer–water complex are presented)

	Relative energies[ab]	Interaction energies[cd]			Relative energies[ab]	Interaction energies[cd]	
Structure	RI-MP2	RI-MP2	RI-MP2$_{TOT}$	Structure	RI-MP2	RI-MP2	RI-MP2$_{TOT}$
u1-(H_2O)	0.00	−10.47	−9.96	t1-(H_2O)	0.00	−10.41	−9.89
u1o2r-(H_2O)	8.31	−12.27	−11.18	t1o2r-(H_2O)	7.94	−12.18	−11.13
u1o4l-(H_2O)	19.70	−11.38	−10.59	t1o4l-(H_2O)	21.21	−11.57	−10.77
u3o2l-(H_2O)	14.61	−13.92	−12.70	t3o2l-(H_2O)	16.68	−13.23	−10.31
u3o4r-(H_2O)	9.32	−12.16	−11.17	t3o4r-(H_2O)	10.02	−12.36	−11.34
uo2lo4r-(H_2O)	13.30	−8.36	−7.43	to2ro4r-(H_2O)	9.29	−11.07	−10.29
uo2ro4r-(H_2O)	9.11	−10.90	−10.14	to2lo4r-(H_2O)	13.35	−8.12	−7.39
u1-(H_2O)$_2$	0.00	−21.96	−20.57	t1-(H_2O)$_2$	0.00	−20.97	−19.73
u1o2r-(H_2O)$_2$	8.52	−24.49	−21.35	t1o2r-(H_2O)$_2$	7.21	−22.36	−20.77
u1o4l-(H_2O)$_2$	19.94	−22.69	−21.01	t1o4l-(H_2O)$_2$	15.91	−30.24	−24.61
u3o2l-(H_2O)$_2$	14.93	−25.99	−23.53	t3o2l-(H_2O)$_2$	11.12	−27.33	−24.36
u3o4r-(H_2O)$_2$	8.97	−23.74	−21.95	t3o4r-(H_2O)$_2$	8.05	−24.00	−22.14
uo2lo4r-(H_2O)$_2$	16.26	−15.82	−15.29	to2ro4r-(H_2O)$_2$	8.66	−22.14	−19.89
uo2ro4r-(H_2O)$_2$	9.81	−21.39	−19.74	to2lo4r-(H_2O)$_2$	15.04	−15.70	−15.23

[a] The order of the relative stability of each tautomer is given with respect to the canonical tautomer. Relative total energies of 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 the TZVPP basis set, while the latter are at the MP2/6-31G** level. [b] For description of abbreviations used for methods, see notes to Table 1. [c] Interaction energies were evaluated with the 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.

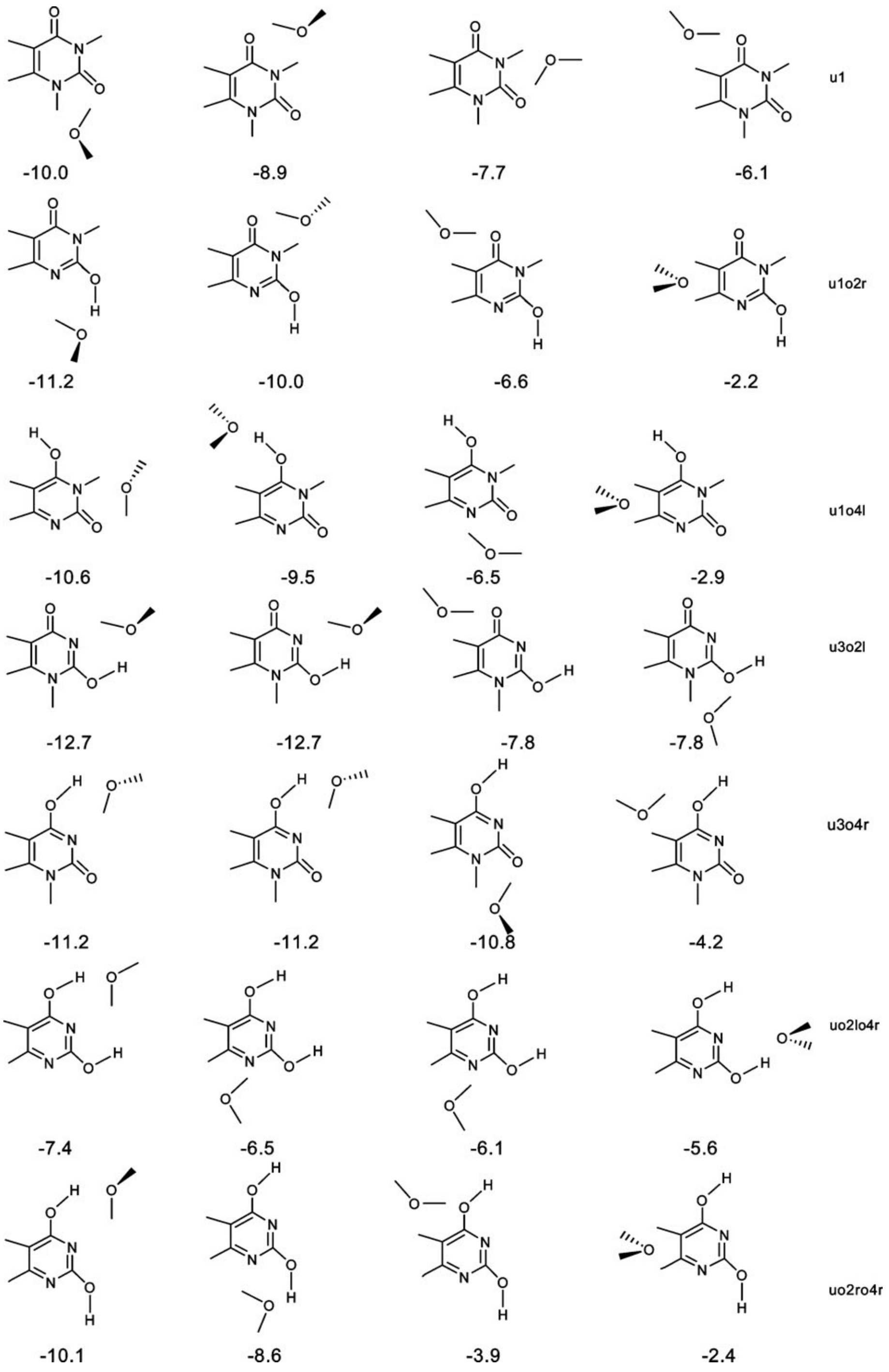

Fig. 3 Seven of the most stable structures of uracil tautomers with one water molecule optimized at the RI-MP2/TZVPP level of theory. The stability decreases from left to right. The total interaction energies in kcal mol^{-1} are presented below the structures. Structures uo2lo4r (2 and 3) and u3o2l (1,2 and 3,4) were detected at empirical level but they belong to the same minimum structure at the *ab initio* level. In the case of bases and water molecules single lines represent hydrogens, hydroxy group hydrogens of bases are depicted explicitly.

presented in Table 3 where, for the sake of comparison, the relative RI-MP2 energies are presented as well. (Only the more stable rotamer structure was studied further.)

Uracil. The water-binding motifs of seven selected monohydrated structures (*cf*. Fig. 3) reasonably agree with patterns found by other authors.[57–60] Our four detected structures correspond to four well recognized hydration sites (w1–w4 in ref. 58). However, hydration motifs related to canonical U1 form cannot be mechanistically transferred to other rare tautomers and a new hydration motif can appear. This concerns tautomers u1o2r and u1o4l where water is located between two CH groups. The stabilization energies of the relevant structures are, however, rather small (−2.2 and −2.9 kcal mol^{-1} for tautomers mentioned), but they indicate that the CH site is active for hydrogen bonding what contradicts Gaigeot and Sprik's opinion.[49,50]

Investigating the entries in Fig. 3 and Table 3, we find that the hydration energies of various tautomers are large and vary dramatically between tautomers. The smallest hydration energy was found for uo2lo4r (−7.4 kcal mol^{-1}) while the largest one (−12.7 kcal mol^{-1}) for u3o2l tautomer. These energies correlate approximately to the dipole moments of tautomers and the fact that no tight correlation between these values exists suggests that other energy terms (other than electrostatic) are also responsible for the stabilization of the hydrated tautomer. Stabilization energies of uracil with water are generally large and they are mostly larger than those of water

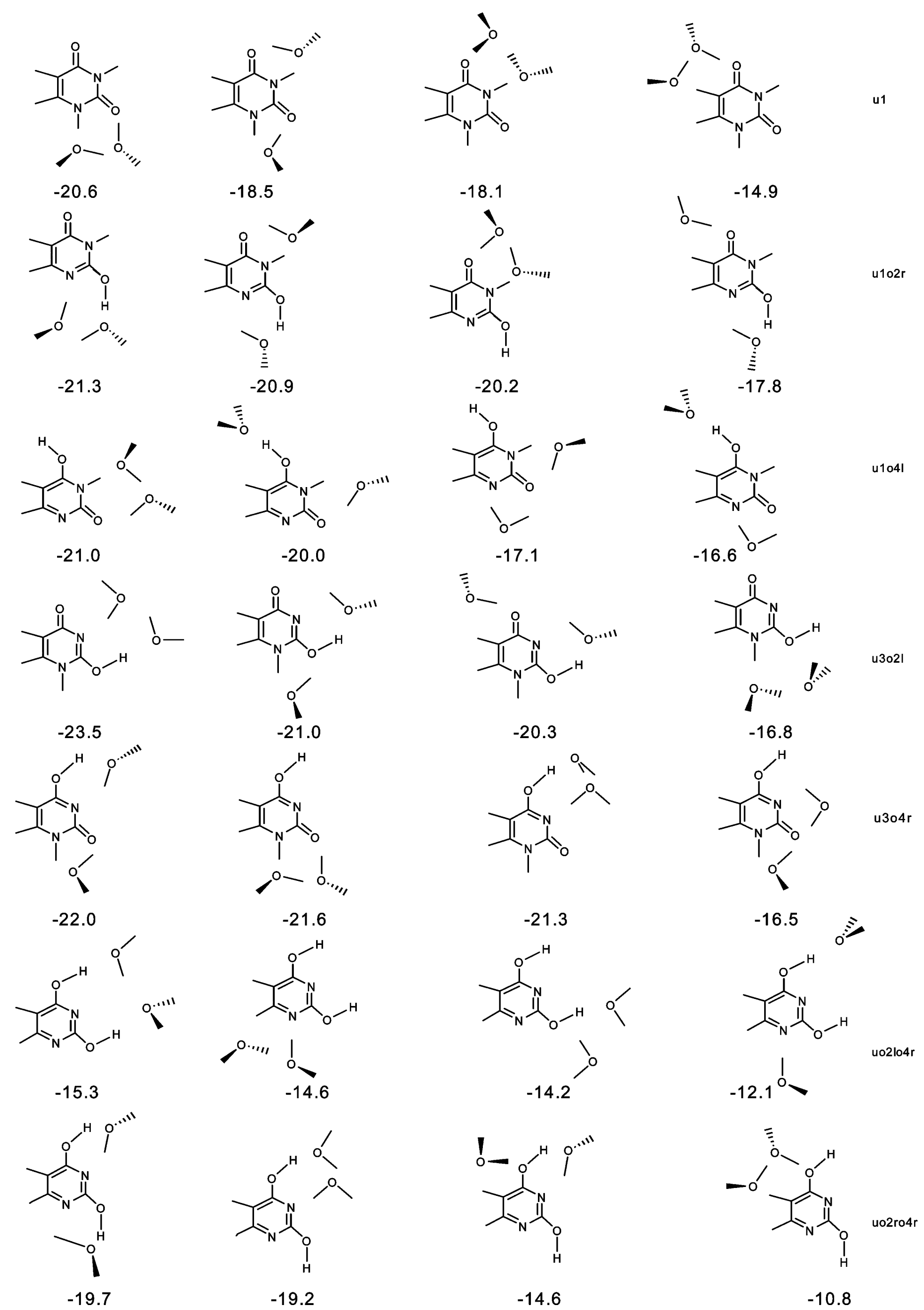

Fig. 4 Seven of the most stable structures of uracil tautomers with two water molecules optimized at the RI-MP2/TZVPP level of theory. The stability decreases from left to right. The total interaction energies in kcal mol^{-1} are presented below the structures. In the case of bases and water molecules single lines represent hydrogens, hydroxy group hydrogens of bases are depicted explicitly.

dimer (−4.5 kcal mol^{-1}; the same calculation level was used). This finding again contradicts the results of Gaigeot and Sprik[49,50] who claimed an opposite trend.

The structures of seven selected dihydrated tautomers are depicted in Fig. 4 and the respective relative and stabilization energies are shown in Table 3.

Investigating the most stable structures of dihydrated tautomers, we found that hydration sites agree with monohydrated ones. In five out of seven cases the second water hydrated the same site and simultaneously made a bridge with the first water. These structures are more stable than structures hydrated at two different sites (without a close contact between water molecules). In all cases, water dimer motif forms were at least the second most stable structures and we found them in 16 out of the 28 cases shown in Fig. 4. The unusual hydration motif found in monohydrated uracil (between two CH groups) is not found among the four most stable dihydrated structures, due to the fact that the number of uracil$\cdots$$2H_2O$ configurations, where water interacts with N and O atoms, is much larger, eliminating low interaction energy sites (*i.e.* CH sites). The CH$\cdots$$OH_2$ contact is retained in two uo2ro4r structures and one u1 and u1o2r structure. The stabilization energies of these structures are smaller than in other structures, but they are definitely evident. Comparing the fourth mono- and dihydrated structures of u1, we found that the second water, which hydrates the CH site along with the first one, causes surprisingly large stabilization of 8.8 kcal mol^{-1}. This clearly tells us that the hydration of the CH group might be substantial and should be taken into consideration.

Let us now look at the changes in the relative energies of uracil tautomers upon hydration and dihydration. The reference values of the relative RI-MP2 energies of uracil tautomers *in vacuo* are summarized in the second column of Table 1. From a comparison between Tables 1 and 3 it becomes clear that inclusion of the monohydrated structures brings some moderate changes that cannot change the order of various

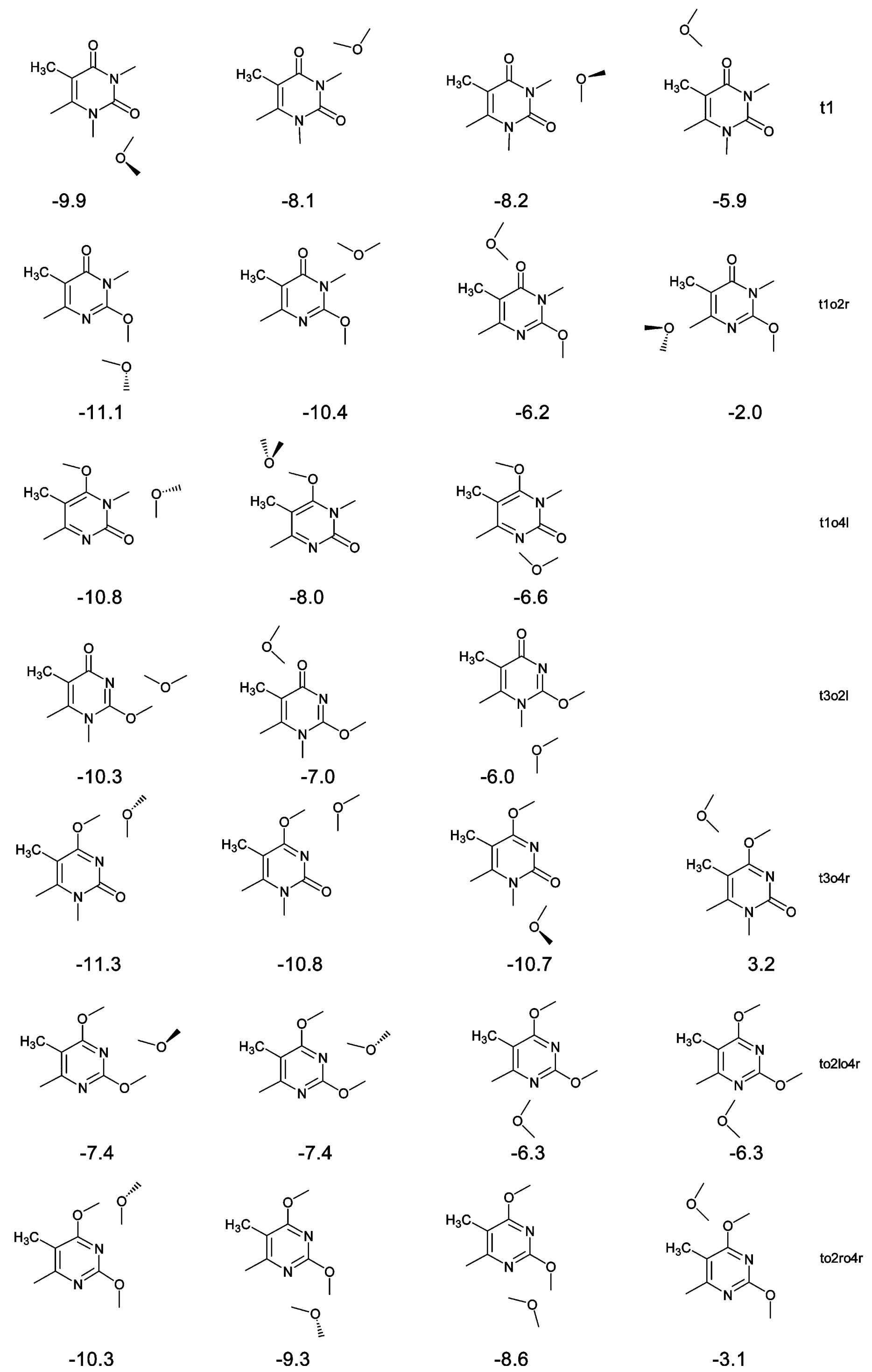

Fig. 5 Seven of the most stable structures of thymine tautomers with one water molecule optimized at the RI-MP2/TZVPP level of theory. The stability decreases from left to right. The total interaction energies in kcal mol^{-1} are presented below the structures. Structures to2lo4r (1,2 and 3,4) were detected at empirical level but they belong to the same minimum structure at the *ab initio* level. For tautomers t1o4l and t3o2l only three structures were localized. In the case of bases and water molecules single lines represent hydrogens, hydroxy group hydrogens of bases are depicted explicitly.

tautomers. The largest change concerns u3o2l, u3o4r and u1o2l where their relative energies, with respect to canonical u1 form, were reduced by 3.1, 1.5 and 1.4 kcal mol^{-1}, respectively. This finding is surprising since u3o2l, u3o4r and u1o2l are *enol* forms known to be stable *in vacuo* while *keto* forms are stabilized in a water environment. The approach of the second water does not yield any further change of relative energies, which indicates the important role of monohydration.

Thymine. The water binding motifs of the seven most stable mono- and dihydrated tautomers are displayed in Figs. 5 and 6, while their relative and interaction energies are summarized in Table 3. Evidently, both energies are similar to those of uracil tautomers and even the hydration binding energies are similar. Changes in the relative energies of various thymine tautomers induced upon monohydration and dihydration are similarly small as in the case of uracil, and water usually reduces the relative energy.

Hydrated tautomers. Gas phase relative free energies, relative hydration free energies determined by MD-TI, COSMO and hybrid approaches for uracil and thymine tautomers are shown in Tables 4 and 5.

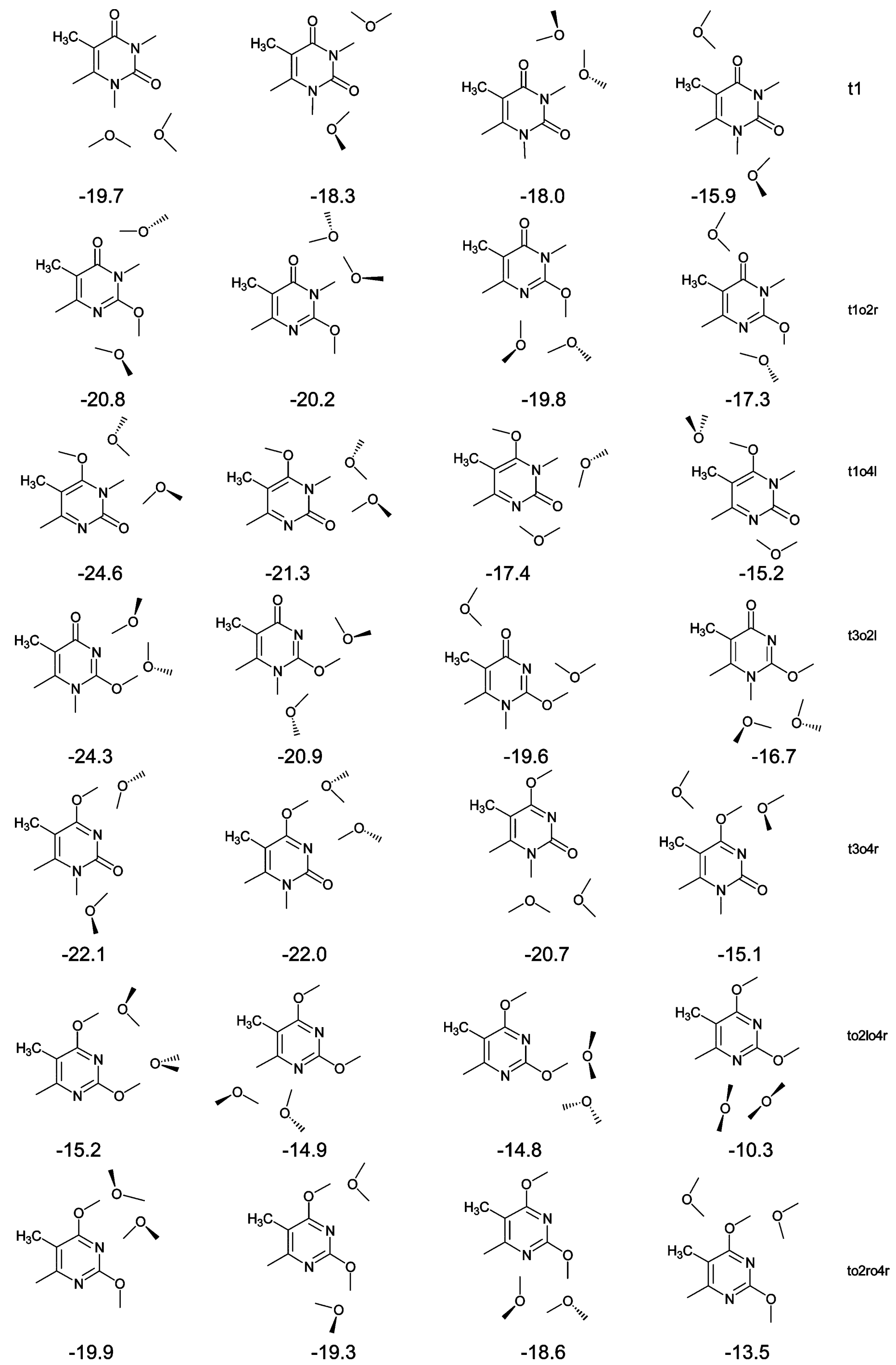

Fig. 6 Seven of the most stable structures of thymine tautomers with two water molecules optimized at the RI-MP2/TZVPP level of theory. The stability decreases from left to right. The total interaction energies in kcal mol^{-1} are presented below the structures. In the case of bases and water molecules single lines represent hydrogens, hydroxy group hydrogens of bases are depicted explicitly.

Hydration free energies. Following the expectations, MD-TI relative hydration free energies (second column of Tables 4 and 5) of uracil and thymine are similar. In both cases the *enol* x1o4l and x3o4r tautomers (x = t or u) are the best hydrated and this finding is in accord with the microhydration results. Let us add that dipole moments of these tautomers are large (though not the largest).

The COSMO results (fourth column of Tables 4 and 5) agree reasonably well with the MD-TI values for all thymine and all uracil tautomers and the relative order of the tautomers is basically retained. Comparing relative ΔG for uracil and thymine tautomers, we can state that methylation at position 5 has a minor effect and ΔG remains about the same for both tautomers. The largest absolute difference was found for t3o2l tautomer (*ca.* 2.1 kcal mol^{-1}), while the absolute average error (AAE) for uracil is slightly smaller than for thymine (0.8 and 1.3 kcal mol^{-1}, respectively). The main conclusion from the MD-TI calculations is that preferential hydration of *enol* forms is fully confirmed at the COSMO level. Taking all these results into account we must state that the bare COSMO technique gives reasonable values of hydration free energies at a much lower cost than computer-time demanding MD-TI calculations. This is, without doubt, a very important conclusion since the COSMO calculations are much simpler. In the following paragraphs we investigate various options within the COSMO technique with the aim to find whether some physically justified modifications result in closer agreement with the reference MD-TI values.

Table 4 Relative gas-phase free energies (ΔG_0^{298}), relative free energies of hydration, evaluated using MD-TI method (ΔG^{TI}), COSMO method ($\Delta G^{\mathrm{C\text{-}PCM}}$) and hybrid model ($\Delta G^{\mathrm{HYB}}$), and relative free energies in aqueous solution ΔG^{298}(TI) (in kcal mol^{-1}) of uracil tautomers

Structure[a]	ΔG_0^{298} [b]	ΔG^{TI}	ΔG^{298} (TI)	$\Delta G^{\mathrm{C\text{-}PCM}}$	$\Delta G^{\mathrm{C\text{-}PCM}}$ opt[c]	ΔG^{HYB} (1 w) vac min[d]	ΔG^{HYB} (1 w) solv min[d]	ΔG^{HYB} (1 w) opt vac[cd]	ΔG^{HYB} (1 w) opt solv[ce]	ΔG^{HYB} (2 w) vac min[d]	ΔG^{HYB} (2 w) solv min [e]	ΔG^{HYB} (2 w) opt vac[cd]	ΔG^{HYB} (2 w) opt solv [ce]
u1	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
u1o2r	11.06	0.78	11.84	0.92	0.39	−0.70	−0.01	−0.85	−0.67	2.53	2.53	−1.24	−1.24
u1o4l	21.62	−6.48	15.14	−5.92	−7.34	−6.91	−7.77	−7.50	−9.05	−7.01	−7.01	−7.49	−7.49
u3o2l	19.20	−3.22	15.98	−4.64	−5.70	−6.12	−5.43	−6.56	−6.38	−5.12	−5.12	−6.89	−6.89
u3o4r	12.46	−3.65	8.81	−2.34	−3.11	−3.18	−2.49	−3.48	−3.29	−0.31	−2.89	−0.38	−3.17
uo2lo4r	12.85	2.31	15.16	1.63	1.23	2.19	1.31	0.76	0.44	1.87	1.87	1.46	1.46
uo2ro4r	11.73	3.08	14.81	2.48	2.08	1.92	2.62	1.79	1.98	4.16	3.39	−1.02	1.65

[a] *Cf.* Fig. 1. [b] See Table 1. [c] Geometries optimized in the continuum solvent. See Methods section. [d] Global minimum in the gas phase. See Results section. [e] Global minimum in the continuum solvent. See Results section.

Table 5 Relative gas-phase free energies (ΔG_0^{298}), relative free energies of hydration, evaluated using MD-TI method (ΔG^{TI}), COSMO method ($\Delta G^{\mathrm{C\text{-}PCM}}$) and hybrid model ($\Delta G^{\mathrm{HYB}}$), and relative free energies in aqueous solution ΔG^{298}(TI) (in kcal mol^{-1}) of thymine tautomers

Structure[a]	ΔG_0^{298} [b]	ΔG^{TI}	ΔG^{298} (TI)	$\Delta G^{\mathrm{C\text{-}PCM}}$	$\Delta G^{\mathrm{C\text{-}PCM}}$ opt[c]	ΔG^{HYB} (1 w) vac min[d]	ΔG^{HYB} (1 w) solv min[d]	ΔG^{HYB} (1 w) opt vac[cd]	ΔG^{HYB} (1 w) opt solv[ce]	ΔG^{HYB} (2 w) vac min[d]	ΔG^{HYB} (2 w) solv min[e]	ΔG^{HYB} (2 w) opt vac[cd]	ΔG^{HYB} (2 w) opt solv[ce]
t1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
t1o2r	10.65	0.98	11.63	0.72	0.09	−0.73	−0.49	−0.69	−0.74	1.30	−0.07	2.44	0.64
t1o4l	24.58	−5.46	19.12	−4.62	−6.06	−5.21	−7.49	−5.36	−5.32	−6.57	−5.96	−4.06	−6.47
t3o2l	18.68	−2.50	16.18	−4.59	−5.72	−2.73	−5.26	−6.47	−6.07	−7.17	−6.50	−6.59	−7.31
t3o4r	13.28	−3.50	9.78	−3.04	−4.01	−3.72	−3.48	−4.01	−4.05	−1.93	−4.12	−1.16	−4.84
to2lo4r	13.13	3.01	16.14	0.96	0.11	1.06	1.29	0.33	0.57	1.42	2.09	1.86	1.14
to2ro4r	13.46	3.90	17.36	1.86	1.30	1.28	1.52	1.57	1.53	0.29	0.56	1.80	0.83

[a] *Cf.* Fig. 1. [b] See Table 2. [c] Geometries optimized in the continuum solvent. See Methods section. [d] Global minimum in the gas phase. See Results section. [e] Global minimum in the continuum solvent. See Results section.

First, the role of geometry optimization was studied. Values given in the fourth column of Tables 4 and 5 were obtained for gas-phase (rigid) structures. It is expected that fully-hydrated structures must relax in the electric field of the solvent. Column 5 of both tables shows the role of this geometry relaxation where, evidently, only modest differences (not exceeding 1.4 kcal mol^{-1}) were found for uracil and thymine tautomers. The optimization of tautomer structures in a solvent causes a systematic shift toward lower relative values. The absolute average error (AAE) values (relative to the MD-TI results) were enlarged for both uracil (1.1 kcal mol^{-1}) and thymine (1.8 kcal mol^{-1}) tautomers.

Let us further investigate the role of specific hydration which was not covered by the original COSMO (C-PCM) treatment. Columns 6–9 of Tables 4 and 5 collect COSMO data where one specific water is considered, and columns 10–13 present these data when two specific water molecules are taken into account (approximations i–iv presented in Computational strategy section). When gas phase geometries and ΔE are considered then results are given in columns 6 and 10, whereas when gas phase geometries and ΔG are taken into account then results are summarized in columns 7 and 11. The AAE of uracil and thymine is larger in all four cases (1.1, 1.5, 1.1, 0.9 and 1.2, 2.1, 1.7, 1.7 kcal mol^{-1}, respectively) than the one found for bare COSMO. Only when gas phase geometries and ΔE criterion were considered (column 6 and level i) did the results come close to those for bare COSMO. This means that reasonable hydration free energies are obtained providing the bare COSMO or simplest hybrid strategy (i) is used. When "more sophisticated" levels (ii), (iii) and (iv) are used, the AAE values are higher.

Summarizing the various COSMO modifications for uracil and thymine, we conclude that the best agreement with MD-TI results is gained by bare COSMO and the simplest hybrid COSMO approaches. The relative orders of the hydration free energies determined by MD-TI (u1o4l, u3o4r, u3o2l, u1, u1o2r; t1o4l, t3o4r, t3o2l, t1, t1o2r), COSMO (u1o4l, u3o2l, u3o4r, u1, u1o2r; t1o4l, t3o2l, t3o4r, t1, t1o2r) and the simplest hybrid COSMO (u1o4l, u3o2l, u3o4r, u1o2r, u1; t1o4l, t3o4r, t3o2l, t1o2r, t1) are very similar and differ only marginally. All these approaches clearly favor *enol* tautomers over canonical u1 and t1 ones in the water environment. We would like to stress here that this conclusion was obtained using different approaches (MD-TI *vs.* COSMO). The validity of the results is thus strongly supported.

Relative free energies in aqueous solution. The third columns of Tables 4 and 5 show the relative free energies of uracil and thymine tautomers in an aqueous solution. The aqueous phase stabilizes several tautomers significantly, but their gas-phase free energy destabilization is too high. For both nucleic acid bases the canonical form is clearly preferred and population of other tautomers was negligible. The only tautomer of uracil and thymine having a relative free energy in aqueous solution below 10 kcal mol^{-1} is x3o4r one.

5. Conclusions

(i) The very low quantum yield of uracil and thymine fluorescence caused the experimental conditions to be very sensitive to all kinds of "side effects". The dominant "side effect" is an intense Raman band strongly disrupting the spectrum. It is thus impossible to use the method of Suwaiyan and Morsy[20] for tautomer detection, even if a very sensitive fluorimeter is used. Their conclusion about the population of rare *enol* tautomers should thus be considered with care.

(ii) Theoretical relative energies, enthalpies, and free energies for isolated uracil and thymine tautomers support the existence of the canonical form only. Both microhydrated environment and bulk solvent stabilize *enol* forms more than the canonical *keto* one, but gas phase destabilization of these *enol* forms is too high. It is thus possible to conclude that population of rare *enol* forms in bulk water is very low and canonical structure is also clearly dominant in this phase.

(iii) The Hybrid model provides, in the case of uracil and thymine, results comparable with the standard C-PCM model only if monohydration without any optimization is used.

Acknowledgements

We would like to thank to Dr Michal Hocek for kindly providing uracil and thymine samples and to Dr Martin Hof and Mgr. Jan Sýkora for helpful discussions and measurement time on the fluorimeter. This study was supported by grants LC 512 from the MŠMT of the Czech Republic and 203/05/0009 from the Grant Agency of the Czech Republic; further it was part of the research project Z4 055 0506.

References

1 M. Yamada and I. Honma, *ChemPhysChem*, 2004, **5**, 724–728.
2 V. H. Harris, C. L. Smith, W. J. Cummins, A. L. Hamilton, H. Adams, M. Dickman, D. P. Hornby and D. M. Williams, *J. Mol. Biol.*, 2003, **326**, 1389–1401.
3 J. Florian and J. Leszczynski, *J. Am. Chem. Soc.*, 1996, **118**, 3010–3017.
4 S. A. Trygubenko, T. V. Bogdan, M. Rueda, M. Orozco, F. J. Luque, J. Šponer, P. Slavíček and P. Hobza, *Phys. Chem. Chem. Phys.*, 2002, **4**, 4192–4203.
5 M. Hanus, F. Ryjáček, M. Kabelac, T. Kubar, T. V. Bogdan, S. A. Trygubenko and P. Hobza, *J. Am. Chem. Soc.*, 2003, **125**, 7678–7688.
6 P. Hobza and J. Šponer, *Chem. Rev.*, 1999, **99**, 3247–3276.
7 J. Sponer, J. Leszczynski and P. Hobza, *J. Mol. Struct. (THEOCHEM)*, 2001, **573**, 43–53.
8 M. Ghomi, A. Aamouche, B. Cadioli, G. Berthier, L. Grajcar and M. H. Baron, *J. Mol. Struct.*, 1997, **410**, 323–326.
9 A. Les, L. Adamowicz, M. J. Nowak and L. Lapinski, *Spectrochim. Acta, Part A*, 1992, **48**, 1385–1395.
10 A. Y. Ivanov, S. A. Krasnokutski, G. Sheina and Y. P. Blagoi, *Spectrochim. Acta, Part A*, 2003, **59**, 1959–1973.
11 M. A. Palafox, N. Iza and M. Gil, *J. Mol. Struct. (THEOCHEM)*, 2002, **585**, 69–92.
12 B. Blicharska and T. Kupka, *J. Mol. Struct.*, 2002, **613**, 153–166.
13 G. Ferenczy, L. Harsanyi, B. Rozsondai and I. Hargittai, *J. Mol. Struct.*, 1986, **140**, 71–77.
14 A. Aamouche, M. Ghomi, C. Coulombeau, H. Jobic, L. Grajcar, M. H. Baron, V. Baumruk, P. Y. Turpin, C. Henriet and G. Berthier, *J. Phys. Chem.*, 1996, **100**, 5224–5234.
15 A. Aamouche, M. Ghomi, C. Coulombeau, L. Grajcar, M. H. Baron, H. Jobic and G. Berthier, *J. Phys. Chem. A*, 1997, **101**, 1808–1817.
16 J. Leszczynski, in *Encyclopedia of Computational Chemistry*, ed. P. v. R. Schleyer, John Wiley, Chicester, 1998, pp. 2951–2960.
17 M. Fujii, T. Tamura, N. Mikami and M. Ito, *Chem. Phys. Lett.*, 1986, **126**, 583–587.
18 Y. Tsuchiya, T. Tamura, M. Fujii and M. Ito, *J. Phys. Chem.*, 1988, **92**, 1760–1765.
19 A. R. Katritsky and A. J. Waring, *J. Chem. Soc.*, 1962, 1540–1544.
20 A. Suwaiyan, M. A. Morsy and K. A. Odah, *Chem. Phys. Lett.*, 1995, **237**, 349–355.
21 M. A. Morsy, A. M. Al-Somali and A. Suwaiyan, *J. Phys. Chem. B*, 1999, **103**, 11205–11210.
22 W. Hauswirth, *Photochem. Photobiol.*, 1971, **13**, 157.
23 B. B. Brady, L. A. Peteanu and D. H. Levy, *Chem. Phys. Lett.*, 1988, **147**, 538–543.
24 A. R. Katritzky, M. Karelson and P. A. Harris, *Heterocycles*, 1991, **32**, 329–369.
25 M. Hanus, M. Kabeláč, J. Rejnek, F. Ryjáček and P. Hobza, *J. Phys. Chem. B*, 2004, **108**, 2087–2097.
26 M. Orozco and F. J. Luque, *Chem. Rev.*, 2001, **101**, 203–203.
27 C. Aleman, *Chem. Phys. Lett.*, 1999, **302**, 461–470.
28 C. Aleman, *Chem. Phys.*, 2000, **253**, 13–19.
29 J. R. Pliego and J. M. Riveros, *J. Phys. Chem. A*, 2001, **105**, 7241–7247.
30 J. R. Pliego and J. M. Riveros, *J. Phys. Chem. A*, 2002, **106**, 7434–7439.
31 M. Cossi, N. Rega, G. Scalmani and V. Barone, *J. Comput. Chem.*, 2003, **24**, 669–681.
32 E. S. Kryachko, M. T. Nguyen and T. Zeegers-Huyskens, *J. Phys. Chem. A*, 2001, **105**, 1288–1295.
33 T. van Mourik, *Phys. Chem. Chem. Phys.*, 2001, **3**, 2886–2892.
34 M. Hanus, F. Ryjáček, M. Kabeláč, T. Kubař, T. V. Bogdan, S. A. Trygubenko and P. Hobza, *J. Am. Chem. Soc.*, 2003, **125**, 7678–7688.
35 M. Feyereisen, G. Fitzgerald and A. Komornicki, *Chem. Phys. Lett.*, 1993, **208**, 359–363.
36 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179–5197.
37 V. Barone, M. Cossi and J. Tomasi, *J. Chem. Phys.*, 1997, **107**, 3210–3221.
38 V. Barone, M. Cossi and J. Tomasi, *J. Comput. Chem.*, 1998, **19**, 404–417.
39 A. Klamt, G. J. P. Krooshof and R. Taylor, *AIChE J.*, 2002, **48**, 2332–2349.
40 S. F. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553.
41 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle and J. A. Pople, *GAUSSIAN 98, (Revision A.7)*, Gaussian, Inc., Pittsburgh PA, 1998.
42 G. W. T. M. J. Frisch, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *GAUSSIAN 03, (Revision A.1)*, 2003, Gaussian Inc., Pittsburgh PA.
43 R. D. Amos, A. Bernhardsson, A. Berning, P. Celani, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, P. J. Knowles, T. Korona, R. Lindh, A. W. Lloyd, S. J. McNicholas, F. R. Manby, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, G. Rauhut, M. Schütz, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson and H.-J. Werner, *MOLPRO, a package of ab initio programs designed by H.-J. Werner and P. J. Knowles, Version 2002.1*, 2002.
44 R. Ahlrichs, M. Bar, M. Haser, H. Horn and C. Kolmel, *Chem. Phys. Lett.*, 1989, **162**, 165–169.
45 T. A. D. D. A. Case, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, B. Wang, D. A. Pearlman, M. Crowley, S. Brozell, V. Tsui, H. Gohlke, J. Mongan, V. Hornak, G. Cui, P. Beroza, C. Schafmeister, J. W. Caldwell, W. S. Ross and P. A. Kollman,Amber 8: Manual and help page, 2004, http://amber.scripps.edu/.
46 C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269–10280.
47 W. D. Cornell, P. Cieplak, C. I. Bayly and P. A. Kollman, *J. Am. Chem. Soc.*, 1993, **115**, 9620–9631.
48 *GAUSSIAN 03* on-line help.
49 M. P. Gaigeot and M. Sprik, *J. Phys. Chem. B*, 2004, **108**, 7458–7467.
50 M. P. Gaigeot and M. Sprik, *J. Phys. Chem. B*, 2003, **107**, 10344–10358.
51 E. Lindahl, B. Hess and D. van der Spoel, *J. Mol. Model.*, 2001, **7**, 306–317.
52 H. J. C. Berendsen, D. Vanderspoel and R. Vandrunen, *Comput. Phys. Commun.*, 1995, **91**, 43–56.

53 T. C. Beutler, A. E. Mark, R. C. Vanschaik, P. R. Gerber and W. F. Vangunsteren, *Chem. Phys. Lett.*, 1994, **222**, 529–539.

54 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

55 H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.

56 J. Leszczynski, *J. Phys. Chem.*, 1992, **96**, 1649.

57 T. van Mourik, D. M. Benoit, S. L. Price and D. C. Clary, *Phys. Chem. Chem. Phys.*, 2000, **2**, 1281–1290.

58 E. S. Kryachko, M. T. Nguyen and T. Zeegers-Huyskens, *J. Phys. Chem. A*, 2001, **105**, 1934–1943.

59 O. Dolgounitcheva, V. G. Zakrzewski and J. V. Ortiz, *J. Am. Chem. Soc.*, 2000, **122**, 12304–12309.

60 A. K. Chandra, M. T. Nguyen and T. Zeegers-Huyskens, *J. Phys. Chem. A*, 1998, **102**, 6010–6016.