At Nonzero Temperatures, Stacked Structures of Methylated Nucleic Acid Base Pairs and Microhydrated Nonmethylated Nucleic Acid Base Pairs are Favored over Planar Hydrogen-Bonded Structures: A Molecular Dynamics Simulations Study

Martin Kabeláč and Pavel Hobza*[a]

Abstract: The dynamic structure of all ten possible nucleic acid (NA) base pairs and methylated NA base pairs hydrated by a small number of water molecules (from 1 to 16) was determined by using molecular dynamics simulations in the NVE microcanonical and NVT canonical ensembles with the Cornell force field (W. D. Cornell, P. Cieplak, C. I. Bayly, I. R.Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. E. Caldwell, P. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179). The presence of one water molecule does not affect the structure of any hydrogen-bonded (H-bonded) nonmethylated base pair. An equal population of H-bonded and stacked structures of adenine ••• adenine, adenine ••• guanine and adenine ••• thymine pairs is reached if as few as two water molecules are present, while obtaining equal populations of these structures in the case of adenine ••• cytosine, cytosine ••• thymine, guanine ••• guanine and guanine ••• thymine required the presence of four water molecules, and in the case of guanine ••• cytosine, six. A comparable population of planar, H-bonded and stacked structures for cytosine ••• cytosine and thymine ••• thymine base pairs was only obtained if at least eight water molecules hydrated a pair. Methylation of bases changed the situation dramatically and stacked structures were favoured over H-bonded ones even in the absence of water molecules in most cases. Only in the case of methyl cytosine ••• methyl cytosine, methyl guanine ••• methyl guanine and methyl guanine ••• methyl cytosine pairs were two, two or six water molecules, respectively, needed in order to obtain a comparable population of planar, H-bonded and stacked structures. We believe that these results give clear evidence that the preferred stacked structure of NA base pairs in the microhydrated environment, and also apparently in a regular solvent, is due to the hydrophilic interaction of a small number of water molecules. In the case of methylated bases, it is also due to the fact that the hydrogen atoms most suitable for the formation of H-bonds have been replaced by a methyl group. A preferred stacked structure is, thus, not due to a hydrophobic interaction between a large bulk of water molecules and the base pair, as believed.

Keywords: AMBER empirical potential • hydrogen bonds • molecular dynamics • nucleobases • solvent effects • stacking interactions

Introduction

The interaction of nucleic acid (NA) bases is of key importance since it determines the structure, and therefore also the functions, of DNA and RNA. Two types of base-pair interactions, planar and vertical, that lead to the formation of H-bonded and stacked structures are relevant. Each of the two interactions is of a different nature. The former are dominated by Coulombic (electrostatic) energy contributions originating in the highly dipolar character of NA bases, while the latter originate mainly in London dispersion energy due to the high polarisabilities of NA bases. H-bonding interactions are highly directional and specific, while stacking interactions are less specific. Both types of NA interactions are, however, important and both contribute to the unique structure and function of DNA and RNA.

In the gas phase (at nonzero temperature), H-bonded pairs are energetically more stable than stacked pairs. Methylation of the bases can, however, change the relationship between these two structural types—due to the increased polarisability of the bases as well as the more favourable role of entropy. Methylation also results in the most favourable H-bonded structures vanishing. Furthermore, the medium plays a crucial role. Bases associate mainly by H-bonding in the gas-phase[1] and in nonpolar solution,[2–4] while in water the stacked structures are clearly dominant.[5–7] Base pairing in the gas phase is now well understood and many theoretical papers

---

[a] Prof. P. Hobza, Dr. M. Kabeláč
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)
E-mail: hobza@indy.jh-inst.cas.cz

have been published on this subject.[8] The different pairing of NA bases in the gas and liquid phases represents a challenging problem and several attempts have been made to explain it.[9–11] All these studies, based on calculated ΔG values, correctly predicted the tendency of stacked structures to be preferred over H-bonded structures in water. Conversion from H-bonded to stacked structure occurred, however, with a rather large amount of solvent ($\sim$870,[9a] $\sim$750,[10] or $\sim$500[11] water molecules). The important question thus arises: whether the transition from a planar, H-bonded structure to a stacked structure is controlled by the hydrophobic interaction of a large bulk of water molecules or is due to the hydrophilic action of a small water cluster, or even due to the nature of the stacking interactions themselves. The problem is now topical since the gas phase hydration of biomolecules (which allows the addition of a specific number of water molecules) is being intensely studied in various laboratories. We have found only two examples in the literature that show the important role of a limited number of water molecules on the structure of base pairs and concern the interaction energy. Using a simple empirical potential, Desfrancois, Carles and Schermann have shown that, in the presence of a single water molecule, the lowest energy structure of the 9-methyl adenine ••• 1-methyl thymine pair becomes stacked.[12a] Sivanesan et al.[12b] studied the hydration of H-bonded and stacked structures of a cytosine dimer. Using empirical electrostatic potential, Hartree–Fock optimisation and MP2 single-point calculations, the authors have shown that stacked dimers in the presence of two or three water molecules were slightly (by about 5%) more stable than H-bonded pairs. In our previous communication[13] we showed that a gradual increase in the hydration number of the adenine ••• thymine NA base pair results in a transition from planar base-pair structures to nonplanar ones and that, for as few as two water molecules, the population of stacked structures is higher than that of planar structures. Since we studied the interaction of free bases, both orientations of bases in the stacked arrangement (face to face or face to back) are possible. Furthermore, only the most stable tautomers (oxo and ammine forms) of bases are considered, since the populations of other tautomers are negligible.

Editorial Board Member:[*] *Pavel Hobza was born in 1946 in Czechoslovakia and graduated from the Czech Technical University, Prague in 1969. In 1974 he received his PhD degree (with Rudolf Zahradník) and in 1988 his DSc degree. After postdoctoral study with Camille Sandorfy at the Université de Montréal, Montréal, he spent several periods as a visiting professor and as a visiting scientist at the Université de Montréal, the Friedrich-Alexander-Universität Erlangen-Nürnberg and the Technische Universität München. In 1989 he become an associated Professor, then in 2001 a full Professor of Charles University in Prague. Since 1989 he has worked at the J. Heyrovský Institute of Physical Chemistry at the Czech Academy of Sciences, Prague and since 2000 has been Head of the Department of Complex Molecular Systems at this institute and Head of the Center of Complex Molecular Systems. He has also been a fellow of the Learned Society of the Czech Republic since 1998. Prof. Hobza has been the author or co-author of more than 200 papers and two books. These studies dealt with molecular interactions and their role in chemistry and biodisciplines.*

The aim of the present paper is to find the dynamic structure of all ten possible NA base pairs and methylated base pairs (N9 methylation of purines and N1 methylation of pyrimidines) when they are hydrated by a small number of water molecules (from one to 16). For the numbering of the atoms see Figure 1.

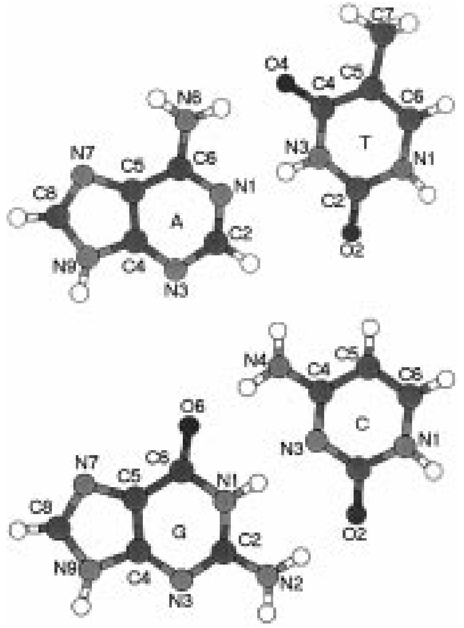

Figure 1. Standard atom numbering for adenine (A), cytosine (C), guanine (G) and thymine (T).

Computational Methods

The molecular dynamics (MD) calculations were performed by using the Cornell force field,[14] since this force field has been shown[8] to give the best energy characteristics for the base–base interaction in comparison with other empirical potentials used for DNA simulations. This is for planar, H-bonded as well as stacked interactions. The atomic charges of all bases and methylated bases (N9 for adenine (A) and guanine (G), N1 for cytosine (C) and thymine (T)) were determined by using a restrained electrostatic potential fit[15] at the HF/6-31G* level. Atomic charges determined at this level are overestimated; the resulting overestimation of electrostatic energy compensates for the missing polarisation energy. The electrostatic dipole–dipole interactions play an important role in stabilising H-bonded and also stacked pairs.

---

[*] Members of the Editorial Board will be introduced to readers with their first manuscript.

0947-6539/01/0710-2068 $ 17.50+.50/0

(In the latter case they determine the structure of the pair, while the stabilisation comes from dispersion energy.) It must be mentioned here that the electrostatic part of the Cornell potential only includes atomic charges and dielectric constants, that is, no parameters are included and care was paid to determine the atomic charges as accurately as possible. From this point of view, the use of the Cornell empirical potential, originally designed for native DNA, is also justifiable for interactions of free bases. This is documented by the close agreement between nonempirical MP2 ab initio and empirical results obtained for free H-bonded and stacked base pairs. In the case of the former complexes, the largest difference in stabilisation energies does not exceed 2.5 kcal mol^{-1}, while for latter this difference is larger, but still below 3.5 kcal mol^{-1}.

Each base pair was solvated by one, two, four, eight and 16 TIP3P[16] water molecules without periodic boundary conditions. No cut-off distance was used for these simulations. To avoid separation of the bases by a distance larger than 16.0 Å, energy barrier restraints were applied. The same restraints were applied to the distances of water molecules from bases. No restraints on the internal geometry of the bases or of the base pair were used. Simulations were carried out in the NVE microcanonical and NVT canonical ensembles with a 0.5 fs integration step (total simulation time was 10 ns). The geometry was recorded every 1 ps. The temperature in the NVT simulations was maintained by using the Berendsen algorithm.[17] NVT simulations were performed at 300 K and, for selected base pairs, also at 400 K. NVE simulations of hydrated and nonhydrated base pairs were performed at various constant total average energies corresponding to an average temperature of 300 K. In the case of the guanine ··· cytosine and 9-methyl guanine ··· 1-methyl cytosine hydrated and nonhydrated base pairs, these simulations were also performed at higher total average energies corresponding to an average temperature of 400 K.

Simulations in the NVE ensemble give the properties of a cluster that does not interact with its surroundings, while the simulations in the NVT ensemble correspond to the situation in which the cluster is in thermal equilibrium with its surroundings. The simulations in different ensembles should provide different results and it is only in the thermodynamics limit that different ensembles yield the same results.

Results and Discussion

To distinguish between the planar, H-bonded and stacked structures of the base pair, we used the distance between the centres of mass of individual bases. Taking the canonical B-DNA structure[18] we found that these distances for planar adenine ··· thymine and guanine ··· cytosine Watson–Crick (WC) base pairs were 6.02 and 5.59 Å, respectively, while for stacked guanine ··· adenine and cytosine ··· thymine pairs they were only 3.89 and 3.79 Å, respectively. We would add that these canonical distances were very close to the distances obtained from optimised, isolated base pairs.[8] Intermolecular distance can thus be used unambiguously to distinguish between H-bonded and stacked structures. Here we present histograms showing the tendency for the bases to adopt various intermolecular distances. Figure 2 shows the results for nonmethylated and methylated base pairs hydrated by 0 to 16 water molecules.

Besides intermolecular distances, we also considered the angle between the two base planes. We distinguished five structural types: i) planar, H-bonded (deviation of base planes is smaller than 25°); ii) nonplanar, H-bonded (deviation of base planes lies in the interval 25–45°); iii) T-shaped (deviation of base planes lies in the interval 45–90°); iv) planar stacked (deviation of parallel planes is smaller than 25°); v) nonplanar stacked (deviation of parallel planes lies in the interval 25–45°). The intermolecular centre-of-mass distance for structural types i)–iii) was between 4.9 and 8.0 Å; in the case of both stacked structures it was below 4.9 Å. Finally, subsystems were considered separated if this distance was larger than 8.0 Å. The population of all base pairs in various structural types is shown in Table 1.

NVE microcanonical ensemble

Nonmethylated base pairs: The distance fluctuation trajectories (not shown) and histograms (Figure 2) for nonhydrated and monohydrated pairs are very similar, and show an intermolecular centre-of-mass distance between 5.25 and 6.5 Å. This means that the H-bonded structures dominate. The shortest distance (5.25 Å) was found for the CC pair, a distance of 5.5 Å was detected for AC, CT and CG pairs, a distance of 6.0 Å was identified for AG and AT pairs and the largest distances of 6.25 and 6.5 Å were characteristic for AA, GG, GT and TT pairs. The difference between nonhydrated and monohydrated pairs is rather small for most base pairs. Analysing the structural types of nonhydrated pairs, we found roughly comparable populations of planar, H-bonded, nonplanar, H-bonded and T-shaped structures. Analysing the H-bonded patterns with two and three H-bonds of various base pairs, we found that adenine utilised mostly the N3 atom and the H atom attached to the N9 atom, while in the case of guanine the O6 atom and the H atom attached to the N1 atom were involved. Cytosine forms H bonds with other bases mainly through the N3 atom and H atoms of the amino group; thymine through the O2 atom and the N1 hydrogen atom. Similar H-bonded patterns were found for most of the nonhydrated as well as for all hydrated structures. Monohydration usually decreased the population of H-bonded pairs and increased the population of T-shaped structures. From an analysis of the H-bonded pattern for purine bases, we have found that, for adenine, the N3 atom and the hydrogen atom attached to the N9 atom are mostly utilised for H-bond formation. For guanine, in most cases the formation of H-bonds concerns the O6 atom and the hydrogen atom attached to the N1 atom . In the case of the pyrimidine bases, cytosine utilises mostly the N3 atom and the hydrogen atoms of an amino group and thymine utilises the O2 atom and the hydrogen atom attached to the N1 atom.

The addition of a second water molecule leads to important changes, and the trajectories and histograms for AA, AG, AT and GT pairs clearly show the existence of different structural types with an intermolecular distance below 4 Å, which correspond to stacking. Table 1 clearly shows that the

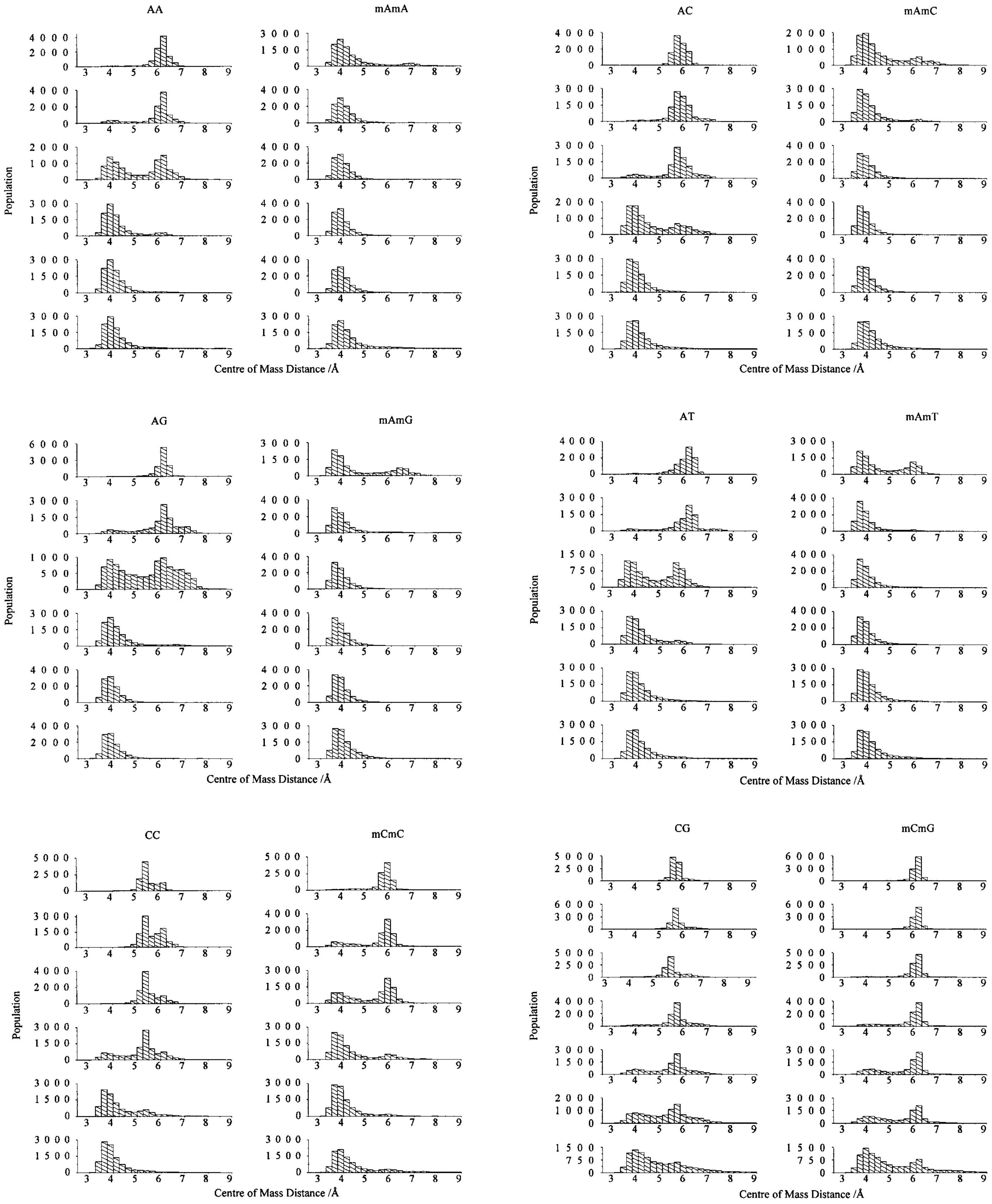

Figure 2. Histograms showing the tendency for various intermolecular base–base centre-of-mass distances [Å] of a base $\cdots$ base $\cdot$ $(\text{water})_n$ cluster in which $n = 0$ (top), 1, 2, 4, 8 and 16 (bottom) for all ten possible NA base pairs and methylated NA base pairs. Histograms were calculated for the whole simulation time (10 ns). For the nonmethylated and methylated pairs of CG, histograms for a water cluster of 6 molecules (third from bottom) are also included.

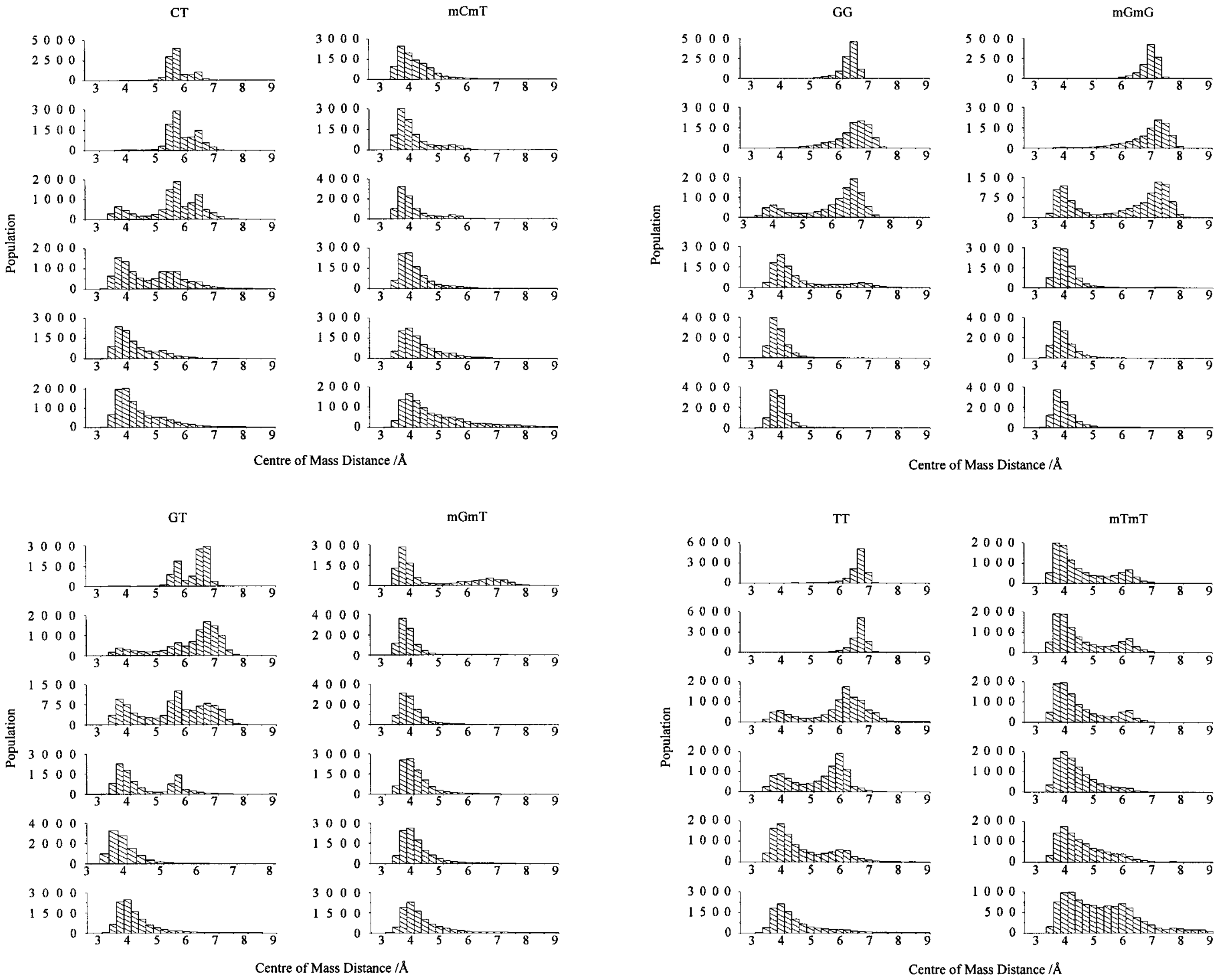

Figure 2. (Continued)

population of H-bonded, stacked and T-shaped structures is now comparable for these pairs. For the remaining pairs, the population of stacked structures increases but is still lower than that of T-shaped and H-bonded structures. The presence of a larger number of water molecules is required in order to obtain a comparable population of H-bonded and stacked structures for AC, CT and GG pairs (four water molecules), the CG pair (six water molecules) and the CC and TT pairs (eight water molecules). The addition of eight water molecules leads to a dominant population (>88%) of stacked structures for all base pairs with the exception of CC, CG, CT and TT. Moreover, in all these cases it is a planar stacked structure that is by far the dominant structure. Passing from eight to 16 water molecules did not bring any important changes, and the trajectories and histograms for all base pairs hydrated by 16 and eight water molecules are similar. In the case of CC and TT pairs, the population of stacked (mainly planar stacked) structures increased. For the CT pair, the population of stacked structures slightly decreased, while the population of T-shaped structures slightly increased.

How can we explain the fact that a higher population of stacked structures of the pyrimidine base pairs (CC, CT and TT) is achieved at the highest hydration number (eight water molecules)? This is certainly not due to the magnitude of the stabilisation energy; the stabilisation energy differences between H-bonded and stacked structures of these pairs are 10.5, 2.9 and 4.1 kcal mol^{-1}, respectively, while those of CG WC and AT WC pairs are 16.5 and 4.2 kcal mol^{-1}, respectively.[8] This fact can only be explained on the basis of an unfavourable geometrical arrangement of proton-acceptor sites in the stacked structures of the pyrimidine base pairs. While these positions in other stacked base pairs are arranged in such a way that a few (two or four) water molecules can easily make an H-bonded bridge between them (cf. Figure 3, below), in the case of the aforementioned base pairs, the formation of such bridges is only possible if more water molecules (six or eight) cooperate.

Increasing the total average energy for the CG pair yields a higher frequency of interconversions, and a comparable population of stacked and H-bonded structures is achieved

Table 1. Populations [%] of various structural types of all ten possible nonmethylated and methylated NA base pairs. HB p = H-bonded planar, HB np = H-bonded nonplanar, T = T-shaped, Stack np = Stacked nonplanar, Stack p = Stacked planar and Sep. = Separated subsystems. For a more detailed description see the main text.

Water No.	HB p	HB np	T	Stack np	Stack p	Sep.
			AA			
0	29	33	35	1	2	0
1	27	28	32	3	10	0
2	10	15	28	7	39	1
4	3	4	6	7	80	0
8	2	2	4	7	85	0
16	2	2	5	7	84	0
			mAmA			
0	2	3	11	8	74	0
1	1	2	3	6	88	0
2	1	2	2	4	91	0
4	1	1	1	4	93	0
8	1	2	2	5	90	0
16	2	3	7	5	80	2
			AC			
0	46	35	19	0	0	0
1	35	32	29	1	3	0
2	28	28	31	3	9	1
4	6	9	22	10	52	1
8	2	1	5	10	82	0
16	2	2	7	11	77	1
			mAmC			
0	6	6	20	9	59	0
1	2	2	9	8	79	0
2	1	1	4	8	86	0
4	1	1	2	5	91	0
8	1	1	3	6	89	0
16	2	3	6	8	80	1
			AG			
0	37	37	26	0	0	0
1	20	24	44	3	9	0
2	8	13	45	9	25	1
4	2	3	8	9	77	1
8	1	1	1	5	92	0
16	1	1	2	5	91	0
			mAmG			
0	3	7	24	5	61	0
1	1	2	7	6	84	0
2	1	1	2	6	88	0
4	1	1	1	4	92	0
8	1	1	1	4	93	0
16	2	2	3	5	88	0
			AT			
0	36	34	28	1	1	0
1	24	28	35	4	9	0
2	16	18	31	11	24	0
4	5	6	17	19	53	0
8	1	2	8	21	68	0
16	2	2	10	20	65	1
			mAmT			
0	11	12	21	7	49	0
1	2	3	10	8	77	0
2	1	1	6	7	85	0
4	1	1	4	6	88	0
8	2	2	5	7	84	0
16	2	3	7	6	81	1
			CC			
0	48	37	15	0	0	0
1	43	35	22	0	0	0
2	39	33	26	1	1	0
4	23	23	31	6	17	0
8	4	6	16	13	61	0
16	1	2	9	12	76	0
			mCmC			
0	46	33	17	1	3	0
1	34	27	22	3	14	0
2	23	20	22	7	28	0
4	5	6	11	11	66	1
8	2	3	6	9	80	0
16	4	5	13	9	66	3
			CG			
0	49	36	15	0	0	0
1	43	34	22	0	1	0
2	36	33	28	1	2	0
4	30	29	32	3	5	1
6	19	20	35	8	17	1
8	12	15	30	10	32	1
16	5	7	27	12	43	4
			mCmG			
0	43	37	20	0	0	0
1	38	36	25	0	1	0
2	31	34	28	2	3	0
4	25	28	32	5	9	1
6	18	22	31	9	18	2
8	15	19	29	13	22	2
16	9	12	20	17	38	4
			CT			
0	49	35	16	0	0	0
1	39	36	23	1	1	0
2	27	26	29	3	14	1
4	11	14	26	9	39	1
8	4	5	16	14	61	0
16	3	5	18	14	59	1
			mCmT			
0	1	1	31	15	52	0
1	2	4	14	10	70	0
2	2	4	10	8	76	0
4	2	2	8	11	77	0
8	2	4	13	11	70	0
16	4	6	22	10	54	4
			GG			
0	43	32	24	0	1	0
1	26	27	45	1	1	0
2	19	22	39	5	15	0
4	4	6	14	9	66	1
8	1	0	1	3	95	0
16	1	0	1	3	95	0
			mGmG			
0	44	33	23	0	0	0
1	19	24	44	1	1	10
2	11	15	31	4	32	9
4	1	1	2	4	89	1
8	1	1	0	3	95	0
16	1	1	1	2	95	0
			GT			
0	56	30	13	0	1	0
1	24	26	38	4	8	0
2	19	18	34	6	23	0
4	15	10	10	7	57	1
8	1	1	4	9	85	0
16	2	3	7	10	78	0
			mGmT			
0	3	6	22	4	63	0
1	1	1	5	5	88	0
2	1	1	4	7	87	0
4	2	2	5	8	83	0
8	2	2	5	7	82	0
16	2	4	9	7	75	3

Table 1. (cont.)

Water No.	HB p	HB np	T	Stack np	Stack p	Sep.
			TT			
0	41	36	23	0	0	0
1	28	33	34	2	2	1
2	20	24	35	7	12	2
4	14	19	37	11	19	0
8	6	9	21	22	41	1
16	3	5	15	24	52	1
			mTmT			
0	6	8	20	8	58	0
1	6	8	18	10	58	0
2	6	7	15	11	61	0
4	3	5	13	13	66	0
8	5	7	18	10	59	1
16	8	11	28	9	38	6

at a higher hydration number (eight rather than six; not shown). Full hydration by 16 water molecules gives the same dominant structure at both average energies—a stacked structure; the population of other structures (H-bonded and T-shaped) is, however, significant at the higher energy. When the total energies for MD simulations are lowered, the sampling of the potential energy surface is limited and the results are strongly dependent on the starting structure.

The change from a planar, H-bonded to a stacked structure is due to the action of a small number of water molecules. Figure 3 shows some typical structural arrangements for two and four water molecules hydrating the AT base pair; this gave the most significant structural variations (see above). In the first arrangement (S1. $2H_2O$), water molecules act separately on the opposite sides of the bases and each water molecule forms one hydrogen bond with one base and a second with the other. The second arrangement (S2. $2H_2O$) is characteristic of a H-bonded water pair and each water molecule interacts with only one nitrogenous base. The H-bonds in the planar base pair are sometimes not disrupted and both water molecules are located in the first hydration shell (HB1. $2H_2O$). Otherwise, one H-bond in the pair is retained while the other is disrupted and a water molecule forms a bridge between adenine and thymine (HB2. $2H_2O$). Similar configurations were also observed in T-shaped structures. Here either one H-bond is retained and the other is disrupted and replaced by a chain of two water molecules (T1. $2H_2O$) or both H-bonds are disrupted and water molecules form a bridge between the bases (T2. $2H_2O$).

The total empirical stabilisation energies for optimised structures S1. $2H_2O$ and HB1. $2H_2O$ are very similar (35.7 and 35.3 $\text{kcal}\,\text{mol}^{-1}$), while the base–base stabilisation energies differ. In the case of the S1. $2H_2O$ complex, the base–base stabilisation energy amounts to 11.9 $\text{kcal}\,\text{mol}^{-1}$ and the sum of all the stabilisation energies between both bases and both waters (and also water–water contributions) equals 23.8 $\text{kcal}\,\text{mol}^{-1}$. This means that the stabilisation from base–water complexes is twice as large as stabilisation from base–base interactions. In the case of HB1. $2H_2O$, the base–base and base–water stabilisation energies amount to 15.5 and 19.8 $\text{kcal}\,\text{mol}^{-1}$, respectively. Evidently, in the case of H-bonded complexes, the geometry arrangement of both water molecules is not as favourable as in the case of stacked pairs.

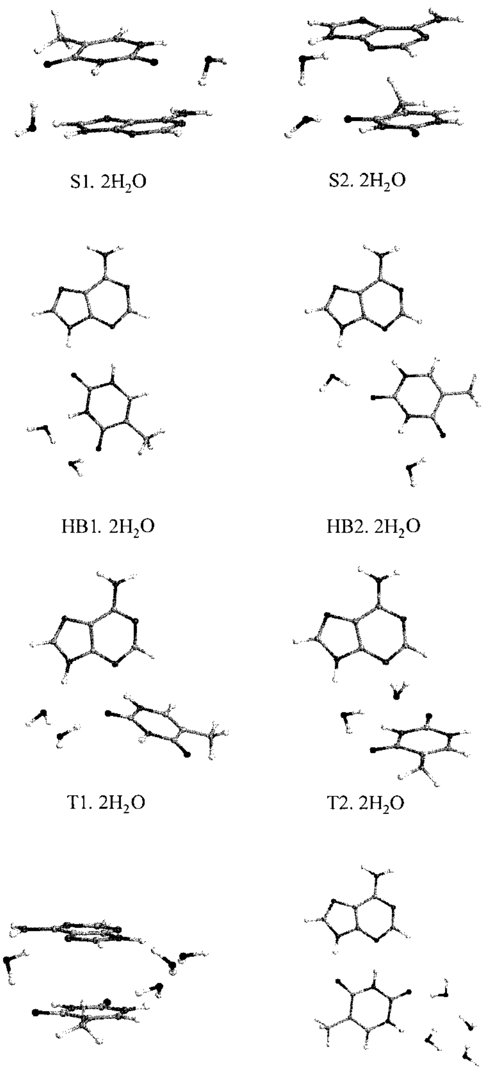

Figure 3. Snapshots from the NVE molecular dynamics simulations of the AT.$n\,H_2O$ cluster. Selected stacked (S), H-bonded (HB) and T-shaped structures (T) are shown. For a full description see the main text.

The basic structural motifs of stacked arrangements with four water molecules are similar to the previous ones (see stacked arrangement S1. $4H_2O$). The H-bonded structures are only rarely present, mostly as structures having a cluster of four water molecules bonded together (HB1. $4H_2O$). The overall picture is not changed for hydration by eight or 16 water molecules (Figure 2b). Similar structural motifs can also be found for the remaining base pairs.

Performing the NVE simulations at higher total average energies (the corresponding average temperature was about 400 K) leads to a slight increase in the population of stacked structures, and, further, the tendency for dissociation is increased. On the other hand, at lower total average energy, the sampling of the potential surface is limited and results depend strongly on the starting structure.

We conclude that the higher relative population of stacked structures, and particularly of planar stacked structures, reflects a gradual increase in hydration number (see Table 1).

Methylated base pairs: The distance fluctuation trajectories (not shown) and histograms (Figure 2) both show a different picture in comparison with nonmethylated pairs, and a higher population of stacked structures is obtained for most methylated base pairs (mAmA, mAmC, mAmG, mAmT, mCmT, mGmT and mTmT) even without the presence of a water molecule. Comparing the populations of various structures of nonmethylated and methylated pairs (Table 1), we find profound differences. While for the former pairs the population of H-bonded (planar and nonplanar) and T-shaped structures is comparable and that of stacked structures is negligible, in the case of the latter pairs the population of planar stacked pairs is clearly dominant. This is because the most favourable H-bonded patterns presented in the nonmethylated structures vanish as a consequence of the substitution of the hydrogen atom attached to the N9 or N1 atom by a methyl group. The only exception is for mCmC, mGmG and mCmG pairs, in which a comparable population of stacked and H-bonded structures was reached by hydration with two, two or six water molecules respectively, similar to the situation for nonmethylated pairs. This gives clear evidence that the most favourable H-bond patterns of the nonmethylated pairs and of the methylated base pairs are identical.

NVT canonical ensemble: The trajectories and histograms for AT.$n\,H_2O$ and CG.$n\,H_2O$ ($n = 0, 1, 2, 4, 8, 16$) pairs generated at lower average energy in the NVE microcanonical-ensemble and in the NVT canonical-ensemble (300 K) simulations exhibit great similarities (not shown). This also applies to CG trajectories and histograms obtained at higher average energy (NVE microcanonical ensemble) and at 400 K (NVT canonical ensemble). Almost identical results from both simulations were also reached for mAmT and mCmG base pairs (not shown).

Conclusion

The effect of hydration on the structure of nonmethylated and methylated NA base pairs was studied by using MD NVE microcanonical-ensemble and NVT canonical-ensemble simulations. In the case of the former pairs, a gradual increase in hydration number results in a transition from planar, H-bonded structures to stacked structures. For some base pairs, the population of stacked structures is higher than that of H-bonded structures for as few as two water molecules. For the remaining base pairs, comparable populations of stacked and H-bonded base pairs were reached with hydration by four, six or eight water molecules. In the case of methylated bases, stacked structures are dominant for the majority of base pairs, even without any water. We believe that these results give clear evidence that the preferred population of stacked structures of NA base pairs in an aqueous environment is due to the hydrophilic interaction of a small number of water molecules with the base pair. In the case of the methylated bases, it is also due to the fact that the most favourable H-bonded structures vanish with the substitution of hydrogen atom by a methyl group. The preferred population of the stacked structures is thus not due to a hydrophobic interaction between a large bulk of water molecules and a base pair, as believed until now.

Hydration by a higher number of water molecules (32) led to similar results and the only difference was the fact that the probability of dissociation of the bases increased. This means that bases are separated by several water molecules. The same tendency was also observed during simulations in a periodic box of 256 water molecules.

Acknowledgements

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

---

[1] I. K. Yanson, A. B. Teplitsky, L. F. Sukhodub, *Biopolymers* **1979**, *18*, 1149.
[2] J. H. Miller, H. M. Sobell, *J. Mol. Biol.* **1967**, *24*, 345.
[3] J. S. Binford, Jr., D. M. Holloway, *J. Mol. Biol.* **1968**, *31*, 91.
[4] L. Katz, S. Penman, *J. Mol. Biol.* **1966**, *15*, 220.
[5] Y. Kyogoku, R. C. Lord, A. Rich, *Science* **1966**, *154*, 518.
[6] T. N. Solie, J. A. Schellman, *J. Mol. Biol.* **1968**, *33*, 61.
[7] T. Imoto, *Biochim. Biophys. Acta* **1977**, *475*, 409.
[8] P. Hobza, J. Šponer, *Chem. Rev.* **1999**, *99*, 3247, and references therein.
[9] a) P. Cieplak, P. Kollman, *J. Am. Chem. Soc.* **1988**, *110*, 3734; b) A. V. Teplukhin, V. I. Poltev, N. V. Shulyupina, G. G. Malenkov, *J. Biomol. Struc. Dyn.* **1989**, *7*, 75; c) V. I. Poltev, A. V. Teplukhin, J. S. Kwiatkowski, *J. Biomol. Struc. Dyn.* **1992**, *9*, 747.
[10] J. Florián, J. Šponer, A. Warshel, *J. Phys. Chem. B* **1999**, *103*, 884.
[11] E. Stofer, C. Chipot, R. Lavery, *J. Am. Chem. Soc.* **1999**, *121*, 9503.
[12] a) C. Desfrancois, S. Carles, J. P. Schermann, *Chem. Rev.* **2000**, *100*, 3943; b) D. Sivanesan, K. Babu, S. H. Gadre, Y. Subramanian, T. Ramasami, *J. Phys. Chem. A*, **2000**, 104, 10887.
[13] M. Kabeláč, F. Ryjáček, P. Hobza, *Phys. Chem. Chem. Phys.* **2000**, *2*, 4906.
[14] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. E. Caldwell, P. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179.
[15] C. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269.
[16] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
[17] H. J. C. Berendsen, J. P. Postina, W. F. van Gunsteren, A. Di Nola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684.
[18] S. Arnott, Z. Chandrasekaran, D. L. Birdsall, A. G. W. Leslie, R. L. Ratliff, *Nature* **1980**, *283*, 743.

Received: December 18, 2000 [F2947]