ELSEVIER

Chemical Physics 256 (2000) 275–287

Chemical Physics

www.elsevier.nl/locate/chemphys

Density functional theory studies of methylated uracil: geometries and energies

C.F. Zhang, X.J. Chen *, Z.S. Yuan, Z.J. Zhang, K.Z. Xu

Laboratory of Atomic and Molecular Physics, Department of Modern Physics, University of Science and Technology of China, P.O. Box 4, Hefei, Anhui 230027, People's Republic of China

Received 21 February 2000

Abstract

Density functional theory studies on the geometries and energies of the methylated derivatives of uracil yield two stable conformations, α and β, for each single-methylated uracil. They are different in the spatial orientation of the substituting methyl group and the molecular total energy. Analyzing the calculated structural parameters, we also found an elongation effect in the methylated uracil, which contributes to the increase of dipole moment and molecular size of molecules such as the methylated derivatives of nucleic acid bases. © 2000 Elsevier Science B.V. All rights reserved.

1. Introduction

Since the double helix structure model of DNA molecule was proposed by Watson and Crick in 1953 [1,2], nucleic acid bases (NAB), being the elementary building blocks of DNA and RNA, have always been the subject of numerous theoretical and experimental research studies. The most recent studies were focused on topics such as the following: the excited state geometry optimization calculations [3], the vibrational spectra experimental [4–7] and theoretical [8] studies, the experimental and theoretical determination of valence electron affinities [9] and dipole-bound electron affinities [9,10], the structure and energy calculations of NAB-amino acid complexes [11], and also the tautomer studies of NAB [12]. Among all these, the influence of methylation on the structures and properties of NAB is an intriguing area. In RNA (DNA), it is on the N1 position that uracil (thymine) is connected with the ribose-phosphate backbone. As a reasonable approximation, the methylated derivative is a simple but satisfactory simulation to the actual RNA (DNA) system. Thus, the studies of these compounds are significant in understanding their properties and activities in the biological systems involving macromolecules [13]. Previous studies have shown that the pattern of the infrared absorption spectra of uracil in the carbonyl region is very sensitive to substitution on the two ring nitrogen atoms (N_1 and N_3) [14]. Recently, Desfrancois et al. [10] investigated the influence of *N*-methylation on the dipole-bound electron affinities of thymine and uracil theoretically using ab initio quantum mechanical calculations, and experimentally using Rydberg electron transfer spectroscopy. Both their theoretical and experimental studies are consistent

---

[*] Corresponding author. Fax: +86-551-3601164.
E-mail address: xjun@ustc.edu.cn (X.J. Chen).

PII: S0301-0104(00)00117-8

in showing that replacement of hydrogen atoms by methyl group reduces electron affinities corresponding to the formation of dipole-bound anions of these systems. The reduction is attributed to an increase of the molecular size due to methylation. The methylation also reduces the distortion of the anion geometries with respect to the geometries of the neutral parents. In a theoretical study on the NAB and their methylated derivatives [13], Stewart et al. carried out high-level ab initio optimizations on the NAB thymine, cytosine, guanine, adenine and their methyl derivatives at both the HF and MP2 levels of theory, applying the standard 6-31G** basis set. The results show a large deviation between the HF (noncorrelated method) and MP2 (correlated method) geometries, implying that full optimization is necessary with the inclusion of electron correlation. In a density functional theory (DFT) study on the molecular and electrostatic properties of the *N*-methylated NAB [15], Bakalarski et al. presented complete geometry optimizations for the dominant tautomeric forms of NAB (thymine, cytosine, guanine, and adenine) and their derivatives methylated at the N1 (pyrimidines) or N9 (purines) positions. However, to the best of our knowledge, there have not been complete studies of the methylated derivatives of uracil at all the possible substitution positions up to now, especially substitution positions at atoms N3, C5, and C6.

In this paper, we focused on the geometrical parameter calculations and total energies analysis of the methylated derivatives of uracil. Applying the DFT method with B3LYP and BLYP exchange-correlation (XC) functionals via the 6-31G, 6-31+G, 6-31G*, and 6-31+G* basis sets, respectively, we studied all the four single-methylated uracil molecules, 1-methyluracil (1MU), 3-methyluracil (3MU), 5-methyluracil (5MU or thymine) and 6-methyluracil (6MU). Also presented here were the two important double-methylated derivatives, 1,5-dimethyluracil (15-DMU) and 1,6-dimethyluracil (16DMU). For every molecule studied, we gave the detailed geometrical parameters. For uracil, 5MU (thymine), 15DMU, the previously available X-ray diffraction and neutron diffraction experimental data were presented for a comparison and a reasonable agreement was obtained.

2. Methods of calculations

All the DFT calculations in this work were performed using the GAUSSIAN 98W program package [16]; various XC functionals and basis sets were applied in the calculations, which will be described in the following. A nonlocal density functional used here was BLYP which is the combination of the Becke's gradient correction to the LSD exchange functional [17] and the most popular one of the dynamical correlation functionals modified by Lee, Yang and Parr (LYP) [18]. Another XC functional used here was the hybrid functional B3LYP which is a three-parameter expression of E_{xc} determined by a least squares fitting to the G1 molecule set [19]. In the previous study of the tautomers of uracil [12], we have found that for the molecules like NAB, the hybrid functional B3LYP will give the most satisfactory results in calculating the geometrical parameters. So, it was also applied in this work. Five basis sets, 6-31G, 6-31+G, 6-31G*, 6-31+G*, and 6-31++G**, were applied, they are different in that whether or not the polarization functions (*) and the diffuse functions (+) were included in the calculations. The geometrical parameters of the single-methylated uracil (1MU, 3MU, 5MU and 6MU) were fully optimized without planarity constraints at the B3LYP/6-31G, B3LYP/6-31+G* and BLYP/6-31++G* levels, respectively. The geometrical parameters of the double-methylated uracil (15DMU and 16DMU) were also optimized in the same condition at the B3LYP/6-31++G** level. In order to clarify the influence of inclusion of polarization functions and diffuse functions on the geometry optimizations of NAB, we took 5MU (thymine) as an example, calculated its geometrical parameters at four different levels (B3LYP/6-31+G*, B3LYP/6-31+G, B3LYP/6-31G*, and B3LYP/6-31G) and the results were compared. The notes to all the bases used here have been presented in Ref. [20]. The total energies of all the four single-methylated uracil molecules were calculated at three different levels (B3LYP/6-31G, B3LYP/6-31+G*, and BLYP/6-31+G*) and their relative stability was ordered and analyzed.

3. Results and discussion

The elaborate geometrical structures of the seven molecules studied here are shown in Fig. 1. For a convenient comparison, we also included here the results of the dominant tautomeric form of uracil, which had been studied earlier [12].

3.1. Uracil

Being the simplest one of the five naturally predominant NAB (uracil, thymine, cytosine, guanine, and adenine), the structures and properties of uracil have been studied extensively [21–37]. Presented here is the full optimization of its (the predominant 2,4-dioxo-dilactam tautomeric form) geometrical parameters (Table 1) and the comparison with the previously available experimental results [21–25] (Table 3).

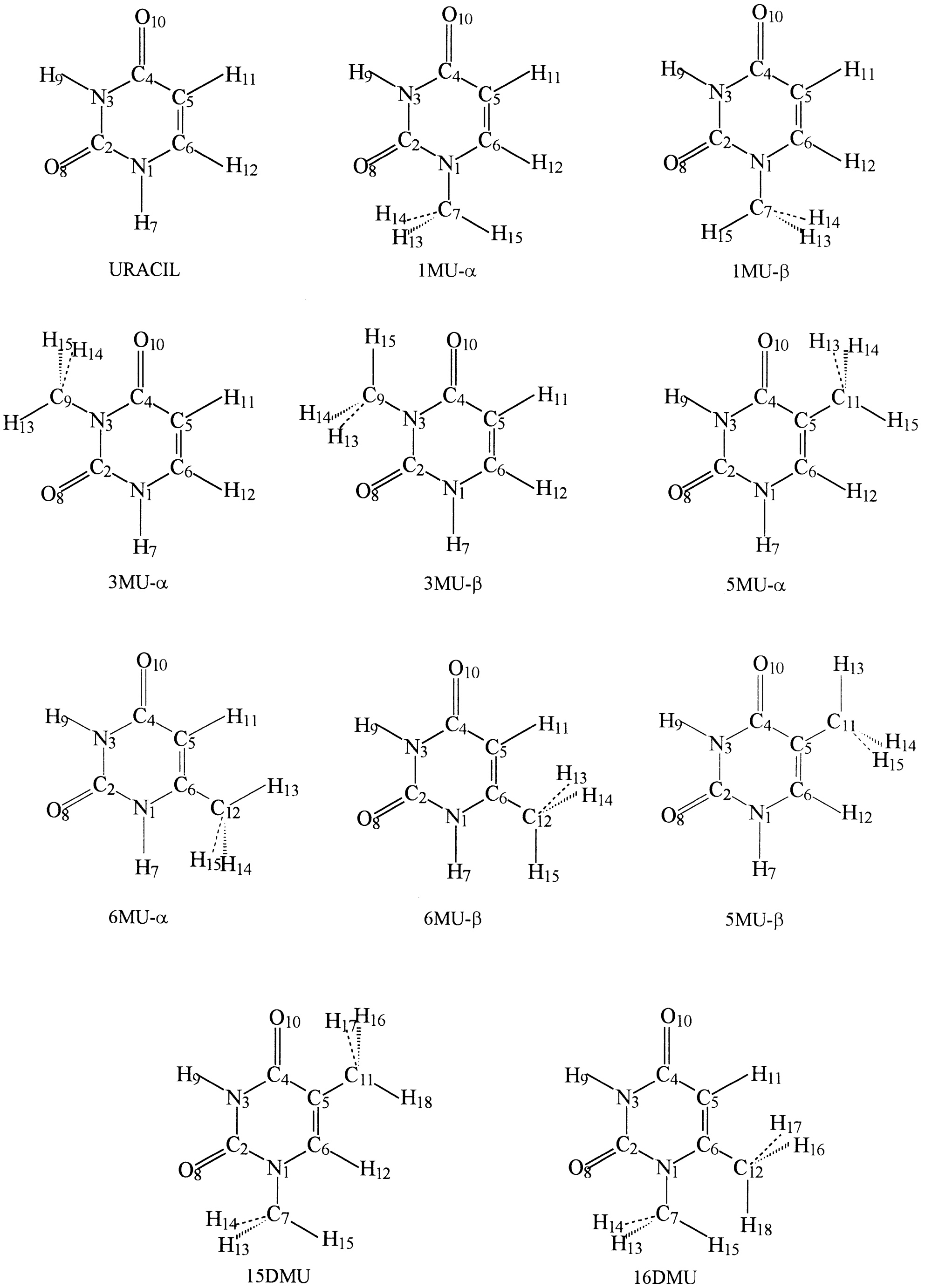
Due to the intermolecular interaction in the crystal, the calculated bond lengths (gas phase) are a little longer than the experimental results (room temperature X-ray diffraction and neutron diffraction data). The differences between calculational values and experimental values vary from as small as 0.001 Å ($C_2{=}O_8$ bond) to as large as 0.18 Å (N_1–H_7 bond). In general, the theoretically predicted bond lengths between heavy atoms (O, N, and C) are more precise than those between light atoms (H). Compared with the present calculational lengths of C–H bond and N–H bond with the only available X-ray diffraction data by Stewart [21], the differences are all above 0.1 Å (0.18 Å for N_1–H_7, 0.14 Å for N_3–H_9, 0.15 Å for C_5–H_{11}, and 0.13 Å for C_6–H_{12}). But for the C=O bond, C–N bond, C–C bond, and C=C bond, respectively, the differences are all below 0.03 Å. This indicates that the influence of the intermolecular interaction in the crystal is more significant on light atoms than on heavy atoms. Compared with the differences of bond lengths, the differences of the calculated bond angles are much less. Almost all the deviations between calculational and experimental values are below 1°, with the two exceptions N_3–$C_2{=}O_8$ bond angle and C_2–N_3–H_9 bond angle, being 1.9° and 2.1°, respectively. It seems that the atom O_{10} will interact strongly with the atoms H_9 and O_8 in crystal, which results in a closer distance between them and thus a larger C_2–N_3–H_9 bond angle and a smaller N_3–$C_2{=}O_8$ bond angle. Most interestingly, the predicted bond angles concerning hydrogen atoms are generally in rather good agreement with the experimental results, indicating that intermolecular interaction has little influence in crystal on the bond angles.

3.2. The single-methylated uracil

When there is only one hydrogen atom being substituted by the methyl group, there could be four possible positions – N_1, N_3, C_5, and C_6; the derivatives will be 1MU, 3MU, 5MU (thymine), and 6MU, respectively (Fig. 1).

Through full geometry optimizations, we find that there are two possible stable conformations for each single-methylated derivative, which we call conformation α and conformation β, respectively. They are different from each other in the spatial orientation of the methyl group, and the total energy (electronic energy plus zero point energy) of α is a little lower than that of β, indicating that α is more stable than β (which will be discussed in detail in Section 3.4). Their detailed structures are shown in Fig. 1 and the geometrical parameters are tabulated in Table 1 (ring structures) and Table 2 (methyl group structures).

The structural optimizations of the methylated uracil demonstrate that, in the geometries of the substituting methyl group, while one of the three hydrogen atoms locates in the plane with the ring, the other two can be at the different sides, with the ring as the symmetrical plane. Both conformations α and β have the same structure like this but the spatial orientation of the methyl group is different (Fig. 1). The bond lengths of conformations α and β of each derivative are very similar to each other. The largest discrepancy occurs at the bonds N_1–C_2 (0.0025 Å), N_3–C_4 (0.0035 Å), C_5–C_{11} (0.0064 Å), and C_6–C_{12} (0.0059 Å) for 1MU, 3MU, 5MU, and 6MU, respectively. This indicates that the spatial orientation of the methyl group influences the ring



URACIL

1MU-α

1MU-β

3MU-α

3MU-β

5MU-α

6MU-α

6MU-β

5MU-β

15DMU

16DMU

Fig. 1. Geometrical structures of the studied molecules.

Table 1

Geometrical parameters (bond lengths R in Å, and bond angles A in degrees) of the studied methylated uracil obtained at B3LYP/6-31+G* level

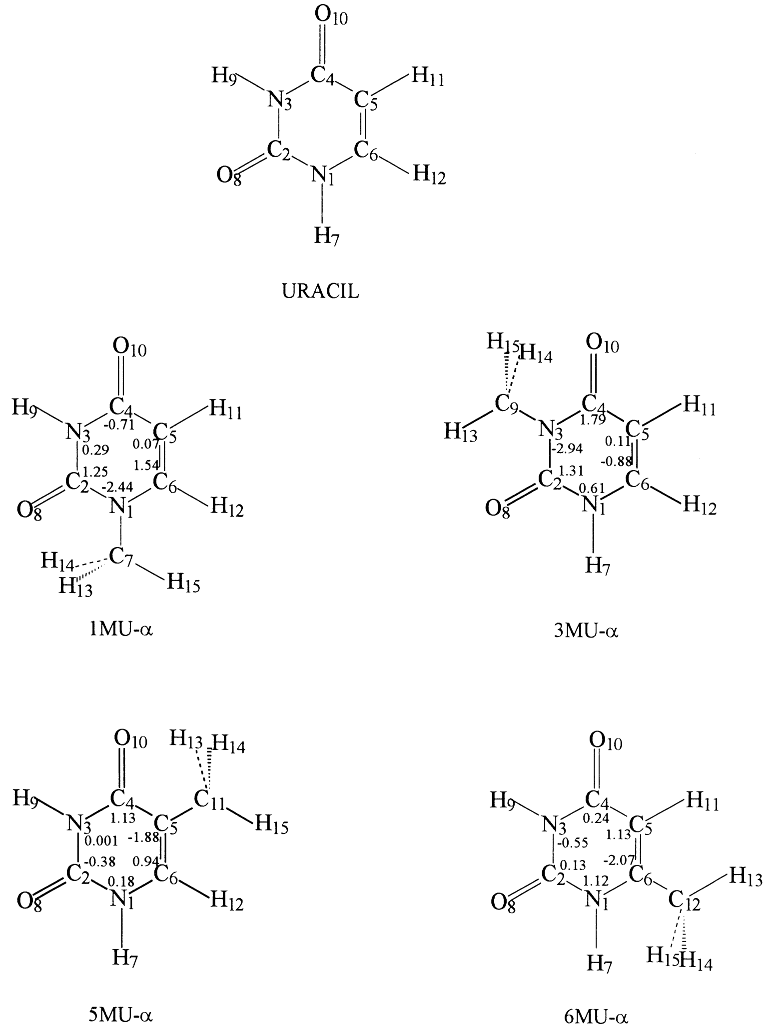
Parameter	Uracil	1MU		3MU		5MU		6MU		15DMU	16DMU
		α	β	α	β	α	β	α	β		
R(1,2)	1.3938	1.3982	1.4007	1.3937	1.3932	1.3884	1.3884	1.3928	1.3943	1.3919	1.4058
R(1,6)	1.3768	1.3772	1.3780	1.3729	1.3726	1.3813	1.3805	1.3843	1.3820	1.3823	1.3926
R(1,7)	1.0118	1.4655	1.4656	1.0118	1.0118	1.0116	1.0117	1.0126	1.0122	1.4652	1.4712
R(2,3)	1.3844	1.3839	1.3848	1.3911	1.3890	1.3851	1.3843	1.3819	1.3803	1.3844	1.3775
R(2,8)	1.2201	1.2227	1.2225	1.2234	1.2231	1.2214	1.2214	1.2212	1.2214	1.2241	1.2242
R(3,4)	1.4123	1.4100	1.4091	1.4198	1.4233	1.4076	1.4082	1.4114	1.4130	1.4044	1.4056
R(3,9)	1.0154	1.0154	1.0154	1.4685	1.4680	1.0154	1.0153	1.0151	1.0150	1.0139	1.0138
R(4,5)	1.4594	1.4546	1.4550	1.4568	1.4577	1.4682	1.4699	1.4562	1.4538	1.4633	1.4458
R(4,10)	1.2227	1.2239	1.2238	1.2253	1.2250	1.2247	1.2249	1.2240	1.2244	1.2260	1.2261
R(5,6)	1.3521	1.3542	1.3533	1.3501	1.3502	1.3541	1.3548	1.3571	1.3583	1.3560	1.3628
R(5,11)	1.0819	1.0820	1.0821	1.0819	1.0819	1.5023	1.5087	1.0823	1.0828	1.5018	1.0822
R(6,12)	1.0850	1.0852	1.0858	1.0850	1.0850	1.0856	1.0860	1.5010	1.5069	1.0851	1.5135
A(2,1,6)	123.6084	121.1728	121.1235	124.2192	123.8085	123.7937	123.7018	124.7341	124.7525	121.3859	121.2907
A(2,1,7)	115.0951	116.7120	118.7782	114.3701	114.6376	115.2263	115.2837	114.5541	114.4695	116.7824	112.9603
A(6,1,7)	121.2965	122.1152	120.0983	121.4107	121.5539	120.9799	121.0145	120.7118	120.7781	121.8317	125.7491
A(1,2,3)	113.0895	114.3398	114.0044	114.3989	114.9202	112.7123	112.6408	113.2193	113.2081	114.0330	115.5534
A(1,2,8)	122.6917	122.3021	123.2943	120.9182	121.4337	123.1672	123.1662	122.5272	122.3612	122.7188	121.6880
A(3,2,8)	124.2188	123.3580	122.7013	124.6828	123.6461	124.1205	124.1930	124.2535	124.4307	123.2482	122.7585
A(2,3,4)	127.9839	128.2709	128.5956	125.0410	124.9766	127.9847	128.1707	127.4261	127.4117	128.1895	127.4676
A(2,3,9)	115.6842	115.2941	115.0480	118.2802	116.0441	115.7751	115.7477	115.9982	116.0703	115.3934	115.4895
A(4,3,9)	116.3319	116.4349	116.3564	116.6788	118.9793	116.2402	116.0816	116.5756	116.5180	116.4170	117.0428
A(3,4,5)	113.6828	112.9690	112.9699	115.4725	115.0040	114.8053	114.7451	113.9192	113.9021	114.1594	112.8555
A(3,4,10)	120.2137	120.4091	120.4506	119.7904	120.8401	120.0990	119.6994	120.0814	119.9332	120.3529	120.3811
A(5,4,10)	126.1036	126.6220	126.5795	124.7371	124.1559	125.0957	125.5555	125.9994	126.1646	125.4877	126.7634
A(4,5,6)	119.8257	119.8987	119.6764	119.9429	120.3708	117.9520	117.7069	120.9613	121.0136	118.0438	122.3002
A(4,5,11)	118.1835	118.4655	118.5461	117.8122	117.5180	118.1896	118.8214	117.5051	117.5018	118.3733	117.3108
A(6,5,11)	121.9909	121.6359	121.7775	122.2448	122.1112	123.8584	123.4717	121.5335	121.4846	123.5828	120.3890
A(1,6,5)	121.8098	123.3488	123.6301	120.9255	120.9199	122.7521	123.0347	119.7399	119.7120	124.1884	120.5325
A(1,6,12)	115.3159	114.8673	114.5669	115.6964	115.7585	114.9222	114.7538	115.7708	116.4526	114.5763	120.9246
A(5,6,12)	122.8743	121.7838	121.8029	123.3781	123.3216	122.3257	122.2116	124.4893	123.8354	121.2353	118.5429

Table 2
Geometrical parameters concerning the substituting methyl obtained at B3LYP/6-31+G* level (bond lengths *R* in Å, and bond angles *A* in degrees)

Parameter	α	β
1MU		
R(7,13)	1.0931	1.0948
R(7,14)	1.0931	1.0948
R(7,15)	1.0912	1.0892
A(1,7,13)	110.21	110.18
A(1,7,14)	110.21	110.18
A(1,7,15)	108.54	107.96
A(13,7,14)	108.42	109.59
A(13,7,15)	109.73	109.45
A(14,7,15)	109.73	109.45
3MU		
R(9,13)	1.0881	1.0930
R(9,14)	1.0925	1.0930
R(9,15)	1.0931	1.0884
A(3,9,13)	108.08	109.85
A(3,9,14)	109.54	109.85
A(3,9,15)	109.54	107.42
A(13,9,14)	110.65	108.56
A(13,9,15)	110.65	110.58
A(14,7,15)	108.37	110.58
5MU		
R(11,13)	1.0961	1.0925
R(11,14)	1.0962	1.0965
R(11,15)	1.0950	1.0965
A(5,11,13)	110.91	109.38
A(5,11,14)	110.93	111.75
A(5,11,15)	111.10	111.75
A(13,11,14)	106.65	107.96
A(13,11,15)	108.56	107.96
A(14,11,15)	108.53	107.89
6MU		
R(12,13)	1.0919	1.0949
R(12,14)	1.0977	1.0949
R(12,15)	1.0977	1.0949
A(6,12,13)	110.58	110.52
A(6,12,14)	110.89	110.52
A(6,12,15)	110.89	111.99
A(13,12,14)	108.30	107.75
A(13,12,15)	108.30	107.95
A(14,12,15)	107.77	107.95

structures distinctly for 1MU and 3MU, but for 5MU and 6MU, it influences the bonds concerning the methyl group significantly. Compared with the minor discrepancy of bond lengths, the differences of bond angles are slightly larger. In 1MU, the C_2–N_1–C_7 bond angle of α is 2° smaller than that of β, and in 3MU, the C_2–N_3–C_9 bond angle of α is 2° larger than that of β. This phenomenon can be explained as follows. Because the attraction between the atom O_8 (negative) and the methyl group (positive) is stronger when the two symmetrical hydrogen atoms are closer to O_8 than they are farther, there will be a smaller distance between O_8 and C_7 in 1MU-α than in 1MU-β and a larger distance between O_8 and C_9 in 3MU-α than in 3MU-β, which will lead to a smaller bond angle C_2–N_1–C_7 in 1MU-α than in 1MU-β and a larger bond angle C_2–N_3–C_9 in 3MU-α than in 3MU-β.

The bond lengths of ring structures of the single-methylated uracil molecules (conformation α) predicted by the B3LY/6-31+G* level are very similar to those of uracil, with a difference of no more than 0.01 Å. This indicates that the substitution of a hydrogen atom by a methyl group has little influence on the geometries of the six-membered ring. While the hydrogen atom is replaced by a methyl group, the previous N_1–H_7 (N_3–H_9) bond in 1MU (3MU) changed into the N_1–C_7 (N_3–C_9) bond and the C_5–H_{11} (C_6–H_{12}) bond in 5MU (6MU) into the C_5–C_{11} (C_6–C_{12}) bond. The bond lengths increase by 0.45 Å (0.45 Å) and 0.42 Å (0.42 Å), respectively. These increases are easy to understand in that the interaction between N–H (C–H) is stronger than that between N–C (C–C). Investigating the calculated inside-ring bond-angle differences between the four single-methylated uracil molecules and uracil (Fig. 2), we find some very interesting phenomena. For every single-methylated derivative studied here (conformation α), in the six inside-ring bond angles (Fig. 2), compared with those of uracil, two of them decrease and the other four increase. In the former two, one decreases slightly (mostly below 0.9° but above 0.4°) and the other decreases much more significantly (2.44° in the bond angle C_2–N_1–C_6 of 1MU-α, 2.94° in the bond angle C_2–N_3–C_4 of 3MU-α, 1.88° in the bond angle C_4–C_5=C_6 of 3MU-α, and 2.07° in the bond angle C_5–C_6–N_1 of 6MU-α). In the latter four, two increase slightly (mostly below 0.3°) and the other two increase much more significantly (mostly above 1.1°). It seems that the six-membered nitrogen-containing ring in the derivative is elongated along the direction of the substituting methyl group, resulting



URACIL

1MU-α

3MU-α

5MU-α

6MU-α

Fig. 2. Inside-ring bond angle differences (unit in degrees) between four single-methylated uracil molecules and uracil.

in a decrease in the bond angles along this direction and an increase in the bond angles perpendicular to this direction. This elongation effect can exactly explain the slight increase of the dipole moment and the molecular size resulting from methylation proposed by Desfrancois et al. [10]. From the larger differences of the bond angles between methylated derivatives and uracil parent compared with those of bond lengths, we can also draw a conclusion that the bond angles are more sensitive to the strong interaction between the substituting methyl group and the remaining section of the molecule than the bond lengths.

To our knowledge, concerning four single-methylated uracil molecules, there are only the experimental parameters of 5MU (thymine) available up to now [25,38]. In Table 3, we listed the comparison of present computational results with the previous available experimental data. There is an excellent agreement between them, especially the prediction of the ring structure. In general, the calculated results are a little higher than the experimental data, just as that discussed in Section 3.1 about uracil.

In applying DFT methods on predicting the geometrical parameters of NAB, we also use

Table 3
Calculated geometrical parameters (obtained at B3LYP/6-31+G* level) in comparison with the available experimental results (bond lengths R in Å, and bond angles A in degrees)

Parameter	Uracil						5MU			15DMU			
	Calculated	Experimental					Calculated	Experimental		Calculated	Experimental		
		a	b	c	d	e		f	g		h	i	j
$R(1,2)$	1.3938	1.371 (2$_5$)	1.374 (19)	1.379 (10)		1.381 (9)	1.3884	1.355	1.376 (8)	1.3919	1.379 (4)	1.378 (3)	1.388 (1)
$R(1,6)$	1.3768	1.358 (2$_2$)	1.370 (22)	1.380 (11)		1.375 (9)	1.3813	1.382	1.378 (7)	1.3823	1.383 (4)	1.373 (3)	1.367 (1)
$R(1,7)$	1.0118	0.836 (18)					1.0116			1.4652	1.470 (4)	1.460 (3)	1.468 (2)
$R(2,3)$	1.3844	1.376 (2$_0$)	1.381 (22)	1.373 (9)		1.373 (7)	1.3851	1.361	1.373 (8)	1.3844	1.379 (4)	1.368 (2)	1.373 (1)
$R(2,8)$	1.2201	1.215 (2$_1$)	1.219 (20)	1.218 (10)			1.2214	1.234	1.220 (8)	1.2241	1.225^a (3)	1.216 (3)	1.215 (2)
$R(3,4)$	1.4123	1.371 (2$_1$)	1.380 (13)	1.383 (10)		1.380 (9)	1.4076	1.391	1.382 (8)	1.4044	1.375 (4)	1.379 (2)	1.377 (1)
$R(3,9)$	1.0154	0.877 (21)					1.0154			1.0139		1.044 (5)	1.032 (3)
$R(4,5)$	1.4594	1.430 (2$_6$)	1.444 (24)	1.440 (11)	1.462 (8)	1.431 (9)	1.4682	1.447	1.445 (9)	1.4633	1.432 (4)	1.428 (3)	1.430 (1)
$R(4,10)$	1.2227	1.245 (1$_9$)	1.233 (23)	1.227 (9)		1.232 (8)	1.2247	1.231	1.228 (9)	1.2260	1.237 (4)	1.225 (3)	1.233 (2)
$R(5,6)$	1.3521	1.340 (2$_3$)	1.343 (26)	1.338 (9)	1.343 (24)	1.337 (9)	1.3541	1.349	1.339 (7)	1.3560	1.346 (4)	1.353 (3)	1.355 (2)
$R(5,11)$	1.0819	0.931 (20)					1.5023	1.503	1.496 (6)	1.5018	1.497 (4)	1.499 (3)	1.496 (3)
$R(6,12)$	1.0850	0.957 (20)					1.0856			1.0851		1.090 (5)	1.085 (5)
$A(2,1,6)$	123.6084	122.7 (0.1)	122.0 (1.4)	121.3 (0.6)	123.2 (1.2)	121.0 (0.6)	123.7937	122.8	121.3 (0.5)	121.3859	120.6 (0.2)	120.9 (0.2)	120.9 (0.1)
$A(2,1,7)$	115.0951	115.1 (1.3)					115.2263			116.7824	118.2 (0.2)	119.2 (0.2)	118.0 (0.1)
$A(6,1,7)$	121.2965	122.1 (1.4)					120.9799			121.8317	121.2 (0.2)	119.9 (0.2)	121.1 (0.1)
$A(1,2,3)$	113.0895	114.0 (0.1)	115.4 (1.5)	114.8 (0.7)	114.6 (2.0)	114.9 (0.6)	112.7123	115.2	114.6 (0.6)	114.0330	115.4 (0.2)	115.4 (0.2)	115.2 (0.1)
$A(1,2,8)$	122.6917	123.7 (0.2)	122.9 (1.3)	123.2 (0.8)	123.8 (1.4)	122.8 (0.7)	123.1672	122.7	123.1 (0.8)	122.7188	123.3 (0.2)	123.6 (0.2)	123.2 (0.1)
$A(3,2,8)$	124.2188	122.3 (0.2)		122.0 (0.7)		122.2 (0.7)	124.1205	122.1	122.3 (0.6)	123.2482	121.3 (0.2)	121.0 (0.2)	121.5 (0.1)

A(2,3,4)	127.9839	126.7 (0.2)	120.4 (1.4)	127.0 (0.6)	126.0 (1.4)	127.0 (0.6)	127.9847	126.3	127.2 (0.6)	128.1895	126.3 (0.2)	126.6 (0.2)	126.5 (0.1)
A(2,3,9)	115.6842	117.8 (1.0)					115.7751			115.3934		116.0 (0.3)	116.1 (0.2)
A(4,3,9)	116.3319	115.5 (1.0)					116.2402			116.4170		117.5 (0.3)	117.4 (0.2)
A(3,4,5)	113.6828	115.5 (0.1)	114.1 (1.6)	114.7 (0.9)	115.5 (1.8)	114.6 (0.6)	114.8053	115.6	115.2 (0.6)	114.1594	116.1 (0.2)	115.6 (0.2)	115.8 (0.1)
A(3,4,10)	120.2137	119.2 (0.2)	120.5 (1.5)	119.8 (0.7)		119.4 (0.7)	120.0990	118.3	119.9 (0.6)	120.3529	120.0 (0.2)	120.3 (0.2)	120.4 (0.1)
A(5,4,10)	126.1036	125.3 (0.2)		125.4 (1.0)	124.3 (2.0)	125.9 (0.6)	125.0957	126.1	124.9 (0.7)	125.4877	123.9 (0.2)	124.2 (0.2)	123.8 (0.1)
A(4,5,6)	119.8257	118.9 (0.2)	120.7 (1.9)	119.2 (1.3)	119.7 (2.1)	119.7 (0.6)	117.9520	118.2	118.0 (0.6)	118.0438	118.3 (0.2)	118.6 (0.2)	118.0 (0.1)
A(4,5,11)	118.1835	118.1 (1.2)					118.1896	119.0	119.0 (0.6)	118.3733	119.3 (0.2)	117.8 (0.2)	119.3 (0.1)
A(6,5,11)	121.9909	123.0 (1.2)					123.8584	122.8	122.9 (0.6)	123.5828	122.4 (0.2)	123.5 (0.2)	122.7 (0.1)
A(1,6,5)	121.8098	122.3 (0.2)	121.2 (1.2)	122.8 (0.9)	122.1 (2.2)	122.7 (0.5)	122.7521	121.8	123.7 (0.6)	124.1884	123.3 (0.2)	122.8 (0.2)	123.5 (0.1)
A(1,6,12)	115.3159	114.5 (1.0)					114.9222			114.5763			115.9 (0.2)
A(5,6,12)	122.8743	123.2 (1.0)					122.3257			121.2353			120.6 (0.2)

a: X-ray diffraction data from Ref. [21]; b: Ref. [22]; c: Ref. [23]; d: electron diffraction data from Ref. [24]. (N–C_{mean} 1.399(6) C=O_{mean} 1.212(3)); e: statistical data from Ref. [25]; f: Ref. [38]; g: statistical data from Ref. [25]; h: X-ray diffraction data from Ref. [42]; i: neutron diffraction data from Ref. [43]; j: neutron diffraction data from Ref. [44]. (Standard deviation in parentheses.)

[a] Temperature movement corrected value.

different basis sets to testify the influence of inclusion of polarization functions (*) and diffuse functions (+) on the calculational results. The geometrical parameters of 5MU (thymine) which are calculated at four different levels (B3LYP/ 6-31G, B3LYP/6-31+G, B3LYP/6-31G*, B3LYP/ 6-31+G*), taken as an example, are presented in Table 4, along with the most recent statistical data from Ref. [24] for a comparison. We do not find significant differences in the bond lengths and bond angles predicted by these four different levels. Almost all of the bond-length differences are below 0.005 Å and the bond-angle differences below 0.5°. The exceptions occur at the double bond $C_2{=}O_8$

Table 4
Geometrical parameters of 5MU-α (thymine) obtained at different calculational levels (B3LYP/6-31G level, B3LYP/6-31+G level, B3LYP/6-31G* level, B3LYP/6-31+G* level, respectively) along with the experimental results (bond lengths *R* in Å, and bond angels *A* in degree) for a comparison

Parameter	Calculated				Experimental[a]
	B3LYP/ 6-31G	B3LYP/6-31+G	B3LYP/ 6-31G*	B3LYP/ 6-31+G*	
R(1,2)	1.3932	1.3908	1.3904	1.3884	1.376(8)
R(1,6)	1.3874	1.3887	1.3802	1.3813	1.378(7)
R(1,7)	1.0091	1.0103	1.0104	1.0116	
R(2,3)	1.3897	1.3882	1.3864	1.3851	1.373(8)
R(2,8)	1.2434	1.2470	1.2174	1.2214	1.220(8)
R(3,4)	1.4104	1.4092	1.4081	1.4076	1.382(8)
R(3,9)	1.0126	1.0139	1.0139	1.0154	
R(4,5)	1.4648	1.4650	1.4683	1.4682	1.445(9)
R(4,10)	1.2477	1.2504	1.2216	1.2247	1.228(9)
R(5,6)	1.3568	1.3588	1.3521	1.3541	1.339(7)
R(5,11)	1.5018	1.5029	1.5010	1.5023	1.496(6)
R(6,12)	1.0831	1.0835	1.0853	1.0856	
R(11,13)	1.0961	1.0962	1.0959	1.0961	
R(11,14)	1.0961	1.0963	1.0959	1.0962	
R(11,15)	1.0942	1.0946	1.0945	1.0950	
A(2,1,6)	123.7106	123.5917	123.9439	123.7937	121.3(0.5)
A(2,1,7)	115.3839	115.5368	114.9941	115.2263	
A(6,1,7)	120.9055	120.8715	121.0620	120.9799	
A(1,2,3)	112.8896	113.2586	112.3985	112.7123	114.6(0.6)
A(1,2,8)	123.0519	122.9257	123.2282	123.1672	123.1(0.8)
A(3,2,8)	124.0584	123.8157	124.3733	124.1205	122.3(0.6)
A(2,3,4)	127.6582	127.3033	128.2685	127.9847	127.2(0.6)
A(2,3,9)	116.0229	116.1578	115.6481	115.7751	
A(4,3,9)	116.3190	116.5389	116.0834	116.2402	
A(3,4,5)	114.9933	115.3336	114.5193	114.8053	115.2(0.6)
A(3,4,10)	120.2506	119.9553	120.4178	120.0990	119.9(0.6)
A(5,4,10)	124.7561	124.7109	125.0627	125.0957	124.9(0.7)
A(4,5,6)	118.3643	118.1378	118.1902	117.9520	118.0(0.6)
A(4,5,11)	117.8336	118.1288	117.7419	118.1896	119.0(0.6)
A(6,5,11)	123.8021	123.7335	124.0679	123.8584	122.9(0.6)
A(1,6,5)	122.3840	122.3750	122.6796	122.7521	123.7(0.6)
A(1,6,12)	115.3992	115.3566	114.9857	114.9222	
A(5,6,12)	122.2168	122.2684	122.3346	122.3257	
A(5,11,13)	110.8756	110.9240	110.8554	110.9137	
A(5,11,14)	110.8809	110.9292	110.8597	110.9316	
A(5,11,15)	111.3850	111.1925	111.3173	111.1011	
A(13,11,14)	106.4228	106.4984	106.4799	106.6494	
A(13,11,15)	108.5563	108.5706	108.5861	108.5637	
A(14,11,15)	108.5513	108.5682	108.5809	108.5325	

[a] Statistical data from Ref. [25]. (Standard deviation in parentheses.)

and $C_4{=}O_{10}$. For $C_2{=}O_8$, the bond length predicted by B3LYP/6-31+G* level is 0.0256 Å smaller than that by B3LYP/6-31+G level, and that by B3LYP/6-31G* level is 0.0260 Å smaller than that by B3LYP/6-31G level. For $C_4{=}O_{10}$, the bond length predicted by B3LYP/6-31+G* level is 0.0257 Å smaller than that by B3LYP/6-31+G level, and that by B3LYP/6-31G* level is 0.0261 Å smaller than that by B3LYP/6-31G level. These indicate that the addition of diffuse functions (+) to the heavy atoms (nitrogen and oxygen atoms) contributes little to the geometry optimizations of the ring structures of NAB, but the addition of polarization functions (*) to the heavy atoms (nitrogen and oxygen atoms) has significant influence on the bonds connected with oxygen atoms.

Although there have been reports concerning the conformational flexibility of NAB [39–41], our results showed that the six-membered nitrogen-containing ring of NAB is predominately planar. The replacement of the hydrogen atom by a methyl group does not change the planarity of the ring significantly.

3.3. The double-methylated uracil

Also studied here are two of the double-methylated uracil – 15DMU(1-methylthymine) and 16DMU. Their structural parameters are tabulated in Table 1.

Similar to what has been discussed in Section 3.2, for 15DMU, the substitution of two methyl groups for two hydrogen atoms at C_1 and C_5, respectively, has little influence on the ring structures. The largest difference for the bond lengths is 0.0079 Å in N_3–C_4 and 0.0055 Å in N_1–C_6. Replacement of H_7 by C_7 changes the N_1–H_7 bond into the N_1–C_7 bond and increases the bond length by 0.45 Å, and replacement of H_{11} by C_{11} changes the C_5–H_{11} bond into the C_5–C_{11} bond and increases the bond length by 0.42 Å (Table 1), just as that demonstrated in the single-methylated uracil.

The comparison of the present work with previous experimental parameters [42–44] are listed in Table 3. The most recent data we have found are the neutron diffraction data in 1974 [44]. Comparing with this work gives a satisfactory agreement. In the bond lengths N_1–C_6 and $C_2{=}O_8$, the present work is more consistent with the X-ray diffraction data in 1963 by Hoogsteen [42], that is 1.382–1.383 Å in the N_1–C_6 bond and 1.224–1.225 Å in the $C_2{=}O_8$ bond.

3.4. Total energies of the single-methylated uracil

The total energies of the four single-methylated uracil molecules are listed in Table 5. The energies predicted by the three different methods (B3LYP/6-31+G*, B3LYP/6-31G, and BLYP/6-31+G*) are very similar to each other, with a difference below 0.2 Hartree/atom (13.3 kJ mol^{-1}). All the three methods yield the same stability order, being 6MU > 5MU > 1MU > 3MU. Even though, the energy differences are very slight. The total energy of 5MU is higher than 6MU only by 0.0036 Hartree (0.24 kJ mol^{-1}). The energy difference between 1MU and 3MU is even more little, being only 0.0005 Hartree (33.3 J mol^{-1}). This indicates that it is perhaps very difficult to identify them by experimental methods.

Also presented in Table 5 are the total energy of conformations α and β of the four molecules predicted by B3LYP/6-31+G* method. The energy differences between α and β are 0.0001 Hartree

Table 5
Total energies[a] (unit in Hartree) of the four single-methylated uracil molecules obtained at different calculational levels

Methods	1MU		3MU		5MU		6MU	
	α	β	α	β	α	β	α	β
B3LYP/6-31+G*	−454.0338	−454.0337	−454.0333	−454.0332	−454.0432	−454.0413	−454.0468	−454.0450
B3LYP/6-31G	−453.8697	na	−453.8683	na	−453.8816	na	−453.8843	na
BLYP/6-31+G*	−453.9107	na	−453.9099	na	−453.9196	na	−453.9227	na

[a] Total energy is the electronic energy plus the zero-point energy.

(6.65 J mol^{-1}), 0.0001 Hartree (6.65 J mol^{-1}), 0.0019 Hartree (126.4 J mol^{-1}), and 0.0018 Hartree (119.7 J mol^{-1}) for 1MU, 3MU, 5MU, and 6MU, respectively. Such a little difference makes it almost impossible to recognize them in experiments.

4. Conclusions

In this work, we performed full geometry optimizations on the methylated derivatives of uracil applying DFT methods. The calculated results were compared with previously available experimental data and a rather satisfactory agreement was obtained. From analyzing the calculated structural parameters, we draw some important conclusions as follows:

1. There are two stable conformations, α and β, for each single-methylated uracil. They are different in the spatial orientation of the substituting methyl group and the molecular total energy. But the energy differences of conformations α and β are very slight, being 0.0001 Hartree, 0.0001 Hartree, 0.0019 Hartree, and 0.0018 Hartree for 1MU, 3MU, 5MU, and 6MU, respectively.
2. The replacement of a hydrogen atom in uracil by the methyl group results in an elongation of the six-membered nitrogen-containing ring. This elongation has little influence on the bond lengths of the ring structures, but it influences the bond angles more significantly. The increase of dipole moment and molecular size [10] can also be explained with respect to this elongation effect.
3. The methylation of uracil does not change the planarity of the six-membered nitrogen-containing ring significantly. The ring is predominantly planar in the derivatives.
4. The inclusion of diffuse and polarization functions in the basis sets contributes little to the geometry optimizations of NAB, with an exception that the addition of polarization functions to heavy atoms (nitrogen and oxygen atoms) decrease the lengths of the double bond connected with oxygen atoms.

Acknowledgements

Financial support for this research work was provided by the National Natural Science Foundation of China (Grant No. 19634040 and No. 19974040). We also gratefully acknowledged the Youth Foundation of University of Science and Technology of China (USTC) for financial support. Here, one of the authors also wants to express his sincere thanks for Dr. Juguang Han in the School of Life Science of USTC for his wholehearted assistance in the discussions.

References

[1] J.D. Watson, F.H.C. Crick, Nature 171 (1953) 737.
[2] J.D. Watson, F.H.C. Crick, Nature 171 (1953) 964.
[3] M.K. Shukla, P.C. Mishra, Chem. Phys. 240 (1999) 319.
[4] K. Szczepaniak, W.B. Person, J. Leszczynski, J.S. Kwiatkowski, Polish J. Chem. 72 (1998) 402.
[5] P. Colarusso, K.Q. Zhang, B. Guo, P.F. Bernath, Chem. Phys. Lett. 269 (1997) 39.
[6] S.L. Zhang, K.H. Michaelian, G.R. Loppnow, J. Phys. Chem. A 102 (1998) 461.
[7] S. Alexjander, D. Deamer, IEEE Engng. Med. Biol. Mag. 18 (1999) 74.
[8] M. Aida, M. Kaneko, M. Dupuis, T. Ueda, K. Ushizawa, G. Ito, A. Kumakura, M. Tsuboi, Spectrochim. Acta A 53 (1997) 393.
[9] C. Desfrancois, V. Periquet, Y. Bouteiller, J.P. Schermann, J. Phys. Chem. A 102 (1998) 1274.
[10] C. Desfrancois, H. Abdoul-Carime, S. Carles, V. Periquet, J.P. Schermann, D.M.A. Smith, L. Adamowicz, J. Chem. Phys. 110 (1999) 11876.
[11] I. Galetich, M. Kosevich, V. Shelkovsky, S.G. Stepanian, Yu.P. Blagoi, L. Adamowicz, J. Mol. Struct. 478 (1999) 155.
[12] S.X. Tian, C.F. Zhang, Z.J. Zhang, X.J. Chen, K.Z. Xu, Chem. Phys. 242 (1999) 217.
[13] E.L. Stewart, C.K. Foley, N.L. Allinger, J.P. Bowen, J. Am. Chem. Soc. 116 (1994) 7282.
[14] M. Szczesniak, M.J. Nowak, H. Rostkowska, K. Szczepaniak, W.B. Person, D. Shugar, J. Am. Chem. Soc. 105 (1983) 5969.
[15] G. Bakalarski, P. Grochowski, J.S. Kwiatkowski, B. Lesyng, J. Leszczynski, Chem. Phys. 202 (1996) 301.
[16] M.J. Frisch et al., GAUSSIAN 98W, Revision A.7, Gaussian Inc., Pittsburgh, PA, 1998.
[17] A.D. Becke, Phys. Rev. A 38 (1988) 3098.
[18] C. Lee, W. Yang, R.G. Parr, Phys. Rev. B 37 (1988) 785.
[19] A.D. Becke, J. Chem. Phys. 98 (1993) 5648.
[20] W.J. Hehre, L. Radom, P. von R. Schleyer, J.A. Pople, Ab Initio Molecular Orbital Theory, Wiley, New York, 1986.

[21] R.F. Stewart, Acta Cryst. 23 (1967) 1102.

[22] D. Voet, A. Rich, Prog. Nucleic Acid Res. Mol. Biol. 10 (1970) 247.

[23] R. Taylor, O. Kennard, J. Mol. Struct. 78 (1982) 1.

[24] G. Ferenczy, L. Harsanyi, B. Rozondai, I. Hargittai, J. Mol. Struct. 140 (1986) 71.

[25] L. Clowney, S.C. Jain, A.R. Srinivasan, J. Westbrook, W.K. Olson, H.M. Berman, J. Am. Chem. Soc. 118 (1996) 509.

[26] A. Padva, P.R. LeBreton, R.J. Dinerstein, J.N.A. Ridyard, Biochem. Biophys. Res. Commun. 60 (1974) 1262.

[27] A. Padva, T.J. D'Donnell, P.R. LeBreton, Chem. Phys. Lett. 41 (1976) 278.

[28] D. Dougherty, K. Wittel, J. Meeks, S.P. McGlynn, J. Am. Chem. Soc. 98 (1976) 3815.

[29] M. Fujii, T. Tamura, N. Mikami, M. Ito, Chem. Phys. Lett. 126 (1986) 583.

[30] Y. Tsuchiya, M. Fujii, M. Ito, J. Phys. Chem. 92 (1988) 1760.

[31] B.B. Brady, L.A. Peteanu, D.H. Leavy, Chem. Phys. Lett. 147 (1988) 538.

[32] M. Kubota, T. Kobayashi, J. Electron Spectrosc. Relat. Phenom. 82 (1996) 61.

[33] M. Szczesniak, M.J. Nowak, K. Szczepaniak, W.B. Person, D. Shugar, J. Am. Chem. Soc. 105 (1983) 5969.

[34] S. Chin, I. Scot, K. Szczepaniak, W.B. Person, J. Am. Chem. Soc. 106 (1984) 3415.

[35] A.R. Katritzky, M. Karelson, J. Am. Chem. Soc. 113 (1991) 1561.

[36] M.J. Scanlan, I.H. Hillier, J. Am. Chem. Soc. 106 (1984) 3737.

[37] J. Leszczynski, J. Phys. Chem. 96 (1992) 1642.

[38] D. Voet, A. Rich, Prog. Nucleic Acid Res. Mol. Biol. 10 (1970) 183.

[39] O.V. Shishkin, J. Mol. Struct. 447 (1998) 1.

[40] J. Sponer, P. Hobza, J. Phys. Chem. 98 (1994) 3161.

[41] J. Sponer, P. Hobza, J. Am. Chem. Soc. 116 (1994) 709.

[42] K. Hoogsteen, Acta Cryst. 16 (1963) 28.

[43] M.N. Frey, T.F. Koetzle, M.S. Lehmann, W.C. Hamilton, J. Chem. Phys. 59 (1973) 915.

[44] A. Kvick, T.F. Koetzle, R. Thomas, J. Chem. Phys. 61 (1974) 2711.