Protonation and Deprotonation Enthalpies of Guanine and Adenine and Implications for the Structure and Energy of Their Complexes with Water: Comparison with Uracil, Thymine, and Cytosine

Asit K.Chandra,[†] Minh Tho Nguyen,[†] Tadafumi Uchimaru,[‡] and Thérèse Zeegers-Huyskens*,[†]

Department of Chemistry, University of Leuven, Celestijnenlaan 200F, B-3001 Heverlee, Belgium and Department of Physical Chemistry, National Institute of Materials and Chemical Research, Tsukuba, 305 Ibaraki, Japan

The optimized geometries, harmonic vibrational frequencies, and energies of the cyclic structures of monohydrated guanine and adenine are computed using density functional theory (B3LYP) combined with the 6-31+G(d,p) basis set. The proton affinity of the O and N atoms and the deprotonation enthalpy of the different NH bonds of guanine and adenine are computed at the same level of theory. The results are compared with recent data on uracil, thymine, and cytosine. The intrinsic acidities and basicities of the five nucleobases are discussed. Complex formation with water results in a moderate change of the pyramidal character of the amino group. For closed complexes where water interacts with the O atom of the nucleobase, the intermolecular distances and the hydrogen bond energies are correlated to the proton affinities and deprotonation enthalpies of the sites involved in complex formation.

Introduction

The complex network of hydrogen bond interactions that modulate recognition of DNA or RNA bases is based on the assumption of specific tautomeric and ionic states.[1–3] The interaction energy between two complementary nucleobases which are held together by NH···O and NH···N hydrogen bonds is expected to depend not only on the intrinsic basicity of the acceptor atoms but also on the acidity of the NH donor groups. Löwdin's mutational mechanism[4–6] which involves a concerted transfer of two protons in the interbase hydrogen bonds, must also be governed by the proton donor and proton acceptor ability of the centers involved in the proton transfer process. In our recent papers, the proton affinities (PA) of the different basic sites and the deprotonation enthalpies of the NH bonds of uracil, thymine, and cytosine have been computed using DFT combined with the 6-31G(d,p) basis set.[7–9] To our knowledge, however, only low level quantum chemical calculations of the protonation and deprotonation enthalpies of guanine and adenine have been reported in the literature. The PA of the different sites of guanine and adenine have been computed at the HF/4-31G//STO-3G level,[10] and the relative acidities have been roughly estimated by AM1 or PM3 methods.[11,12] In this work, we will consider only the tautomers present in the Watson–Crick structure. Several papers have discussed the tautomerism of guanine[13–19] and adenine,[20–22] but a review of these studies is beyond the scope of this work.

Understanding the biological role of the nucleic acids depends on a knowledge of physical and chemical behavior of the complexes between nucleobases and water. A precise experimental evaluation of the relative stability of monohydrated guanine and adenine complexes still does not exist, and ab initio calculations have been carried out only at a low level.[23] More recently, the cyclic structure of monohydrated oxo-amino-N9H

guanine computed by MP2/6-31G(d) calculations has been predicted to be more stable than the two other enol–water complexes.[24] No other calculations at similar levels for water interacting with the other sites of guanine are available in the literature.

Nonplanarity of the nucleobases has recently attracted considerable attention.[25–33] Water can also have an influence on the amino group nonplanarity, and this has been recently discussed for the oxo-amino-N9H tautomer of guanine complexed with one or two water molecules at the C=O,N1H side.[34] The influence of water on the geometry of the amino group for the other possible complexes on the other sites of guanine has not been investigated. The same remark also holds for the different complexes between adenine and water.

The major result of this work is an accurate calculation of the PAs and deprotonation enthalpies of the different sites of guanine and adenine. The geometries and energies of these nucleobases complexed with one water molecule are considered in relation to these parameters and the correlations previously established for uracil,[7] thymine,[8] and cytosine[9] generalized. Calculations have recently reported the energies of the different complexes formed between uracil and *one* water molecule,[35,36] but no attempt has been made to explain the relative order of these stabilities.

Computational Methods

The geometries of the isolated guanine and adenine molecules and their corresponding water complexes were optimized using the density functional theory (DFT) with the B3LYP[37,38] exchange correlation functional and 6-31+G(d,p) basis functions. Basis set superposition errors (BSSE) were estimated by the counterpoise (CP) method.[39] The proton affinities and deprotonation enthalpies were computed at the same level. For the purpose of comparison, parameters previously computed with the 6-31++G(d,p) basis set[7–9] for uracil, thymine, and

---

[†] University of Leuven.
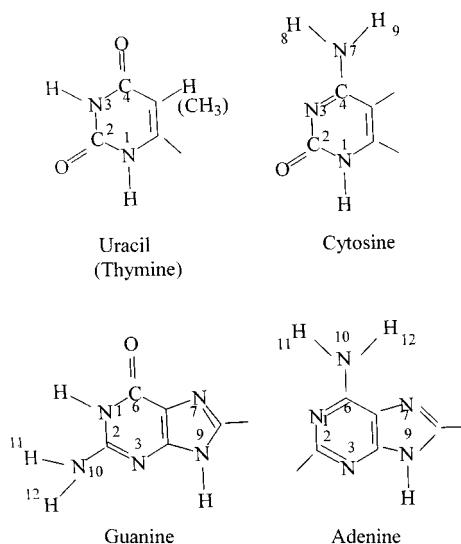[‡] National Institute of Materials and Chemical Research.

Figure 1. Atom labeling in uracil, cytosine, guanine, and adenine.

cytosine were recomputed at this level. The *Gaussian 94* package[40] was used for all of the calculations. Harmonic vibrational frequencies were calculated at the same level of theory to characterize the stationary points and to evaluate the frequency shifts resulting from complex formation with water. High-level DFT methods have significantly narrowed the gap between the computed and the experimental frequencies, and this has been discussed elsewhere.[41-45] Although the DFT method strongly underestimates the stabilization energies of stacked DNA base pairs,[46] it gives good results for hydrogen-bonded complexes. As shown in previous work,[7] the energy and intermolecular distances obtained from B3LYP calculations for the uracil−water interaction are comparable with the ones calculated from MP2 calculations performed with the same basis set.[35]

Results and Discussion

1. Structure and Vibrational Properties of the Guanine−Water Complexes. The B3LYP/6-31+G(d,p) optimized geometries of the four closed guanine−water complexes are shown in Figure 2. In these optimized geometries, only the amino group hydrogen atoms are considered as nonplanar (NPA). Characteristic geometrical parameters of free guanine and its four water complexes are indicated in Table 1.[47]

The distances and angles in free guanine are very similar to those computed at the MP2/6-31G(d,p) level.[32] Our DFT calculations show, however, a lesser degree of pyramidalization of the amino group. This appears clearly from the sum of the angles around the N2 atom (ΣAH) and the N1C2N10H11 dihedral angle (339.1° and 38.7°)[30] as compared with the values of 348.1° and 27.3° found in this work. This lesser degree of pyramidalization is also reflected by the larger C2N10 distance obtained in this work (1.384 Å) as compared with the value of 1.374 Å cited in ref 32. After full geometry optimization, we found the N10 atom 2.7° out of the plane of the guanine ring.

In complexes A, B, and C, one of the water hydrogens is out of the plane of the ring. In structure D, water acts as a bidonor and lies in the plane of the guanine ring. In complex A, the $(O_w)H'\cdots O$ and $(N)H\cdots O_w$ distances are very similar to the values of 1.930 and 1.890 Å obtained from MP2/6-31G* geometry optimization.[26,36] Owing to anticooperativity, the elongation of the C6=O bond and of the OH′ bond of water are smaller in complex D than in complex A. The intermolecular
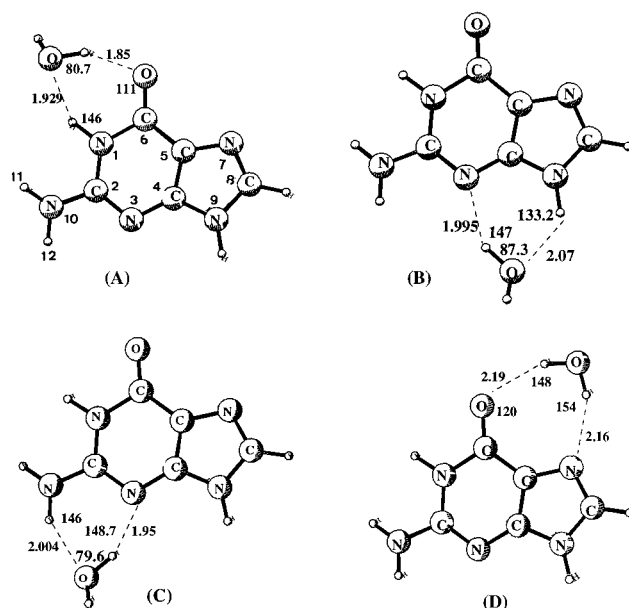

Figure 2. Optimized structures for the A, B, C, and D complexes between guanine and water obtained from B3LYP/6-31+G(d,p) geometry calculations (distances in Å, angles in deg).

distances $(O_w)H'\cdots N$ and $(O_w)H'\cdots O$ are also longer for complex D. Complex formation results in moderate changes of the pyramidalization of the amino group, and this will be discussed in more in detail in section 3.

The vibrational spectrum of guanine has been thoroughly discussed.[13,15,48-52] No data are, however, available for the guanine−water complex in low-temperature matrices, and this probably results from the fact that amino-oxo-N9H, amino-oxo-N7H, and amino-hydroxy tautomers are simultaneously present in these low-temperature materials and thus make the assignment of the absorptions in the water complexes rather speculative. Relevant B3LYP/6-31+G(d,p) unscaled vibrational frequencies in free guanine and the four water complexes are listed in Table 2. The νNH, νNH2, and νC=O vibrations of the bonds involved in the interaction with one water molecule are shifted downward. The opposite effect is observed for the in-plane δNH modes, the rocking and torsion vibrations of the NH2 group being coupled with other modes. Interestingly, in both complexes A and D, where the NH2 bonds are not involved in complex formation, the νNH2 vibrations are blue shifted by 7 to 11 cm⁻¹, and this effect results from the increasing sp² character of the amino nitrogen atom. Upward shifts of more than 200 cm⁻¹ are predicted for the almost pure γN1H vibration (complex A) and for the γN9H (complex B). The δNH vibrations appear to be less sensitive to complex formation and are coupled with other modes. In complexes A, B, and D, the torsion mode coupled with the wagging mode is shifted to lower frequencies and this may be related to the increased planarity of the amino group. In complex C where one NH bond of the amino group is directly involved in complex formation, three components having a high γN10H11 character are predicted at 693, 685, and 661 cm⁻¹.

2. Structure and Vibrational Properties of the Adenine−Water Complexes. B3LYP/6-31+G(d,p) optimized geometries of free adenine and the three closed adenine−water complexes are shown in Figure 3. Relevant geometrical parameters are indicated in Table 3.[47] As for guanine, the deviation from 360° of the ΣAH values and the dihedral angles are smaller than those computed from MP2/6-31G(d) calculations (349.3, 18.7, and − 21.1).[32] The larger ΣAH value and the shorter C6−N10

Guanine and Adenine Complexes with Water

J. Phys. Chem. A, Vol. 103, No. 44, 1999 **8855**

TABLE 1: Results of B3LYP/6-31+G(d,p) Geometry Optimization of Free Guanine and the Four Water Complexes A, B, C, and D (Bond Lengths in Å, Bond Angles in Deg)[a]

	free guanine	complex A	complex B	complex C	complex D
C6=O	1.221	1.236	1.220	1.221	1.226
N1−H	1.013	1.025	1.013	1.014	1.014
C2−N10	1.374	1.370	1.374	1.361	1.370
N9−H	1.010	1.010	1.018	1.010	1.010
N10−H11	1.010	1.009	1.010	1.008	1.010
N10−H12	1.009	1.009	1.009	1.019	1.009
∠C2N10H11	118.9	119.3	118.6	119.9	119.6
∠C2N10H12	114.2	115.1	114.7	116.5	114.8
∠H11N10H12	115.3	116.3	114.9	117.9	115.9
∠N1C2N10H11	+27.3	+22.1	+25.5	+18.3	+24.4
∠N3C2N10H12	−11.5	−12.3	−13.4	−8.3	−10.8
δ	11.6	9.3	11.8	5.7	9.7
Δ	15.8	9.8	12.1	10	13.6

	free water	complex A	complex B	complex C	complex D
O−H	0.965	0.964	0.964	0.964	0.971[b]
O−H′	0.965	0.984	0.979	0.981	0.969[c]
∠HOH′	105.7	107.2	107.3	107.2	102.3

Intermolecular Parameters

complex A		complex B		complex C		complex D	
H′···O	1.852	H′···N3	1.995	H′···N3	1.954	H′···N7	2.161
H···O$_w$	1.929	H···O$_w$	2.070	H···O$_w$	2.004	H′···O	2.199
∠N1HO$_w$	146.0	∠N9HO$_w$	133.2	∠N3H′O$_w$	148.7	∠O$_w$H′O	147.8
∠O$_w$H′O	145.5	∠O$_w$H′N3	147.0	∠O$_w$H12N10	146.2	∠O$_w$H′N7	153.8

[a] Only the amino group nitrogen atoms are nonplanar. [b] Water bonded to the O atom. [c] Water bonded to the N7 atom.

TABLE 2: Unscaled Characteristic B3LYP/6-31+G(d,p) Vibrational Frequencies (cm^{-1}) in Free Guanine and Water and the Water Complexes A, B, C, and D[a,b]

assignment	free guanine	complex A	complex B	complex C	complex D
νNH2	3714	3725(+11)	3713(−1)	3699(−15)	3721(+7)
νN9H	3658	3658(0)	3523(−135)	3660(+2)	3658(0)
νN1H	3600	3391(−209)	3601(+1)	3560(−40)	3594(−6)
νNH2	3592	3601(+9)	3594(+2)	3461(−131)	3601(+9)
νC=O	1796	1767(−29)	1800(+4)	1798(+2)	1783(−13)
δNH2 + δR	1665	1669(+4)	1668(+3)	1687(+22)	1682(+17)
δN1H	1337	1366(+29)	1337(0)	1347(+10)	1340(+3)
δR + rNH2	1138	1156(+18)	1138(0)	1147(+9)	1144(+6)
rNH2 + δN1H	1045	1061(+16)	1052(+7)	1056(+11)	1042(−3)
γN1H	594	815(221)	591(−3)	587(−7)	599(+5)
γN9H	535	536(+1)	746(+211)	521(−14)	542(+7)
ωNH2 + τNH2	335	329(−6)	296(−39)[c]	364(+29)	291(−44)

Water Modes

assignment	free water	complex A	complex B	complex C	complex D
νOH^{as}	3931	3891(−40)	3892(−39)	3887(−44)	3825(−106)
νOH^{s}	3809	3501(−308)	3586(−223)	3542(−267)	3742(−67)
δHOH	1603	1636(+23)	1634(+31)	1630(+27)	1623(+20)

[a] ν = stretching, δ = in-plane deformation, γ = out-of-plane deformation, τ = torsion, ω = wagging, r = rocking vibration. [b] The numbers between parentheses indicate the frequency shifts resulting from complex formation with water.

distance of 1.354 Å in free adenine as compared with the value of 1.370 Å in free guanine indicate a greater delocalization of the N10 lone pair into the aromatic ring.

Relevant vibrational frequencies in free adenine and the three water complexes are indicated in Table 4. The infrared and Raman spectra of adenine in several environments are well documented.[21,22,53−58] Our data are, in general, in good agreement with previous calculations.[47] In complex A, unlike in guanine, the νNH2 vibrations remain unchanged, and this results from the fact that the NH2 group in free adenine is already nearly sp^2 hybridized. In complexes B and C where the NH2 group is involved in complex formation, downward shifts of the νNH2 vibrations and upward shifts of the in-plane, rocking and wagging vibrations of the NH2 group are predicted. No clear frequency shifts of the torsion and wagging vibrations in

the adenine−water complexes could emerge from our calculations, these modes being strongly coupled with the water vibrations.

Comparison of earlier results on uracil,[7] thymine,[8] and cytosine[9] with the present data allows us to deduce, for the five nucleobases, a correlation between $\Delta\nu$(OH), the mean frequency shifts of the ν^{as}(OH) and ν^{s}(OH) vibrations of water, and the elongation of the OH′ bond involved in complex formation

$$\Delta\nu(\text{OH})(\text{cm}^{-1}) = 740 + 144 \ln \Delta r(\text{OH})\ (\text{Å})\ (r = 0.9874) \tag{1}$$

This correlation is rather predictable on intuitive arguments, but it is interesting to observe that the anticooperative structure D of guanine where water acts as a bidonor strongly deviates from the straight line of Figure 4.

TABLE 3: Results of B3LYP/631+G(d,p) Optimization of Free Adenine and the Three Water Complexes A, B, and C (Bond Lengths in Å, Bond Angles in Deg)

	free adenine	complex A	complex B	complex C
N6−C10	1.355	1.353	1.347	1.347
N10−H11	1.008	1.008	1.008	1.017
N10−H12	1.008	1.007	1.020	1.008
N9−H	1.010	1.019	1.010	1.010
∠C6N10H11	118.9	119.1	117.5	119.5
∠C6N10H12	120.2	120.3	121.9	119.4
∠H11N10H12	120.2	120.3	120.6	121.1
∠C5C6N10H12	−5.1	−2.9	−1.7	−1.2
∠N1C6N10H11	4.7	3.2	2	0.7
δ	0.7	0.3	0	0
Δ	0.4	0.3	0.3	0.5

	free water	complex A	complex B	complex C
O−H	0.965	0.964	0.964	0.964
O−H′	0.965	0.982	0.985	0.984
HOH′	107.2	107.4	107.1	107.1

Intermolecular Parameters

complex A		complex B		complex C	
N3···H′	1.969	N7···H′	1.886	N1···H′	1.918
H9···O$_w$	2.059	H12···O$_w$	1.945	H11···O$_w$	2.045
∠O$_w$H′N3	145.5	∠N10H12O$_w$	161.3	∠O$_w$H′N10	142.4
∠N9H···O$_w$	130.3	∠O$_w$H′N7	155.4	∠N7H′O$_w$	151.2

(A)   (B)

(C)
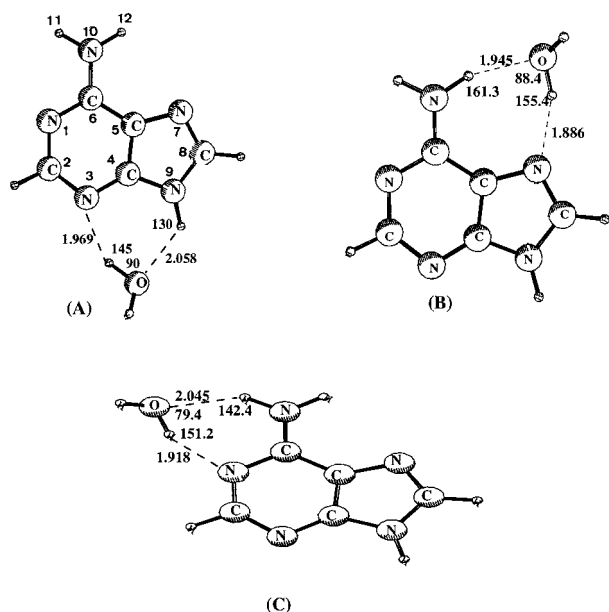
Figure 3. Optimized structures for the A, B, and C complexes between adenine and water obtained from B3LYP/6-31+G(d,p) calculations (distances in Å, angles in deg).

3. Pyramidal Character of the Amino Group in the Water Complexes. As discussed in recent work,[36] there are two structural sources of nonplanarity for the nucleobases. The nonplanarity of the first type is related to the partial sp^3 hybridization of the amino group and can be estimated as the deviation of the sum of the angles around the nitrogen atom (ΣAH) from 360° (δ in Tables 1 and 3). The nonplanarity of the second type is related to the interaction of one of the H atoms of the amino group with the closest H atom and can be estimated from the difference of the absolute values of the dihedral angles (Δ in Tables 1 and 3), provided that the NPA structure of the amino group is considered. The effects are very weak in adenine. In guanine, both effects are operating, the effect of the first type being somewhat larger for complex C where one of the NH bonds of the amino group is directly involved in complex formation and the effect of the second type being

TABLE 4: Unscaled Characteristic B3LYP/6-31+G(d,p) Vibrational Frequencies (cm^{-1}) in Free Adenine and in the Three Water Complexes A, B, and C[a]

assignment	free adenine	complex A	complex B	complex C
νNH2^{as}	3756	3755(−1)	3692(−64)	3711(−45)
νN9H	3662	3510(−152))	3662(0)	3661(−1)
νNH2^s	3615	3615(0)	3424(−191)	3458(−157)
δNH2	1665	1669(+4)	1694(+29)	1678(+13)
rNH2	1011	1011(0)	1044(+33)	1036(+25)
γR + γN9H	666	695(+29)	662(−4)	657(−9)
γN9H	575	648(+73)	576(+1)	575(0)
ωNH2 + τNH2	537	541(+4)	721	685(N10H11)
				657(N10H11)

Water Modes

assignment	free water	complex A	complex B	complex C
νOH^{as}	3931	3898(−33)	3890(−41)	3889(−42)
νOH^s	3809	3542(−267)	3487(−322)	3511(−298)
δHOH′	1603	1633(+30)	1638(+35)	1634(+31)

[a] Same remarks as below Table 2.

somewhat larger for complexes A and C. The difference between the Δ values in free guanine and complex A obtained after full geometry optimization is 9.5°,[34] somewhat larger than the value of 6° found in this work. This cannot affect the general discussion of the present work which was intended to compare the different water complexes of the nucleobases. The degree of nonplanarity, which originates from the balance between the sp^2-sp^3 hybridization of the amino group nitrogen, can be well predicted from the C−N(H2) bond length.[5,33]

Comparison with the data of the present work on free guanine, free adenine and their water complexes (Tables 1 and 3) shows that C−N(H2) distances decrease with the Σ(AH) values. This correlation is illustrated in Figure 5. The pyramidal properties of the amino group have also a marked influence on the νNH2 stretching frequencies. For the cytosine, guanine, and adenine complexes where the amino group is not involved in complex formation, the average of the ν^{as}(NH2) and ν^s(NH2) frequencies are correlated to Σ(AH) (Figure 6).

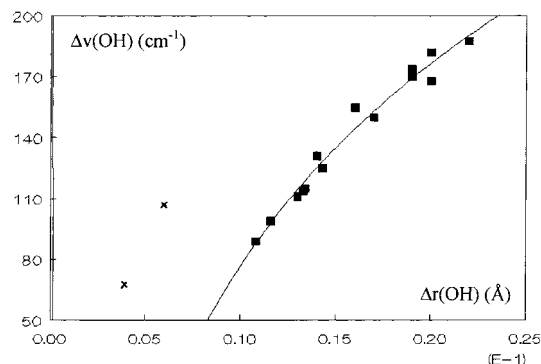Guanine and Adenine Complexes with Water

J. Phys. Chem. A, Vol. 103, No. 44, 1999 **8857**

Figure 4. $\Delta\nu(OH)(cm^{-1})$ as a function of $\Delta r(OH)(\text{Å})$ for uracil (U), thymine (T), (ref 8), cytosine (Cy) (ref 9), guanine (G), and adenine (Ad) (this work) complexes. The complexes are indicated by the corresponding letters between parentheses. The data for the guanine D complex where water acts as a bidonor are indicated by cross.
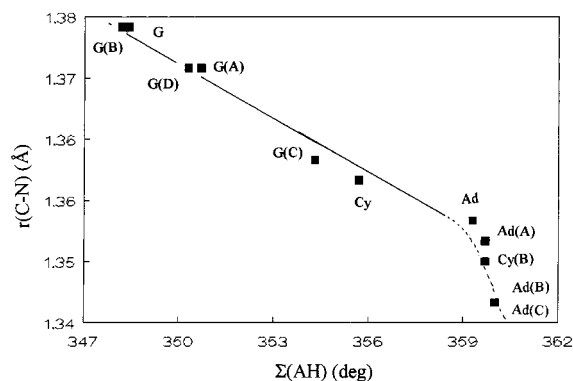

Figure 5. C-NH(2) distance (in Å) as a function of $\Sigma(AH)$ (in deg). Cy, G, and Ad refer to the free molecules. The complexes are indicated by the corresponding letters between parentheses. Cy(B) is formed between the N3 atom and the N7H8 bond of cytosine.
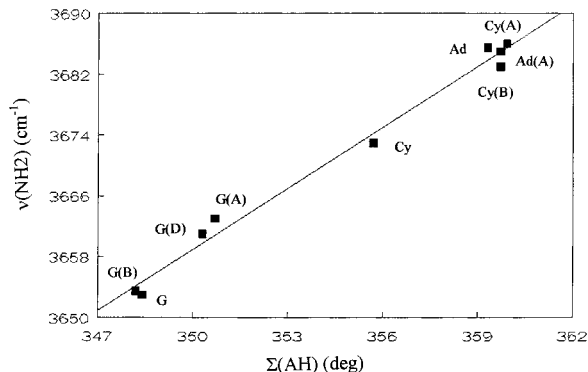

Figure 6. Mean values of the $\nu^{as}(NH2)$ and $\nu^s(NH2)$ (in cm^{-1}) as a function of $\Sigma(AH)$ (in deg) for cytosine (ref 9), guanine and adenine (this work) complexes. Cy, G, and Ad refer to the free molecules and the letters between parentheses to the corresponding water complexes.

4. Energies of the Complexes Formed between the Nucleobases and Water.

The B3LYP/6-31+G(d,p) binding energies for the different complexes of guanine and adenine including ZPE and BSSE corrections are reported in Table 5. This table also reports the energies for the uracil, thymine, and cytosine complexes calculated at the same level. These energies differ by only 0.1 to 1.4 kJ mol^{-1} from the ones computed with the 6-31++G(d,p) basis set.[7–9] The energies of the guanine—water complexes obtained after full geometry optimization of the NH2 group are the same for complexes C and D and are 0.2 kJ mol^{-1} lower for complexes A and B.

TABLE 5: B3LYP/6-31+G(d,p) Binding Energies (kJ mol^{-1}) Including ZPE and BSSE Corrections for the Different Cyclic Structures of Uracil, Thymine, Cytosine, Guanine, and Adenine Complexed with Water[a]

nucleobase	complex A	complex B	complex C	complex D
uracil	32.7[b]	24.4[c]	26.7[d]	
thymine	32.2[b]	24.6[c]	26.1[d]	
cytosine	37.2[e]	34.4[f]		
guanine	37.7	25.9	25.1	23.9
adenine	33.6	29.8	26.8	

[a] The numbers in italic refer to the cyclic complexes formed between the O atom and the vicinal NH bond of the nucleobase. [b] Complex formed between the N1H bond and the O2 atom. [c] Complex formed between the N3H bond and the O2 atom. [d] Complex formed between the N3H bond and the O4 atom. [e] Complex formed between the N1H bond and the O atom. [f] Complex formed between N3 atom and the NH7 bond.

TABLE 6: B3LYP/6-31+G(d,p) Proton Affinities (PA(B)) and Deprotonation Enthalpies (PA(A$^-$)) (kJ mol^{-1}) of Uracil, Thymine, Cytosine, Guanine, and Adenine[a]

nucleobase	PA(B)		
uracil	O4(N3 side): 849		O2(N1 side): 815.1
	O4(C5 side): 859.8		O2(N3 side): 820.1
thymine	O4(N3 side): 854.4		O2(N1 side): 830.2
	O4(C5 side): 865.7		O2(N3 side): 843.1
cytosine	N3: 955.5		O(N1 side): 921.7
			O(N3 side): 956.8
guanine	N3: 887.4	N7: 960.1	O(N1 side): 900.8
			O(N7 side): 936.7
adenine	N3: 937.6	N7: 909.6	N1: 943.8

nucleobase	PA(A$^-$)	
	NH	NH2
uracil	N1H: 1391.0	
	N3H: 1447.1	
thymine	N1H: 1398.1	
	N3H: 1450.0	
cytosine	N1H: 1444.5	N7H8: 1481.8
		N7H9: 1457.1
guanine	N1H: 1415.8	N10H11: 1412.4
	N9H: 1407.4	N10H12: 1435.0
adenine	N9H: 1409.1	N10H11: 1488.5
		N10H12: 1486

[a] Including ZPE energies computed at the same level. The proton affinity PA(B) is defined as the negative enthalpy change associated with the gas-phase protonation reaction B + H$^+$ ⇌ BH$^+$, and the deprotonation enthalpy PA(A$^-$) is defined as the enthalpy change associated with the gas-phase deprotonation reaction AH ⇌ A$^-$ + H$^+$

5. Proton Affinities and Deprotonation Enthalpies of Guanine and Adenine.

The B3LYP/6-31+G(d,p) computed proton affinities (PA(B)) and deprotonation enthalpies (PA(A$^-$)) of the different proton donor and proton acceptor sites of guanine and adenine are indicated in Table 6 together with the same parameters for uracil, thymine, and cytosine computed at the same level of theory. These values differ from 0.5 to 2 kJ mol^{-1} from the ones computed by the same DFT method but using the slightly larger 6-31++G(d,p) basis.[8] The PA(B) of the N1 atom of adenine recently computed at the HF/6-31G level 6 is 993.8 kJ mol^{-1}, being more than 50 kJ mol^{-1} higher that the one computed in this work. Our values differ by less than 10 kJ mol^{-1} from the experimental values reported for guanine (950 kJ mol^{-1}) and adenine (937.2 kJ mol^{-1})[59] which probably refer to the most basic sites, namely the N7 atom of guanine and the N1 atom of adenine. Recent higher level calculations (MP4/6-311++G(d,p)//MP2/6-31G(d)) have also shown that the amino-oxo-N9H tautomer of guanine is protonated at N7.[20] Comparison of PA(B) values of cytosine and guanine deter-

TABLE 7: Intermolecular Distances (Å) in the Closed NH···O$_w$···O=C Structures of the Complexes of Uracil, Thymine, Cytosine, and Guanine with Water

complex	$r(O_w(H'···O))$	$r(NH···O_w)$
uracil(A)[a]	1.941	1.927
uracil(B)[b]	1.975	1.988
uracil(C)[c]	1.921	1.968
thymine(A)[a]	1.929	1.941
thymine(B)[b]	1.947	1.999
thymine(C)[c]	1.914	1.979
cytosine[d]	1.813	1.947
guanine[e]	1.852	1.929

[a] Cyclic complex formed between the NlH bond and the O2 atom (ref 8). [b] Cyclic complex formed between the N3H bond and the O2 atom (ref 8). [c] Cyclic complex formed between the N3H bond and the O4 atom. [d] Cyclic complex formed between the N1H bond and the O2 atom (ref 9). [e] This work.

mined at the MP4 level of theory suggests that protonation of cytosine is easier by about 5 kJ mol^{-1}, whereas an opposite trend is found in this work. Unfortunately, the experimental results are controversial. Thus, the PA(B) value of cytosine has been reported to be 3 kJ mol^{-1} larger than that of guanine,[60,61] but another study found the PA(B) of guanine to be greater by 6 kJ mol^{-1}.[59] This suggests that the difference in PA(B) is too small to fully guarantee the reliability of either experimental or theoretical estimates. Perhaps more relevant is the order of basicity for the sites involved in the Watson−Crick structure which is as follows: cytosine (N3 and O(N3 side)) > adenine (N1) > guanine (O(N1 side)) > thymine (O4(N3 side). Such an energy order is the same as that predicted from HF/4-31G calculations,[62] but all of the PA(B) values are about 50 kJ mol^{-1} larger than the ones computed in this work.

No experimental data for deprotonation energies or enthalpies are available in the literature, and results from ion cyclotron resonance mass spectroscopy give only the relative order of acidity.[11] The reliability of our calculations can be judged from the fact that the experimental value of the deprotonation enthalpy of 2-aminopyridine is 1516 kJ mol^{-1} [63] which differs by only 1.7 kJ mol^{-1} from our computed value of 1517.7 kJ mol^{-1}. Calculations carried out at the HF/6-31G(d) level gave PA(A$^-$) values of 1495.6 and 1494.3 kJ mol^{-1} for the N1H and N10H11 sites of guanine;[5] these values are about 80 kJ mol^{-1} higher than the values computed in this work, although their difference is of the same order of magnitude.

The deprotonation enthalpies of the NH bonds in the five nucleobases computed in this work vary across a broad range, from 1391 to 1486 kJ mol^{-1}. In guanine and cytosine, the PA(A$^-$) values of the two NH bonds differ by 22.6 and 24.7 kJ mol^{-1} respectively, and this effect results from the repulsion between the NH bond and the vicinal CH bond. In adenine, where this effect is not operating, the difference between the PA(A$^-$) of the two NH bonds amounts only to 2.5 kJ mol^{-1}. Further, the PA(A$^-$) values which are the lowest in guanine seem also to depend on the pyramidalization of the amino group which is the highest in guanine and the lowest in adenine.

Some attempts have been made to discuss the acidities of the nucleobases or nucleotides, but the experimental and theoretical results disagree. Experimental results obtained from Fourier transform ion cyclotron spectroscopy have suggested that the *relative* acidities of the nucleic bases follow the order: adenine > thymine > guanine > cytosine.[11] This order does not correspond to PM3 calculations which suggest the following order: thymine > guanine > adenine > cytosine.[11] The order found in this work for the most acidic site is: uracil > thymine > guanine > adenine > cytosine. Perhaps more relevant is the

order of acidity of the NH sites involved in the Watson−Crick structure, which is as follows: guanine (N10H11 and N1H) > thymine (N3H) > cytosine (N7H8) > adenine (N10H11).

6. Correlations between the Hydrogen Bond Parameters and the Protonation and Deprotonation Enthalpies. The correlations between the hydrogen bond parameters and the proton affinity are well documented for neutral and ionic hydrogen bonds.[26,64−71] The correlations between the hydrogen bond properties and the intrinsic acidity have been much less discussed. We will consider the closed (N)H···O$_w$(H2)···O=C complexes formed between uracil, thymine, cytosine, and guanine and water. For the purpose of comparison, it is useful to note here that the six-membered structure is planar or nearly so and that the (O$_w$)H'···O and (N)H···O$_w$ angles range between 143° and 151°. The data of Tables 1, 3, 6, and 7 show that the intermolecular distances are not ordered according to the PA(B) values. In these six-membered-ring structures, the two hydrogen bonds are mutually strengthened by cooperativity. It can then be expected that the $r(O_w)H'···O$ distances will also depend, although to a lesser extent, on the acidity of the NH bonds involved in the formation of the ring structure. The best correlation coefficients are found for the following exponential equations:

$$r(O_w)H'···O = 2.532e^{-0.008\,[PA(B)\,-\,0.35(PA(A^-)]}\quad r = 0.9976$$
(2)

$$r(N)H···O_w = 1.064e^{0.000548\,[PA(A^-)\,-\,0.37PA(B)]}\quad r = 0.9683$$
(3)

These correlations recently proposed for the uracil−water and thymine−water complexes can thus be extended in a much broader PA(B)−PA(A$^-$) range. They indicate that the mutual influence of the two hydrogen bonds on each other is about the same. The cooperativities which in the present case can be evaluated from the ratio of the coefficients of PA(B) and PA(A$^-$) of eqs 2 and 3 are 0.35 and 0.37, respectively.

The hydrogen bond energies in the closed (O$_w$)H'···O complexes can also be represented as a function of the intrinsic acidities or basicities of the sites involved in hydrogen bond formation. Comparison of the data of Tables 5 and 6 leads to the following exponential expression (energies in kJ mol^{-1}):

$$E_{HB} = 5347e^{-0.00401\,[(1.5PA(A^-)\,-\,PA(B)]}\quad r = 0.9929\quad(4)$$

The data, illustrated in Figure 7, show that the correlation established for the interaction between water and uracil or thymine[8] can be now extended to a broader domain. In most of the correlations between the hydrogen bond properties and the intrinsic basicities, the coefficients of the PA of the two partners may be taken as one; for homomolecular hydrogen bonds ($\Delta PA = 0$) one usually considers the average complexation energies in a given series. Equation 4 shows some similarities with the dual substituent regression equation proposed by Caldwell et al.[72] for the alcohol−alkoxide complexes in the gas phase.

Last, we could not deduce, for closed (N)H···O$_w$H'···N structures, expressions similar to eqs 2−4. The reason lies probably in the fact that the angular properties in these structures are different, the NH···O$_w$ angle varying from 130° (adenine (A)) to 161° (adenine (B)).

Concluding Remarks

In the present work, the PAs of the different basic sites and the deprotonation enthalpies of the NH bonds of the five nucleobases are for the first time computed at an accurate level
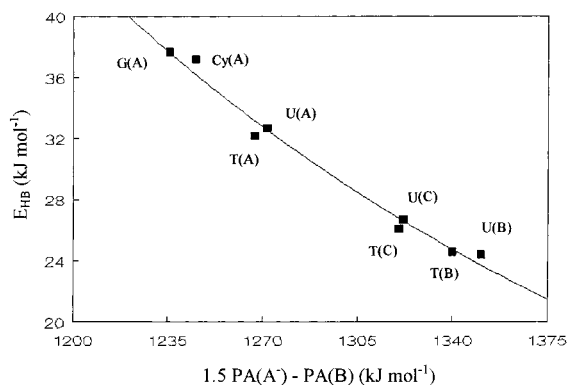
Guanine and Adenine Complexes with Water

J. Phys. Chem. A, Vol. 103, No. 44, 1999 **8859**

Figure 7. E_{HB} as a function of $1.5(PA(A^-)-PA(B))$ (both in kJ mol^{-1}) for the closed $(N)H\cdots O_wH'\cdots O$ structures of uracil, thymine, cytosine, and guanine complexed with water. U(A) and T(A) refer to the cyclic complexes formed between water and the N1H bond and the O2 atom of uracil or thymine; U(B) and T(B) refer to the cyclic complexes formed between water and the N3 bond and the O2 atom of uracil or thymine; U(C) and T(C) refer to the cyclic complexes formed between water and the N3H bond and the O4 atom of uracil and thymine (ref 7 and 8).

of theory and compared. It is shown that intrinsic acidities and basicities can be considered as key factors for understanding the energies and vibrational properties of hydrogen bonds formed between nucleobases and a water molecule. The present discussion is based solely on computational results on singly aquated nucleobases. Solvation effects are known to be vital in DNA base pair interactions.[73,74] Specific solvation of the N3 and N7 sites of guanine and adenine can modify the intrinsic acidities or basicities of the other sites. However, the solvation effects on the structure of the different uracil−adenine base pairs have been recently investigated by explicit inclusion of seven water molecules on the first coordination sphere,[6] showing that the *relative* stability order of the complexes remains unchanged upon interaction with this number of water molecules. This is also likely to be the case for the intrinsic basicities or acidities discussed in the present work.

References and Notes

(1) Watson, J. D.; Crick, F. H. *Nature* **1963**, *171*, 737.
(2) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
(3) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin, Heidelberg, New York, 1991 and references therein.
(4) Löwdin, P. O. *Rev. Mod. Phys.* **1963**, *35*, 724.
(5) Florian, J.; Leszczynski, J. *J. Am. Chem. Soc.* **1996**, *118*, 3010.
(6) Zhanpeisov, N. U.; Leszczynski, J. *J. Phys. Chem. A* **1998**, *102*, 6167.
(7) Ngyuen, M. T.; Chandra, A. K.; Zeegers-Huyskens, Th. *J. Chem. Soc., Faraday Trans.* **1998**, *94*, 1277.
(8) Chandra, A. K.; Nguyen, M. T.; Zeegers-Huyskens, Th. *J. Phys. Chem. A* **1998**, *102*, 6010.
(9) Chandra, A. K.; Ngyuen, M. T.; Zeegers-Huyskens, Th. *J. Mol. Struct. (THEOCHEM)*, in press.
(10) Del Bene, J. E. *J. Phys. Chem.* **1983**, *87*, 367.
(11) Rodgers, M. T.; Campbell, S.; Marzluff, E. M.; Beauchamp, J. L. *Int. J. Mass Spectrom. Ion Processes* **1994**, *37*, 121.
(12) Chen, E. S. D.; Chen, E. C. M.; Sane, N. *Biochem. Biophys. Commun.* **1998**, *246*, 228.
(13) Gould, I. R.; Vincent, M. A.; Hillier, I. H. *Spectrochim. Acta* **1993**, *49A*, 1927.
(14) Le Breton, P. R.; Yang, X.; Urano, S.; Fetzer, S.; Yuy, M.; Leonard, J.; Kumar, S. *J. Am. Chem. Soc.* **1990**, *112*, 2138.

(15) Sczcepaniak, K.; Szczesniak, M. *J. Mol. Struct. (THEOCHEM)* **1987**, *156*, 29.
(16) Sczcepaniak, K.; Szczesniak, M.; Person, W. B. *Chem. Phys. Lett.* **1988**, *153*, 39.
(17) Latajka, Z.; Person, W. B.; Morokuma, K. *J. Mol. Struct. (THEOCHEM)* **1986**, *135*, 253.
(18) Szczepaniak, K.; Szczesniak, M.; Szajda, W.; Person, W. B.; Leszczynski, J. *Can. J. Chem.* **1991**, *69*, 1718.
(19) Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811.
(20) Nowak, M. J.; Rostkowska, H.; Lapinski, L.; Kwiatkowski, J. S.; Leszczynski, J. *J. Phys. Chem.* **1994**, *98*, 2813.
(21) Nowak, M. J.; Rostkowska, H.; Lapinski, L.; Kwiatkowski, J. S.; Lesczynski, J. *Spectrochim. Acta* **1994**, *50A*, 1081.
(22) Gonnella, N. C.; Natanishi, H.; Holtwick, J. B.; Horowitz, D. S.; Kanamoni, K.; Leonard, N. J.; Roberts, J. D. *J. Am. Chem. Soc.* **1983**, *105*, 2050.
(23) Del Bene, J. E. *J. Mol. Struct. (THEOCHEM)* **1984**, *108*, 179.
(24) Broo, A.; Holmen, A. *J. Phys. Chem.* **1997**, *101*, 3589.
(25) Leszczynski, J. *Chem. Phys. Lett.* **1990**, *174*, 347.
(26) Riggs, N. V. *Chem. Phys. Lett.* **1991**, *177*, 447.
(27) Nowak, M. J.; Lapinski, L.; Fulara, J.; Les, A.; Adamowicz, L. *J. Phys. Chem.* **1992**, *96*, 1562.
(28) Sponer, J.; Hobza, P. *J. Am. Chem. Soc.* **1994**, *116*, 709.
(29) Sponer, J.; Hobza, P. *J. Mol. Struct. (THEOCHEM)* **1994**, *304*, 35.
(30) Sponer, J.; Hobza, P. *J. Phys. Chem.* **1994**, *98*, 3161.
(31) Bludsky, O.; Sponer, J.; Leszczynski, J.; Spirko, V.; Hobza, P. *J. Chem. Phys.* **1996**, *105*, 11042.
(32) Sponer, J.; Florian, J.; Hobza, P.; Leszczynski, J. *J. Biomol. Struct.* **1996**, *13*, 827.
(33) Leszczynski, J. *J. Phys. Chem. A* **1998**, *102*, 2357.
(34) Gorb, L.; Leszczynski, J. *J. Am. Chem. Soc.* **1998**, *120*, 5024.
(35) Smets, J.; McCarthy, M. J.; Adamowicz, L. *J. Phys. Chem.* **1996**, *100*, 14655.
(36) van Mourik, T.; Price, S. L.; Clary, D. C. *J. Phys. Chem. A* **1999**, *103*, 1611.
(37) Lee, C.; Yank, W.; Parr, R. G. *Phys. Rev.* **1988**, *37B*, 785.
(38) Becke, A. D. *Chem. Phys.* **1993**, *98*, 5648.
(39) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19,* 553.
(40) Frisch, M. F.; Trucks, G. W.; Schegel, H. B.; Gill, P. M. W.; Jonhson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Peterson, G. A.; Montgomery, J. A.; Raghavachary, K.; Al-Laham, M. A.; Zakrewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Ciolowski, J.; Stefanov, B. B.; Nanay-akkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Repogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzales, C.; Pople, J. A. *Gaussian 94*, Revision C3; Gaussian Inc.: Pittsburgh, PA, 1995.
(41) Latajka, Z.; Bouteiller, Y. *J. Chem. Phys.* **1994**, *101*, 9763.
(42) Rauhut, G.; Pulay, P. *J. Phys. Chem.* **1995**, *99*, 3093.
(43) Latajka, Z.; Bouteiller, Y.; Scheiner, S. *Chem. Phys. Lett.* **1995**, *234*, 159.
(44) Barone, V.; Adamo, C. *J. Phys. Chem.* **1995**, *99*, 15062.
(45) Kwiatowski, J. S.; Leszczynski, J. *J. Phys. Chem.* **1996**, *100*, 941.
(46) Sponer, J.; Leszczynski, J. *J. Phys. Chem.* **1996**, *100*, 1965.
(47) All the optimized B3LYP/6-31+G(d,p) geometries, vibrational frequencies and intensities in free guanine and adenine and their water complexes can be obtained from the authors.
(48) Majoube, M. *J. Chim. Phys.* **1984**, *81*, 303.
(49) Majoube, M. *J. Mol. Struct.* **1984**, *81*, 303.
(50) Letellier, R. Ghomi, M.; Taillandier, E. *Eur. Biochem. J.* **1987**, *14*, 243.
(51) Florian, J. *J. Phys. Chem.* **1993**, *97*, 10649.
(52) Florian, J.; Leszczynski, J. *Int. Quantum Chem. Biol. Symp.* **1995**, *22*, 207.
(53) Lord, R. C.; Thomas, G. J. *Spectrochim. Acta* **1967**, *23A*, 2551.
(54) Majoube, M. *J. Raman Spectrosc.* **1985**, *16*, 98.
(55) Hirokawa, A. Y.; Okada, H.; Sasagawa, S.; Tsuboi, M. *Spectrochim. Acta* **1985**, *41A*, 209.
(56) Majoube, M. *Biopolymers* **1985**, *24*, 2357.
(57) Florian, J. *J. Mol. Struct. (THEOCHEM)* **1992**, *253*, 83.
(58) Dhaouadi, Z.; Ghomi, M.; Austin, J. C.; Girling, R. B.; Hester, R. E.; Mojzes, P.; Chinsky, L.; Turpin, P. Y.; Coulombeau, C.; Jobic, H.; Tomkinson, J. *J. Phys. Chem.* **1993**, *97*, 1074.
(59) Greco, F.; Liguori, A.; Sindona, G.; Ucella, N. *J. Am. Chem. Soc.* **1990**, *112*, 9092.
(60) Meot-Ner (Mautner), M. *J. Am. Chem. Soc.* **1979**, *101*, 2396.
(61) Lias, S. G.; Liebman, J. F.; Levin, R. D. *J. Phys. Ref. Data* **1984**, *13*, 695.
(62) Del Bene, J. E. *J. Phys. Chem.* **1983**, *87*, 376.

Chandra et al.

(63) Meot-Ner (Mautner) M. *J. Am. Chem. Soc*. **1988**, *110*, 3071.

(64) Davidson, W. R.; Sunner, J.; Kebarle, P. *J. Am. Chem. Soc*. **1979**, *101*, 1675.

(65) Meot-Ner (Mautner), M.; Sieck, L.L. *J. Phys. Chem*. **1985**, *89*, 5222.

(66) Larson, J. W.; McMahon, T. B. *J. Am. Chem. Soc*. *1984*, *106*, 517.

(67) Zeegers-Huyskens, Th. *Chem. Phys. Lett*. **1986**, *129*, 172.

(68) Mo, O.; Yanez, M.; Elguero, J. *J. Org. Chem*. **1987**, *52*, 1713.

(69) Zeegers-Huyskens, Th. *J. Mol. Struct*. (THEOCHEM) **1988**, *177*, 125.

(70) Zeegers-Huyskens, Th.; Huyskens, P. In *Intermolecular Forces. An Introduction to Modern Methods and Results*; Huyskens, P. L., Luck, W. A. P., Zeegers-Huyskens, Th., Eds.; Springer-Verlag: Berlin, Heidelberg, **1991**; p24 and references therein.

(71) Zeegers-Huyskens, Th. *J. Mol. Liq*. **1995**, *67*, 33.

(72) Caldwell, G.; Rozeboom, M. D.; Kiplinger, J. P.; Bartmess, J. E. *J. Am. Chem. Soc*. **1984**, *106*, 4660.

(73) Schneider, B.; Cohen, D.; Berman, H. M. *Biopolymers* **1992**, *32*, 725.

(74) Schneider, B.; Berman, H. M. *Biophys. J*. **1995**, *69*, 2661.