Theoretical Study of the Interaction between Thymine and Water. Protonation and Deprotonation Enthalpies and Comparison with Uracil

3 AUTHORS:

Asit K. Chandra
North Eastern Hill University
134 PUBLICATIONS **2,447** CITATIONS

SEE PROFILE

Minh Tho Nguyen
University of Leuven
746 PUBLICATIONS **10,620** CITATIONS

SEE PROFILE

Thérèse Zeegers-Huyskens
University of Leuven
208 PUBLICATIONS **2,927** CITATIONS

SEE PROFILE

Theoretical Study of the Interaction between Thymine and Water. Protonation and Deprotonation Enthalpies and Comparison with Uracil

Asit K. Chandra, Minh Tho Nguyen, and Thérèse Zeegers-Huyskens*

Department of Chemistry, University of Leuven, Celestijnenlaan 200F, B-3001, Heverlee, Belgium

Received: February 24, 1998; In Final Form: May 6, 1998

The optimized geometries, harmonic vibrational frequencies, and energies of the three cyclic structures of the thymine−water complex are computed using density functional theory (B3LYP) combined with the 6-31++G(d,p) basis set. The proton affinity of the oxygen atoms and the deprotonation enthalpy of the NH bonds of thymine are computed at the same level and compared with recent data on uracil. In both uracil and thymine, the deprotonation enthalpies are lower (1391−1449 kJ mol^{-1}) than those of the biological NH donors of the peptide links (1470−1485 kJ mol^{-1}) (Mautner, M. *J. Am. Chem. Soc.* **1988**, *110*, 3075). Harmonic vibrational frequencies are also reported for the uracil−water complexes. In both uracil and thymine complexes, the most stable hydrogen bond is formed at the O site characterized by the smallest proton affinity and at the NH site characterized by the highest acidity. The intermolecular distances and the energies of the hydrogen bonds formed at the different sites of thymine and uracil depend on the proton affinity and the deprotonation enthalpy of these sites. New correlations between these parameters are presented and the cooperativity in the closed structures discussed.

Introduction

The specific interactions between the purine and pyrimidine bases are one of the cornerstones of the molecular biology. These interactions, which underly the transmission of genetic information, are governed in large part by the hydrogen bonds between the appropriate bases. For example, the Watson−Crick pairing scheme[1] involves two hydrogen bonds for the uracil (or thymine)−adenine base pair, the O8 atom of uracil being involved in a hydrogen bond with the amino group of adenine and the N3H bond of uracil interacting with the heteroaromatic N of adenine.[2,3]
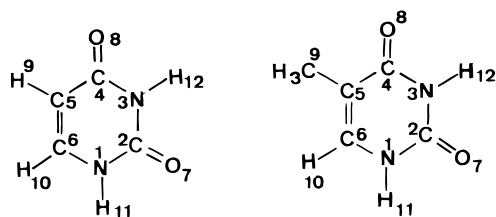
Atom labelling in uracil and thymine

It can be anticipated that the strength of the hydrogen bond between uracil and adenine will be governed not only by the proton acceptor ability of the O atom but also by the proton donor ability of the NH group. The basicity of uracil has been discussed in several works.[4−10] There are also experimental[11,12] and theoretical data on the proton affinity of thymine.[13,14] The computations have been carried out using rather low levels (HF/4-31G[13] or MINDO[14]). The correlations between the proton affinity and hydrogen bond parameters are also well documented,[15−24] and it has been recently outlined that the nature of the proton transfer to O8 and the decarboxylation process in orotidine monophosphate is of general significance for understanding the activity of many enzymes.[25]

It should be mentioned here that the *intrinsic acidity* emerges as a common property in the protein-α helix and enzymes. This property may be in part responsible for the natural selection of these biological molecules,[26,27] and this is likely to be the case for the bases of the nucleic acids. There are no experimental data on the deprotonation energies or enthalpies of these bases. In a recent work, the proton affinity of the two O atoms of uracil and the deprotonation enthalpies of the two NH bonds of uracil have been computed using the density functional theory (B3LYP) and a 6-31++G(d,p) basis set. It has been suggested that the energy of the three stable uracil−water complexes depends not only on the proton acceptor ability of the oxygen atoms but also on proton donor ability of the NH groups.[28] To establish more general correlations between the acidity and basicity of an amphoteric molecule and its hydrogen-bonding ability, the energy of the thymine−water complexes and the protonation and deprotonation energies of thymine are computed in this work, using the same level of theory. The density functional theory combined with an appropriate basis set reproduces the energy of the complex very well; in fact, the energy and the intermolecular distances obtained for the uracil−water interaction[28] are comparable with the ones calculated at the MP2/6-31++G(d,p) level.[29] For the thymine−water interaction, only the energy and geometrical parameters have been reported from rather low level (HF/STO-3G calculations).[30] For the purpose of comparison, the vibrational frequencies of both uracil and thymine−water complexes are computed in this work. It must be noticed here that the vibrational spectra of uracil and thymine in several environments (argon matrix, polycrystalline phase, aqueous solutions) have been discussed in numerous papers.[31−48] No theoretical studies have, however, addressed the assignment of the vibrational spectra for uracil or thymine complexed with *one* water molecule, and only the frequencies of the stretching modes of water complexed with uracil have been recently reported.[28] For uracil complexed with two water molecules, some characteristic vibrational modes such
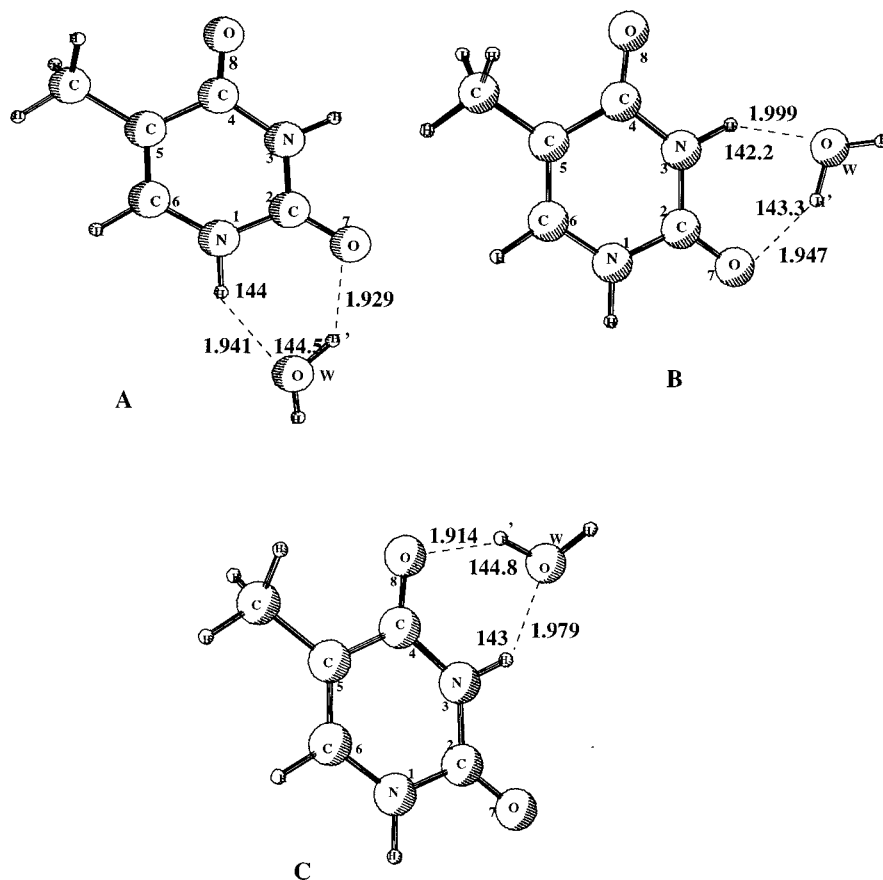
Interaction between Thymine and Water

J. Phys. Chem. A, Vol. 102, No. 29, 1998 **6011**

Figure 1. Optimized structures for the **A**, **B**, and **C** complexes between thymine and water obtained from B3LYP/6-31++G(d,p) calculations (distances in Å, angles in deg).

as the in-plane and out-of-plane vibrations of the NH and C=O groups have been recently computed.[49] Characteristic vibrations of uracil embedded in a dielectric continuum ($\epsilon = 78$) have also been reported.[49] In this case, most of the normal modes of uracil are red shifted and the results do not reflect the *selectivity* of the hydrogen bond interaction on the different sites of uracil. The same remark also holds when uracil or thymine is dissolved in aqueous solutions. It should be pointed out that the hydration of nucleotides can be described as a dynamic phenomenon and that highly organized water molecules have been observed in many nucleotide structures.[50]

The main scope of this work is to compare the optimized geometries, the characteristic vibrational modes, and the energies of the uracil− and thymine−water complexes and to discuss these parameters as a function of the basicity of the O atoms and the proton donor ablity of the NH groups. All the results are obtained by B3LYP/6-31++G(d,p) calculations. It must be stressed here that the scope of this work is neither to compare the vibrational parameters of isolated uracil and thymine obtained at different levels nor to discuss the optimal scaling factors. High-level density functional theory methods have significantly narrowed the gap between the computed and the experimental frequencies, and this has been discussed in other works.[51−55]

Computational Methods

The geometry of the isolated uracil and thymine molecules and their corresponding water complexes was fully optimized by the density functional theory using B3LYP[56,57] exchange correlation functional and 6-31++G(d,p) basis functions. The proton affinities and deprotonation enthalpies were computed

by the same procedure, using the same 6-31++G(d,p) basis function. The Gaussian package[58] was used for all the calculations. Harmonic vibrational frequencies were calculated at the same level of theory to characterize the stationary points and to evaluate the frequency shifts due to complex formation with water. Anharmonic contributions whose evaluation requires the calculation of successive energy derivatives beyond second order have been neglected.

Results and Discussion

1. Geometry and Vibrational Characteristics of the Uracil−Water and Thymine−Water Complexes. The B3LYP/6-31++G(d,p) optimized geometries of the three closed thymine−water complexes are shown in Figure 1. As for uracil−water interaction, three stable closed complexes are the ones in which one water accepts the acidic NH proton while donating a proton to the carbonyl oxygen of thymine. The three complexes are characterized by C_1 symmetry, one of the hydrogen atoms being out-of-plane of the thymine ring. Characteristic geometrical parameters of free uracil and thymine and their three water complexes are indicated in Tables 1 and 2. The CH distances in uracil and in the CH_3 group of thymine and the angles in the ring are very insensitive to complex formation and are not indicated in these tables.[59] The geometry of free uracil is very similar to that obtained recently by B3LYP calculations carried out at the 6-31G(d) level.[45] We can also notice that the greatest variation in the distances is observed for N1−C2 in complex **A**, C2−N3 for complex **B**, and N3−C4 for complex **C**, in other words for the N−C bonds involved in the formation of the pseudo-ring structure. As expected for both complexes, the greatest elongation of the C=O and NH
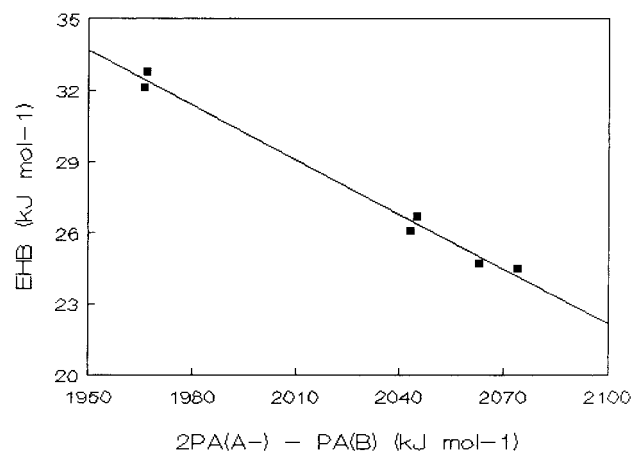
6012 *J. Phys. Chem. A, Vol. 102, No. 29, 1998*

Chandra et al.

Figure 2. E_{HB} as a function of $2PA(A^-) - PA(B)$ for the three closed structures of the uracil− and thymine−water complexes.

TABLE 1: Results of B3LYP/6-31++G(d,p) Geometry Optimization of Free Uracil and the Uracil−Water Complexes A, B, and C. Bond Lengths in Å, Bond Angles in Deg

	free uracil	complex A	complex B	complex C
N1−C2	1.3935	1.3838	1.3878	1.3968
C2−N3	1.3838	1.3780	1.3747	1.3838
N3−C4	1.4121	1.4146	1.4121	1.401
C4−C5	1.4593	1.4568	1.4615	1.4550
C5=C6	1.3519	1.3535	1.3506	1.3530
C2=O7	1.2202	1.2323	1.2316	1.2193
C4=O8	1.2227	1.2228	1.2218	1.2343
N1H11	1.0104	1.0232	1.0103	1.0106
N3H12	1.0141	1.0141	1.0256	1.0262

	free water	complex A	complex B	complex C
OH	0.9652	0.9645	0.9640	0.9640
OH′	0.9652	0.9786	0.9760	0.9782
∠HOH′	105.72	107.56	107.62	107.64

Intermolecular Parameters[a]

complex A		complex B		complex C	
H11···Ow	1.927	H12···Ow	1.988	H12···Ow	1.968
∠N1H11Ow	144.30	∠N3H12Ow	142.57	∠N3H12Ow	143
H′···O7	1.941	H′···O7	1.975	H′···O8	1.921
∠OwH′O7	142.56	∠OwH′O7	141.6	∠OwH′O8	144

[a] Ref 28.

bonds is obtained for the groups directly involved in hydrogen bond formation. A small contraction of the C2=O7 bond is calculated for complex **C** and a small contraction of the C4=O8 bond for structure **B**. In all the complexes, hydrogen bond formation results in a small contraction of the free OH bond of water. As shown in a recent work,[60] this contraction parallels a small increase of the force constant of the nonbonded OH group. A small but meaningful difference between the intermolecular parameters in the uracil and thymine complexes can be noticed. The H11···O$_w$ and H12···O$_w$ distances are shorter for the uracil complexes, but the H′···O7 and H′···O8 distances are shorter for the thymine complexes. These differences will be discussed in section 2.

Tables 3 and 4 contain the unscaled characteristic vibrational frequencies in free uracil and thymine and in their three water complexes. For isolated uracil, the frequencies are slightly different from those obtained by B3LYP/6-31G* calculations, which predict the νNH vibrations at frequencies about 10 cm^{-1} lower and the νC=O vibrations at frequencies about 30−40 cm^{-1} higher. Comparison with the experimental data in an argon matrix[38] (Table 5) reveals that for isolated uracil the

TABLE 2: Results of B3BLYP/6-31++G(d,p) Geometry Optimization of Free Thymine and the Three Water Complexes A, B, and C. Bond Lengths in Å, Bond Angles in Deg

	free thymine	complex A	complex B	complex C
N1−C2	1.3878	1.3783	1.3820	1.3911
C2−N3	1.3848	1.3785	1.3759	1.3847
N3−C4	1.4070	1.4096	1.4072	1.3968
C4−C5	1.4684	1.4657	1.4706	1.4338
C5=C6	1.3540	1.3556	1.3528	1.3552
C6−N1	1.3813	1.3775	1.3787	1.3828
C2=O7	1.2216	1.2339	1.2331	1.2205
C4=O8	1.2248	1.2248	1.2239	1.2363
N1H11	1.0102	1.0235	1.0102	1.0104
N3H12	1.0140	1.0141	1.0253	1.0257

	free water	complex A	complex B	complex C
OH	0.9652	0.9646	0.9642	0.9640
OH′	0.9652	0.9795	0.9768	0.9785
∠HOH′	105.72	107.02	107.43	107.63

Intermolecular Parameters

complex A		complex B		complex C	
H11···Ow	1.9409	H12···Ow	1.999	H12···Ow	1.979
∠N1H11Ow	144.3	∠N3H12Ow	142.22	∠N3H12Ow	142.97
H′···O7	1.929	O7···H′	1.947	H′···O8	1.914
∠OwH′O7	144.5	∠O7H′Ow	143.30	∠O8H′Ow	144.80

TABLE 3: Unscaled Characteristic B3LYP/6-31++G(d,p) Vibrational Frequencies (cm^{-1}) in Free Uracil and Water and the Three Water Complexes

assignment[a]	free uracil[b]	complex A[b]	complex B[b]	complex C[b]
νN1H	3650(109)	3426(505)	3650(154))	3647(104)
νN3H	3608(69)	3607(145)	3415(360)	3400(356)
νC2=O7	1807(642)	1784(308)	1781(306)	1809(735)
νC4=O8	1774(792)	1770(1117)	1774(1042)	1747(659)
νC5=C6	1677(62)	1677(55)	1681(53)	1673(30)
δN1H + νN1−C6	1501(96)	1534(59)	1498(80)	1511(95)
δC6H + δN3H + δN1H	1420(14)	1440(123)	1457(13)	1460(15)
δN1H + δN3H + νC2−N3 + νN3−C4	1407(101)	1415(25)	1423(130)	1419(60)
δC6H + δN3H + νring	1385(36)	1392(19)	1404(18)	1393(75)
νring + δC5H + δN1H	1233(5)	1249(39)	1236(2)	1240(1)
δC6H + δN1H + νring	1200(107)	1217(117)	1211(112)	1212(100)
νring + γC2=O7	989(7)	994(7)	996(8)	995(7)
δC6H + δC5H + δN3H	968(10)	975(6)	984(7)	985(17)
γC6H	966(1)	973(1)	964(1)	967(1)
νring	771(3)	777(6)	776(4)	777(1)
γC2=O7 + γC5H	743(48)	749(24)	742(19)	740(21)
γN3H	676(86)	675(35)	844(143)	849(145)
γN1H	562(46)	799(187)	562(64)	576(60)
δring	558(5)	572(6)	564(6)	563(2)
δring	541(6)	547(10)	546(17)	553(15)

Water Modes

assignment	free water	complex A	complex B	complex C
ν3	3927(54)	3891[c](103)	3900[c](98)	3897[c](95)
ν1	3804(5)	3611[c](314)	3654(217)	3612[c](390)
ν2	1601(87)	1613(268)	1604(232)	1607(247)
ωHOH′		633(233)	603(215)	641(248)

[a] ν = stretching, δ = in-plane deformation, γ = out-of-plane deformation, ω = wagging vibration. [b] The numbers between parentheses indicate the infrared intensities in km mol^{-1}. [c] Ref 28.

scaling factors of the frequencies are between 0.952 and 0.978 for the in-plane modes and 0.980 for the γNH out-of-plane modes.

The νCH vibrations predicted at 3262 and 3222 cm^{-1} and the δC5H mode computed at 1090 cm^{-1} in isolated uracil are not shifted in the uracil−water complex. This is also the case

Interaction between Thymine and Water

J. Phys. Chem. A, Vol. 102, No. 29, 1998 **6013**

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

assignment[a]	free thymine	complex A	complex B	complex C
νN1H	3651(106)	3419(501)	3651(92)	3649(102)
νN3H	3608(68)	3607(52)	3417(357)	3407(340)
νC2=O7	1802(784)	1776(607)	1774(707)	1806(828)
νC4=O8	1757(654)	1758(845)	1761(648)	1732(558)
νC5=C6	1701(40)	1699(38)	1704(32)	1698(5)
δN1H + δCH$_3$	1503(64)	1529(36)	1501(59)	1510(86)
δCH$_3$ + δN1H	1493(32)	1498(5)	1493(32)	1495(6)
δN1H	1423(98)	1452(134)	1433(106)	1426(72)
δN3H	1404(7)	1403(1)	1452(8)	1458(21)
δC6H + δN1H	1202(142)	1232(104)	1209(144)	1218(129)
rCH$_3$ + δN1H	1026(2)	1032(3)	1030(3)	1025(5)
γC4=O8 + γC2=O7 + γN3H	762(5)	754(89)	760(1)	763(2)
γC2=O7 + γN3H	740(48)	744(2)	741(9)	745(4)
γN3H	677(84)	677(44)	844(122)	846(121)
δring	604(1)	612(1)	610(19)	611(2)
γN1H	555(62)	808(102)	558(80)	569(78)
δring	546(7)	556(16)	550(10)	550(9)
δring	460(19)	464(36)	467(17)	467(21)

	Water Modes			
assignment	free water	complex A	complex B	complex C
ν3	3927(54)	3887(96)	3896(97)	3897(96)
ν1	3804(5)	3595(452)	3639(315)	3607(423)
ν2	1601(87)	1623(226)	1610(229)	1607(252)
wHOH'		664(300)	622(240)	648(246)

^{a} Same remarks as below Table 3, r = rocking vibration.

TABLE 5: Experimental and B3LYP/6-31++G(d,p) Frequencies for Isolated Uracil and the Water Complexes

isolated uracil		uracil−water complex		
observed[a]	calculated	observed[b]	calculated	assignment
3485	3650	3480	3650, 3647	νN1H (**B** and **C**)
		3292	3426	νN1H (**A**)
3435	3608	3431	3607	νN3H (**A**)
		3270	3415, 3400	νN3H (**B** and **C**)
1764	1807	1767	1809	νC2=O7 (**C**)
		1742	1784, 1781	νC2=O7 (**A** and **B**)
1706	1774	1712	1774, 1770	νC4=O8 (**A** and **B**)
		1698	1747	νC4=O8 (**C**)
1389	1420 or 1407	1408	1460 or 1419	δN3H (**B** and **C**)
662	676	805	844, 849	γN3H (**B** and **C**)
757	743	767	749, 742	γC2=O7 (**A** and **B**)
		759	740	γC2=O7 (**C**)
551	562	710(?)	799	γN1H (**A**)

	Water Vibrations			
free water		uracil−water		
observed	computed	observed	computed	assignment
3783	3927	3704	3891, 3897, 3900	ν3
3638	3804	3525	3654	ν1 (**B**)
		3513	3611, 3612	ν1 (**A** and **C**)
1591	1601	1614	1613, 1604, 1607	ν2 (**A**, **B**, and **C**)

^{a} Ref 38. ^{b} Ref 61.

for the νCH$_3$ modes at 3125, 3104, and 3045 cm^{-1}, for the νCH mode at 3213 cm^{-1}, for the δCH$_3$ modes at 1475 and 1430 cm^{-1}, for the δC6H mode at 1375 cm^{-1}, and for the rCH$_3$ mode at 1070 cm^{-1}, in the thymine−water complex. Between 400 and 100 cm^{-1}, several vibrational modes are computed corresponding to the out-of-plane ring deformation or to the torsion of the methyl group in thymine. In the complexes, these modes are strongly mixed with the water vibrations (intermolecular in-plane and out-of-plane vibrations, in-plane and out-of-plane

butterfly modes) and will no longer be discussed in the present work.

The frequency shifts of the νNH vibrations of the bonded NH group take values between −191 and −232 cm^{-1}. These vibrations are almost 100% pure. As a matter of fact, comparison of the data reported in Tables 1, 2, 3, and 4 shows that the frequency shifts are linearly related to the elongations of the NH bonds:

$$-\Delta\nu(NH) = -0.47 + 17.8 \times 10^3 \Delta r(NH)$$

$$r = 0.9995 \quad (1)$$

A good relation is also obtained between the frequency shifts of the γNH vibrations and the elongations of the NH bonds:

$$\Delta\gamma(NH) = -0.94 + 16.24 \times 10^3 \Delta r(NH)$$

$$r = 0.9832 \quad (2)$$

Similar correlations are obtained between the mean frequency shifts ($\Delta\nu$OH) of the ν1 and ν3 vibrations of water and the elongations of the OH bonds:

$$-\Delta\nu(OH) = -19 + 10 \times 10^3 \Delta r(OH)$$

$$r = 0.9987 \quad (3)$$

The slopes of eqs 1−3 depend on the systems. For *N*-(methoxycarbonyl)glycine complexed with water, the calculated frequency shifts are also linearly correlated to the elongation of the OH bond of the carboxylic group involved in hydrogen bond formation, but the slope is 14×10^3 in this case.[61] A linear correlation between the frequency shifts and the elongation of the OH bond of water complexed with N-bases can also be anticipated from the results of Gould and Hillier.[62] For these systems the slope is sensibly higher (19×10^3).

The νC2=O7 and νC4=O8 vibrations may be coupled together, creating in-phase and out-of-phase vibrations.[41] Recent theoretical calculations have however shown on one hand that the highest frequency mode has a predominant νC2=O7 character and is coupled with the δNH bending vibration and on the other hand that the second mode is predominantly a νC4=O8 vibration coupled with the δNH, νC−N or δCH vibrations.[42−44] The vibration predicted at 1677 cm^{-1} involves mainly a C5=C6 stretching mode slightly coupled with the νC4=O8 vibration. Our calculations also show that the νC2=O7 vibration is coupled mainly with the δN1H mode, and the C4=O8 mode is mixed with the δN3H mode. The blue shift of the νC2=O7 vibration in complexes **A** and **B** and the blue shift of the νC4=O8 mode in complexes **C** of thymine and uracil show that the mode computed at 1807 (uracil) and 1802 cm^{-1} (thymine) has a predominant νC2=O7 character, and the mode at 1774 (uracil) and 1757 cm^{-1} (thymine) a predominant νC4=O8 character. It must be noticed however that for the uracil complex **A** the contribution of the δN1H and νC2=O7 vibration to the mode computed at 1784 cm^{-1} increases, while in complex **B**, both the νC2=O7 and νC4=O8 vibrations contribute to the modes predicted at 1781 and 1774 cm^{-1}; in complex **C**, the vibrational mode predicted at 1809 cm^{-1} involves mainly a νC2=O7 stretching motion, while the mode at 1747 cm^{-1} contains more δN3H character than in the isolated molecule. In the three uracil complexes, the vibrations are also weakly coupled with the scissoring vibrations of the water molecules. The same remark also holds for thymine. In isolated thymine, the νC2=O7 vibration is mixed with the δN1H mode and the νC4=O8 vibration is coupled with the δN3H mode

and to a weaker extent with the δN1H mode. The contribution of these modes increases in the three thymine−water complexes.

The δNH vibrations contribute to several modes in the 1500−900 cm^{-1} range. In the uracil complex, the vibrations with a predominant δN1H character at 1501, 1233, and 1200 cm^{-1} are more shifted in complex **A**, and the reverse holds for the modes at 1420 and 968 cm^{-1}, which have a more pronounced δN3H character. In isolated thymine, the modes at 1475−1430 cm^{-1} are almost pure δCH$_3$ modes. The modes at 1503 and 1493 cm^{-1} are mixed δN1H and δCH$_3$ vibrations, the principal component of the δN1H vibration at 1423 cm^{-1} is red shifted by 26 cm^{-1} in complex **A**, and the main component of the δN3H vibration at 1404 cm^{-1} is red shifted by 48 and 54 cm^{-1} in complexes **B** and **C**.

No clear shifted γC=O vibration could be found. In isolated uracil and thymine, the main contributions of the γC2=O7 mode are predicted at 743 and 740 cm^{-1}, respectively. This mode is mixed with the γN3H vibration. In the complex **A** of uracil and thymine, the contribution to the γN1H vibration is increased from the isolated molecules. In the two other complexes, the vibrations predicted at 743 and 740 cm^{-1} are strongly coupled with the out-of-plane ring deformation.

Although the computed infrared intensities represent only a qualitative trend, it is worth noting that the greatest intensity increase is predicted for the νNH and γNH vibrations of the NH groups involved in hydrogen bond formation. Hydrogen bond formation on a carbonyl group generally increases the intensity of the νC=O vibration. The apparent anomalies of the computed intensities of the νC=O modes in the uracil and thymine complexes can be explained by a strong mixing of the νC2=O7 and νC4=O8 vibrations as previously discussed.

Table 5 reports some characteristic frequencies observed in an argon matrix for the uracil−water complex;[63] the frequencies in isolated uracil are also indicated for comparison. It must be pointed out here that the interpretation of the experimental infrared spectra in this low-temperature material is difficult owing to the presence of numerous bands originating from other transitions than the fundamental ones. In the infrared spectrum of isolated uracil, at least eight different absorptions assigned to Fermi resonance are observed between 1800 and 1650 cm^{-1}. As a consequence, only some bonding trends could be extracted from the experimental spectra. The existence of closed complexes is difficult to prove experimentally, but the perturbations of the vibrations in argon matrixes containing a low water content allow one to identify the interaction sites. The results of Barnes et al.[35] based on the perturbations of the two γC=O modes have suggested that water is hydrogen bonded to the C2=O7 group of uracil. Our own results[56] have suggested that water forms both C=O···HO and NH···O complexes in an argon matrix and that both C2=O7 and C4=O8 groups are involved in hydrogen bond formation. This statement is based on the observation of two shifted νN1H, νN3H, νC2=O7, and νC4=O8 vibrations. The splitting of the ν1 vibration of water also strongly suggests that the O7 and O8 atoms of uracil are involved in hydrogen bond formation. The absorptions observed at 3525 and 3513 cm^{-1} have been assigned to the ν1 vibration of water bonded to the O4 and O2 atoms of uracil. The present calculations suggest that this assignment may probably be reversed, the computed water frequencies being lower in complexes **A** than in complexes **B** or **C**.

The comparison between the present calculations and the available experimental data suggests that complexes **A** and **C** are present in an argon matrix. The existence of complex **B**

TABLE 6: B3LYP/6-31++G(d,p) Binding Energies (kJ mol^{-1}) Including BSSE Corrections for the Three Structures of the Uracil−Water and Thymine−Water Complexes[a]

	complex **A**	complex **B**	complex **C**
uracil−water	−42.5(32.8)	−33.3(24.5)	−35.8(26.7)
thymine−water	−42.4(32.1)	−33.7(24.7)	−35.2(26.1)

[a] The values between parentheses indicate the binding energies with ZPE corrections.

TABLE 7: B3LYP/6-31++G(d,p) Proton Affinities (PA(B)) and Deprotonation Enthalpies (PA(A$^-$)) (kJ mol^{-1}) of Uracil and Thymine[a]

	PA(**B**)			
	O7		O8	
	N1 side	N3 side	C5 side	N3 side
uracil[b]	815	820	860	849
thymine	830.1	835.2	866.1	855

	PA(**A**$^-$)	
	N1H	N3H
uracil[b]	1391	1447
thymine	1398	1449

[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$^+$.

cannot be ruled out in view of the small energy difference between the structures **B** and **C** (section **2**).

The scaling factors take values between 0.948 and 0.972 for the in-plane vibrations of uracil and between 0.952 and 1 for the water vibrations. The factors are about the same as in the free molecules. The scaling factor is between 1.024 and 1.033 for the γC2=O7 mode and 0.953 for the γN3H mode. The great discrepancy between the calculated (799 cm^{-1}) and the experimental (710 cm^{-1}) frequencies for the γN1H mode results probably from a erroneous assignment in ref 61.

The vibrational frequency shifts resulting from the interaction between thymine and water are similar to those observed for the uracil complexes.[63] The experimental results have suggested that more water complexes are formed on the O7 atom in thymine than in uracil.

2. Protonation and Deprotonation Enthalpies of Thymine and Bonding Trends in the Water Complexes. Table 6 reports the hydrogen bond energies in the three thymine−water complexes. The binding energies for the uracil−water complexes[28] are also indicated for comparison. The basis set superposition errors (BSSE) were computed by the counterpoise (CP) method.[64] The results indicate that for structures **A** and **C** the hydrogen bond energies are slightly higher for the uracil complexes, the reverse trend being computed for structure **B**. The differences between the hydrogen bond energies can be discussed as a function of the PA of the two O atoms and of the N-deprotonation enthalpies of the two NH bonds of uracil and thymine, which are indicated in Table 7.

These results show that the PA of the O7 atom at the two N sides and the PA of the O8 atom at the C5 and N3 side are higher for thymine than for uracil. We can also notice that the computed PA of thymine agrees well with the experimental value of 873 kJ mol^{-1},[12] which probably corresponds to the protonation of the most basic site, namely, the O8 atom at the C5 side. The results of Table 7 also indicate that the acidity of the two NH bonds of thymine is somewhat lower than the acidity

Interaction between Thymine and Water

J. Phys. Chem. A, Vol. 102, No. 29, 1998 **6015**

of the corresponding bonds in uracil. It must be noticed that the NH bonds in both molecules are characterized by a relatively high intrinsic acidity, which is sensibly higher than that of formamide (1500 kJ mol^{-1}), *N*-methylformamide (1510 kJ mol^{-1}), *N*-methylacetamide (1510 kJ mol^{-1}), and the biological NH donors of the peptide links (1470−1484 kJ mol^{-1}).[26] The deprotonation enthalpy of the N3H bond is of the same order of magnitude as the experimental value for diacetamide (1449 kJ mol^{-1}),[65] showing that the addition of a carbonyl (N-acylation) has a major effect on the acidity. The reason is probably that the N-deprotonated form of diacetamide is stabilized by delocalization of the negative charge over the two carbonyls and the bridging N^{-} center.[26]

As previously stated, the intermolecular distances are somewhat different in the uracil− and thymine−water complexes, the (N)H···O$_w$ distances being slightly shorter in the uracil complexes and the (O$_w$)H′···O distances slightly shorter in the thymine complexes. The (N)H···O$_w$ distances are mainly determined by the acidity of the NH bonds, and the (O$_w$)H′···O distances depend mainly on the basicity of the O atoms. In these closed six-membered-ring structures, the two hydrogen bonds are strengthened by cooperativity. It can be expected that the r(O$_w$)H′···O distance 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:

$$r(\text{O}_w)\text{H}'\cdots\text{O} = 2.60e^{-0.00088[\text{PA}(\textbf{B})-0.35\text{PA}(\textbf{A}^-)]}$$

$$r = 0.9870 \quad (4)$$

$$r(\text{N})\text{H}\cdots\text{O}_w = 1.08e^{0.00053[\text{PA}(\textbf{A}^-)-0.37\text{PA}(\textbf{B})]}$$

$$r = 0.9580 \quad (5)$$

These two correlations 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 obtained from the ratio of the coefficients of PA(**A**$^-$) and PA(**B**) (eq 4) or PA(**B**) and PA(**A**$^-$) (eq 5) are 0.35 and 0.37, respectively. It should also be mentioned that the slope and intercept of eqs 4 and 5 are different. This is in line with the calculations of Desmeules and Allen,[66] who have shown that the correlations between the hydrogen bond energies or intermolecular distances and the difference between the PA of the two partners depend on the nature of the atoms involved in hydrogen bond formation.

Complex formation induces a red shift in both the N1H and N3H stretching vibrations. The results reported in Tables 3 and 4 indicate that the NH shifts of both thymine and uracil complexes are higher for N1H (between −224 and −232 cm^{-1}) than for N3H (between −191 and −208 cm^{-1}), and these differences result from the larger predicted acidity of the N1H bond. Also, the red shifts of the water stretching vibrations are higher in the thymine than in the *corresponding* uracil complexes. For example, in structure **A**, the shifts of the ν3-(OH) and ν1(OH) vibrations are −36 and −193 cm^{-1} for uracil and −40 and −209 cm^{-1} for thymine. This can be accounted for by the higher basicity of the O atoms in thymine. There is, however, for the six complexes studied in the present work no general correlation between the frequency shifts induced by the interaction with water and the PA(**B**) or PA(**A**$^-$) values of the corresponding interacting sites. This clearly appears when comparing, for example, the frequency shifts of the ν1 vibration in structures **A** and **C** of uracil. The shifts are nearly identical (−193 and −192 cm^{-1}) and the PA(**B**) values of the corresponding sites are equal to 815 and 849 kJ mol^{-1}, respectively.

Comparison of the results of Tables 6 and 7 reveals that in the interaction between uracil or thymine and water *the most stable hydrogen bond is formed at the lone pair of the O7 atom (N1 side) which is characterized by the lowest PA.* The lower proton acceptor power of the O7 atom is more than compensated by the higher acidity of the N1H bond. The best correlation between the hydrogen bond energies (E_{HB}) and the PA(**B**) and PA(**A**$^-$) values of the corresponding sites (both expressed in kJ mol^{-1}) is the following one:

$$E_{\text{HB}} = 183.5-0.0768[2\text{PA}(\textbf{A}^-) - \text{PA}(\textbf{B})]$$

$$r = 0.9940 \quad (6)$$

which has a better correlation coefficient than when considering the difference PA(**A**$^-$) − PA(**B**) ($r = 0.9265$). This correlation is illustrated in Figure 2. Equation 6 shows the dominance of the proton donor in determining the hydrogen bond energies.

Somewhat similar results have been obtained for the hydrogen bond complexes involving diacetamide and guest molecules such as water, methanol, and ammonia. In this case also, the proton donor ability of the guest molecule plays a more important role than its proton acceptor ability.[67] In a broader E_{HB}−PA range, the correlation is not linear but takes a polynomial form of the second degree.

Cyclic structures for the interaction between amphoteric bases (pyridone[68] *N*-methoxycarbonylglycine[61]) and water are by no means unique. The results however cannot be compared with the present ones because the PAs of the corresponding anions have not been determined. The only complex we could compare is the cyclic formamide−water complex. The binding energy computed from DFT calculations (without ZPE corrections) is 36.6 kJ mol^{-1} [69] and is very similar to the energy of the uracil−water complex **B**. The PA of the trans lone pair of the O atom which is the hydrogen bond interaction site is 866 kJ mol^{-1}, and the deprotonation energy of the NH bond is 1504 kJ mol^{-1}.[70] In this case also, the greater basicity of the O atom is compensated by the lower acidity of the NH bond.

Concluding Remarks

The most interesting results obtained in the present theoretical study are that the most stable cyclic complexes between uracil or thymine with one water molecule are formed on the O atom having the lowest proton affinity and the NH site having the highest acidity. The intermolecular distances and binding energies are mainly determined by the proton affinity and deprotonation enthalpies of the corresponding sites. The hydrogen bond interactions that determine the specificity of recognition betweeen nucleobases must also depend on these two factors. Further, the Löwdin's mutational mechanism,[71] which involves a concerted transfer of two protons in the interbase hydrogen bonds, must also depend on the intrinsic acidity and basicity of the centers involved in the proton-transfer reaction. This will be discussed in a forthcoming paper.[72]

Note

After submission of our paper on the protonation and deprotonation enthalpies of uracil,[28] a publication of Ilich, Hemann, and Hille appeared in *J. Phys. Chem.*[73] According to in vacuo ab initio calculations (HF/DF B3LYP/6-31+G(d) method), the N1 anion is more stable by 58.5 kJ mol^{-1} than the N3 species and the 4-ol cation more stable by 48.1 kJ mol^{-1} than the 2-ol one. The influence of a medium of high dielectric constant on the protonation or deprotonation equilibria is discussed.

References and Notes

(1) Watson, J.D.; Crick, F.H. *Nature* **1953**, *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) D'Albis, A.; Wickens, M. P.; Gratzer, W. B. *Biopolymers* **1975**, *14*, 1423.

(5) Scheiner, S. *Biopolymers* **1983**, *22*, 731.

(6) Falk, M.; Hartman, K. A.; Lord, R. C. *J. Am. Chem. Soc.* **1983**, *85*, 387.

(7) Spencer, J. N.; Barton, S. W.; Smith, W. S.; Wolbach, J. F.; Powell, J. F.; Kirshenbaum, M. R.; Firth, D. W.; Harris, E. M.; Judge, T. A. *Can. J. Chem.* **1983**, *61*, 2695.

(8) Kasende, O.; Zeegers-Huyskens, Th. *J. Phys. Chem.* **1984**, *88*, 2636.

(9) Pranata, J.; Wierschke, S.G.; Jorgensen, W.J. *J. Am. Chem. Soc.* **1991**, *113*, 2810.

(10) Weisz, K.; Jänchen, J.; Limbach, H.H. *J. Am. Chem. Soc.* **1997**, *119*, 6436.

(11) Meot-Ner (Mautner), M. *J. Am. Chem. Soc.* **1979**, *101*, 2396.

(12) Greco, F.; Liguiri, A.; Ucelle, N. *J. Am. Chem. Soc.* **1990**, *101*, 2396.

(13) Del Bene, J. E. *J. Phys. Chem.* **1983**, *87*, 367.

(14) Voituyk, V.A.; Blizyuk, A.A. *J. Mol. Struct (THEOCHEM)* **1990**, *205*, 119.

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

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

(17) Meot-Ner (Mautner), M. *J. Am. Chem. Soc.* **1984**, *106*, 1257 and 1265.

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

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

(20) Mó,O.; Yañez, M.; Elguero, J. *J. Org. Chem.* **1987**, *52*, 1713.

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

(22) 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; p 24.

(23) Dega-Szafran, Z.; Hrynio, A.; Szafran, M. *J. Mol. Struct.* **1990**, *240*, 159.

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

(25) Lee, J. K.; Houk, K. N. *Science* **1997**, *276*, 942.

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

(27) Meot-Ner (Mautner), M. *J. Am. Chem. Soc.* **1988**, *110*, 3075.

(28) Ngyuen, M. T.; Chandra, A. K.; Zeegers-Huyskens, Th. *J. Chem. Soc., Faraday Trans.* **1998**, *94*, 1277.

(29) Smets, J.; McCarthy, J.; Adamowicz, L. *J. Phys. Chem.* **1996**, *100*, 14655.

(30) Del Bene, J. E. *J. Chem. Phys.* **1982**, *76*, 1058.

(31) Susi, H.; Ard, J. S. *Spectrochim. Acta* **1971**, *27A*, 1549.

(32) Nishimura, M.; Tsuboi, M.. Kato, S.; Morokuma, K. *J. Am. Chem. Soc.* **1981**, *103*, 1354.

(33) Chin, S. C.; Scott, I.; Szczepaniak, K.; Person, W. B. *J. Am. Chem. Soc.* **1984**, *106*, 3415.

(34) Harsanyi, L.; Csaszar, P. *Acta Chim. Acad. Sci. Hung.* **1983**, *113*, 257.

(35) Barnes, A. J.; Stuckey, M. A.; Le Gall, L. *Spectrochim. Acta* **1984**, *40A*, 419.

(36) Harsanyi, L.; Csaszar, P.; Csaszar, A.; Boggs, J. E. *Int. J. Quantum. Chem.* **1986**, *XXIX*, 799.

(37) Csaszar, P.; Harsanyi, L.; Boggs, J.E. *Int. J. Quantum Chem.* **1988**, *XXXIII*, 1.

(38) Graindourze, M.; Smets, J.; Zeegers-Huyskens, Th.; Maes, G. *J. Mol. Struct.* **1990**, *222*, 345.

(39) Gould, I. R.; Vincent, M. A.; Hillier, I. H. *J. Chem. Soc., Perkin Trans. 2* **1992**, 69.

(40) Person, W. B.; Szczepaniak, K.; Szczesniak, M.; Del Bene, J. E. In *Recent Experimental and Computational Advances in Molecular Spectroscopy*; Fausto, R., Ed.; Kluwer Acad. Publ.: Dordrecht, 1993; p 141.

(41) Person, W. B.; Szczepaniak, K. *Vibrational Spectra and Structure*; Durig, J. R., Ed.; Elsevier: Amsterdam, 1993; Vol. 20, p 239.

(42) Les, A.; Adamowicz, L.; Nowak, M. J.; Lapinski, L. *Spectrochim. Acta* **1992**, *48A*, 1385.

(43) Estrin, D. A.; Paglieri, L.; Corongiu, G. *J. Phys. Chem.* **1994**, *98*, 5653.

(44) Rush, T., III; Petitcolas, W. L. *J. Phys. Chem.* **1995**, *99*, 14647.

(45) Aamouche, A.; Ghomi, M.; Coulombeau, C.; Jobic, H.; Grajcar, L.; Baron, M. H.; Baumruk, V.. Turpin, P. Y.; Henriet, C.; Berthier, G. *J. Phys. Chem. A* **1996**, *100*, 5224.

(46) Aamouche, A.; Ghomi, M.. Coulombeau, C.; Grajcar, L. Baron, M.; Jobic, H.; Berthier, G. *J. Phys. Chem. A* **1997**, *101*, 1808.

(47) Aamouche, A.; Berthier, G.; Coulombeau, C.; Flament, J. P.; Ghomi, M.; Henriet, C.; Jobic, H.; Turpin, P. Y. *Chem. Phys.* **1996**, *204*, 353.

(48) Ghomi, M.; Aamouche, A.; Cadioli, B.; Berthier, G.; Grajcar, L.; Baron, M.H. *J. Mol. Struct.* **1997**, *390*, 11.

(49) Ghomi, M.; Aamouche, A.; Cadioli, B.; Berthier, G.; Grajcar, L.; Baron, M. H. *J. Mol. Struct.* **1997**, *410*, 323.

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

(51) Latajka, Z.; Bouteiller, Y. *J. Chem. Phys.* **1994**, *101*, 9793.

(52) Rauhut, G.; Pulay, P. *J. Phys. Chem.* **1995**, *99*, 3093.

(53) Latajka, Z.; Bouteiller, Y.; Scheiner, S. *Chem. Phys. Lett.* **1995**, *234*, 159.

(54) Barone, V.; Adamo, C. *J. Phys. Chem.* **1995**, *99*, 15062.

(55) Kwiatowski, J. S.; Leszczynski, J. *J. Phys. Chem.* **1996**, *100*, 941.

(56) Lee, C.; Yank, W. Parr, R. G. *Phys. Rev.* **1988**, *37B*, 785.

(57) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(58) Frisch, M. F.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Peterson, G. A.; Montgomery, J. A.; Raghavachary, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Ciolowski, J.; Stefanov, B. B.; Nanayakkara, 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.

(59) All the optimized B3LYP/6-31++G(d,p) geometries in free uracil and thymine and in their water complexes can be obtained from the authors.

(60) Nguyen, M. T.; Leroux, N.; Zeegers-Huyskens, Th. *J. Chem. Soc., Faraday Trans.* **1997**, *93*, 33.

(61) Ngyuen, M. T.; De Wael, K.; Zeegers-Huyskens, Th. *J. Phys. Chem.* **1995**, *99*, 9739.

(62) Gould, I. R.; Hillier, I. H. *J. Mol. Struct. (THEOCHEM)* **1994**, *314*, 1.

(63) Graindourze, M.; Grootaers, T.; Smets, J.; Zeegers-Huyskens, Th.; Maes, G. *J. Mol. Struct.* **1991**, *243*, 37.

(64) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(65) Cumming, J.B.; Kebarle, P. *Can. J. Chem.* **1978**, *56*, 1.

(66) Desmeules, P.J.; Allen, L.C. *J. Chem. Phys.* **1980**, *72*, 4731.

(67) Nguyen, M. T.; Leroux, N.; Zeegers-Huyskens, Th. *J. Mol. Struct.* **1997**, *404*, 75.

(68) Del Bene, J. E. *J. Phys. Chem.* **1994**, *98*, 5902.

(69) Sim, A.; St-Amant, A.; Papai, I.; Salahub, D. R. *J. Am. Chem. Soc.* **1992**, *114*, 4391.

(70) Dewar, M. J. S.; Dieter, K. M. *J. Am. Chem. Soc.* **1986**, *108*, 8075.

(71) Löwdin, P.-O. *Rev. Mod. Phys.* **1963**, *35*, 724.

(72) Chandra, A. K.; Nguyen, M. T.; Zeegers-Huyskens, M. T. *Biopolymers,* submitted.

(73) Illich, P.; Hemann, C.F.; Hille, R. *J. Phys. Chem. B* **1997**, *101*, 10923.