Computational Study of Hydrogen-Bonded Complexes between the Most Stable Tautomers of Glycine and Uracil

Iwona Dąbkowska,[†] Janusz Rak,*[,†] and Maciej Gutowski*[,†,‡]

Department of Chemistry, University of Gdańsk, Sobieskiego 18, 80−952 Gdańsk, Poland, and Environmental Molecular Sciences Laboratory, Theory, Modeling & Simulation, Pacific Northwest National Laboratory, Richland, Washington 99352

Received: April 10, 2002; In Final Form: June 4, 2002

A total of 23 hydrogen bonded complexes between the lowest energy tautomers of uracil and glycine have been characterized at the density functional level of theory with a hybrid B3LYP exchange-correlation functional and 6-31++G** basis sets. The most stable complexes are formed when the carboxylic group of glycine is bound through two hydrogen bonds with a NH proton donor and an O proton acceptor of uracil, and stabilization energies for these cyclic structures span a range of 12.3−15.6 kcal/mol. Interplay between the topological match among proton donor and acceptor sites involved in cyclic structures and their preference to form single hydrogen bonds, measured by the values of proton affinity and deprotonation enthalpy, has been discussed. Upon formation of a uracil−glycine complex, the elongations of proton donor bonds and vibrational red shifts for proton donor stretching modes can reach 0.05 Å and 650 cm^{-1}, respectively. These perturbations of proton donor bonds correlate with the magnitude of two-body interaction energy terms. A qualitative correlation was demonstrated between the values of proton affinity and deprotonation enthalpy of the sites involved in hydrogen bonds and the values of both the two-body interaction energy term and elongation of the proton donor bond.

I. Introduction

Binding of proteins to DNA plays an important role in the regulation and control of gene expression. It has recently been proved that proteins are capable of specific recognition of DNA sequences with extremely high precision (4−8 base pairs).[1] Hydrogen bonds between peptide bonds or hydrophilic side chains of amino acids and DNA bases are among the most important interactions responsible for the amazing specificity of protein binding. Therefore, basic knowledge concerning the interactions between the building blocks of proteins and DNA, amino acids and nucleic acid bases, is of great interest. Although these systems are definitely simpler than real biochemical targets, quantitative information regarding the interactions between amino acids and nucleobases can provide insight into biochemical problems. Such information can also aid in the parameterization of high quality molecular mechanics force fields, which can be applied to the modeling of complex macromolecules.

There is a paucity of theoretical information about interactions between amino acids and nucleobases. The early studies concentrated on the interaction of proteins with nucleobase pairs at the Hartree−Fock (HF) level of theory.[2,3] It was demonstrated that external hydrogen bonds stabilize or destabilize the base pairs, depending on the type of interacting residues and the site of the interaction. The interaction between single- and double-stranded B-DNA helices and polyglycine has been studied at the HF level.[4] It was found that these complexes, in which the agreement between the helical symmetry of B-DNA and polyglycine is realized, represent the most stable configurations. More recently, free energies of the interaction between the hydrophilic side chain of asparagine and nucleobase pairs have been studied by extensive conformational sampling using a molecular force field.[5] The difference in interaction specificity of asparagine toward A−T and G−C was demonstrated and the role of both structural flexibility of the side chain and entropic interactions was emphasized.

On the experimental side, much effort has been spent to analyze the role of the amide group in the untwisting of the DNA double helix.[6−8] In computational studies, the hydrophilic side chains of asparagine and glutamine are frequently replaced by acrylamide. The interaction energy between 1-methyluracil and acrylamide was studied in a combined experimental and theoretical effort.[6] The enthalpy of formation of this complex, determined in the temperature-dependent field ionization mass spectrometry experiments, was found to be 9.7 ± 1.0 kcal/mol. Furthermore, anionic complexes of a nucleobase and a model molecule have been studied using the Rydberg electron-transfer method.[1] The comparison of experimental and calculated electron binding energies for adenine bound to imidazole, pyrole, and methanol (models for serine and threonine) and cytosine bound to imidazole provided information about the structure of these neutral complexes.

Here we report on the results of electronic structure calculations concerning the simplest amino acid−nucleobase complex, i.e., the dimer of glycine and uracil. Glycine is the smallest amino acid, and uracil is a building pyrimidine nucleobase of RNA (see Figure 1). Despite its simplicity, glycine has proved to be a challenging system for experimental[9−18] and theoretical[19−27] studies. First, it can formally exist either as a zwitterionic or as a canonical tautomer. Second, there are many possible conformers for the canonical tautomer, which are

* To whom correspondence should be addressed. E-mail: maciej.gutowski@pnl.gov and janusz@raptor.chem.univ.gda.pl.
† University of Gdańsk.
‡ Pacific Northwest National Laboratory.

10.1021/jp020947i CCC: $22.00
Published on Web 07/19/2002

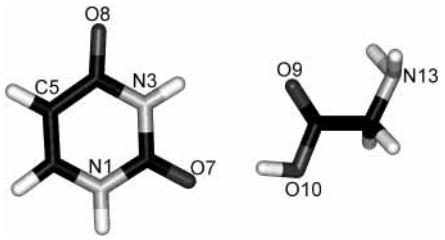

Figure 1. Lowest energy tautomers and conformers of glycine and uracil.

stabilized by different sets of intramolecular hydrogen bonds. It was demonstrated in millimeter wave spectroscopy experiments[9,10] and by measurements of substituent effects on the gas-phase basicity[12] that glycine is not a zwitterion in the gas phase. These results are consistent with ab initio calculations, which indicate that the isolated zwitterionic structure is not a minimum on the potential energy surface.[22,26] Newer studies concentrated on hydrated,[26] as well as on protonated and alkali-cationized[16−18] glycine. The strong preference of glycine to exist as a canonical structure in the gas phase has been established, which prompted us to consider only this tautomer in the current study.

The relative stability of conformers of the canonical glycine was deciphered in the course of a fruitful interplay between theoretical[19−25] and experimental[9−15] studies. The microwave studies of glycine were hampered by the coexistence of low-energy conformers with different dipole moments.[9−11,14,15] Because the microwave intensities strongly depend on the magnitude of the dipole moment of the molecule in addition to the abundance of the conformer present, Iijima et al. performed electron diffraction studies to characterize the low energy structures of glycine.[13] This study allowed for the identification of the lowest energy conformer, depicted in Figure 1, and confirmed the coexistence of a higher-energy conformer. The computational studies provided an invaluable help in solving the structure of glycine; see ref 27 for an excellent review. The most recent results obtained at the coupled-cluster level of theory with single, double, and noniterative triple excitations indicate that the lowest energy conformer (Figure 1) is separated by less than 1 kcal/mol from the next conformer in which OH acts as a proton donor and the lone pair of N as a proton acceptor.[24]

In RNA, uracil is bonded to the sugar−phosphate backbone through the nitrogen N1; see Figure 1. In DNA, it is replaced by thymine, its 5-methyl derivative. The 2,4-dioxo tautomer of uracil, depicted in Figure 1, is the most stable in the gas phase, solutions, and solids as concluded from nuclear magnetic resonance,[28] ultraviolet,[29] infrared,[30] and microwave spectroscopic[31] studies. The interest in higher energy tautomers of nucleic acid bases is dictated by a potential link between their occurrence and spontaneous point mutations developing during transformations of DNA and RNA.[32] The most systematic study of tautomers of uracil has recently been performed at the B3LYP/6-31+G(d,p) level of theory, in which 13 structures were examined.[33] The diketo tautomer was found to be stable by 11 and 13 kcal/mol with respect to the lowest monohydroxy and dihydroxy tautomer, respectively, in excellent agreement with the earlier MP4/6-31G**//HF/6-31G** predictions.[34] The significant energy gap between the diketo and other tautomers of uracil prompted us to consider only the former in the current study.

The computational studies of hydrogen bonds that develop between nucleobases have recently been reviewed,[35] and several studies on the interactions between nucleobases and water molecules have been reported.[36−41] The uracil molecule displays four characteristic regions of the neighboring proton donor and acceptor sites that are capable to form double hydrogen bonds;[39] see Figure 2a. It was recognized that the most stable uracil−water complex is formed when the O7 uracil site, characterized by the smaller proton affinity than the O8 site, and the N1H site, characterized by the highest acidity, are involved in a double hydrogen bond.[38] The involvement of a proton acceptor with the smaller proton affinity was an unexpected finding and suggested that the strength of a hydrogen bond might be more sensitive to the acidity of a proton donor than to the basicity of a proton acceptor.

The present study was inspired by the recent photoelectron spectroscopy experiments on the anion of the uracil-glycine complex.[42] The main goal of our current computational effort is to identify decisive factors responsible for the stability of neutral complexes formed by the most stable tautomers of glycine and uracil. Our preliminary results for the four most stable structures have already been reported.[43] In future reports, we will extend this study to neutral complexes formed by less stable tautomers of glycine and uracil and to anionic species.[44]

II. Computational Method

We applied primarily the DFT method with a hybrid B3LYP functional[45−47] and 6-31++G** basis sets[48,49] to study structure and stability of the uracil−glycine complexes. Our recent work on the neutral and charged arginine,[50,51] as well as other reports on complexes between nucleic acid bases and water,[52,53] and pyridine and water,[54] demonstrated the usefulness of this approach in studying systems with intra- and intermolecular hydrogen bonds. In addition to hydrogen bonds involving two highly electronegative atoms (N or O), we also explored complexes with the C5H group of uracil acting as a proton donor. The relatively weak hydrogen bonds formed by the C5H group may require an explicit treatment of intermolecular dispersion effects. For this reason, two selected complexes have also been examined at the second-order Møller−Plesset level of theory (MP2) using the same 6-31++G** basis sets. The core 1s orbitals of C, N, and O were excluded from electron correlation treatments at the MP2 level. The dependence of calculated stabilization energies for the uracil−glycine complexes on the selection of one-electron basis set was tested by performing B3LYP calculations with aug-cc-pVDZ basis sets[55] for two selected structures.

The stability of uracil−glycine (UG) complexes is expressed in terms of E_{stab}, H_{stab}, and G_{stab}. E_{stab} is defined as a difference in electronic energies of the monomers and the dimer

$$E_{stab} = E^{U}(\mathrm{Geom}^{U}) + E^{G}(\mathrm{Geom}^{G}) - E^{UG}(\mathrm{Geom}^{UG}) \quad (1)$$

with the electronic energy E^X (X = U, G, or UG) computed for the coordinates determining the optimal geometry of X (i.e., the geometry where E^X is at the minimum). E_{stab} is decomposed as[56]

$$E_{stab} = E^{U}_{dist} + E^{G}_{dist} + E^{UG}_{int} \quad (2)$$

where E^{X}_{dist} is a repulsive one-body component related to a distortion of the monomer X (X = U or G) in the dimer

$$E^{X}_{dist} = E^{X}(\mathrm{Geom}^{X}) - E^{X}(\mathrm{Geom}^{UG}) \quad (3)$$

and E^{UG}_{int} is a two-body interaction energy between the distorted monomers[57]

$$E^{UG}_{int} = E^{U}(\mathrm{Geom}^{UG}) + E^{G}(\mathrm{Geom}^{UG}) - E^{UG}(\mathrm{Geom}^{UG}) \quad (4)$$

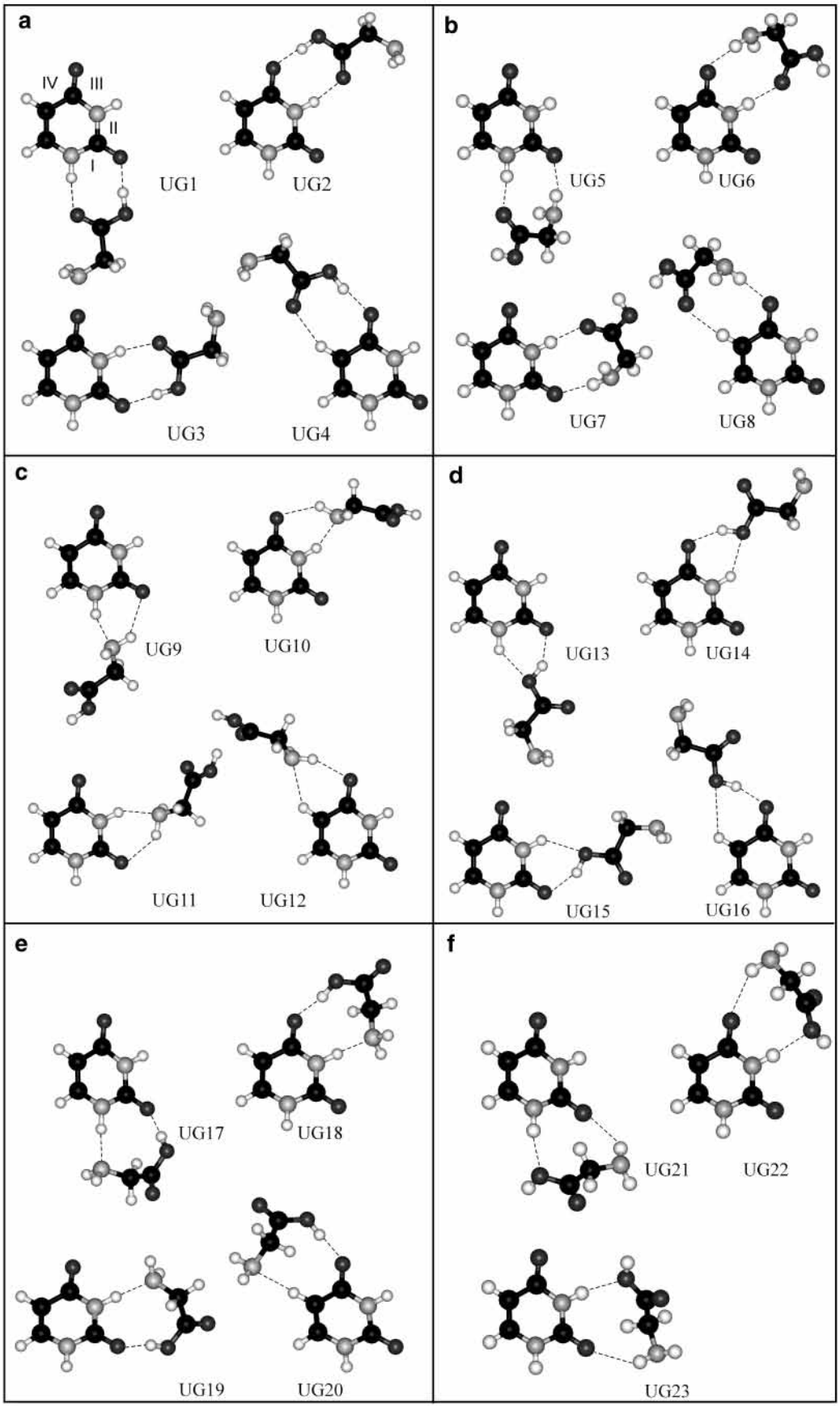

Figure 2. (a−f) B3LYP/6-31++G** optimized structures of dimers UG1-UG23. I, II, III, and IV denote regions of the uracil monomer capable of forming two adjacent hydrogen bonds.

The E_{int}^{UG} component was corrected for basis set superposition error (BSSE) using the counterpoise method of Boys and Bernardi.[58,59] In this method, the energy of each monomer is evaluated in the basis set of the dimer. The values of the E_{dist}^{X} terms, on the other hand, were calculated with monomer centered basis sets.[60,61] The stabilization enthalpy H_{stab} results

TABLE 1: Proton Affinites (PA) of the O and N Atoms and Deprotonation Enthalpies (DPE) of the NH, OH, and CH Bonds of Uracil and Glycine, Obtained at the B3LYP/6-31++G Level of Theory[a]**

		PA			
N13	212.3	**O7 (N3 side)**	196.1	**N3**	175.8
O8 (C5 side)	205.6	**O7 (N1 side)**	194.7	**N1**	170.9
O8 (N3 side)	202.7	**O9**	183.4	**O10**	175.4
		DPE			
N1H	332.8	**N3H**	346.2	**N13H**	390.4
O10H	340.2	**C5H**	378.7		

[a] All quantities in kcal/mol.

from correcting E_{stab} for zero-point vibration terms; thermal contributions to energy from vibrations, rotations, and translations; and the pV terms. Finally, the stabilization Gibbs energy G_{stab} results from supplementing H_{stab} with the entropy term. The values of H_{stab} and G_{stab} discussed in section III were obtained for $T = 298$ K and $p = 1$ atm.

All calculations were carried out with the MOPAC 2000,[62] Gaussian 98,[63] and NWChem[64] codes on a DEC Alpha 533au two-processor workstation, SGI Origin2000 numerical servers, and a cluster of 32 bit Xeon/SCI Dolphin processors.

III. Results

III.A. Properties of Isolated Monomers. The properties of isolated glycine and uracil are well reproduced at the B3LYP/6-31++G** level. The proton affinities of proton acceptor sites and deprotonation enthalpies of proton donor sites are provided in Table 1. These values will be used in section IV to analyze properties of the cyclic hydrogen bonded complexes of uracil and glycine. The proton affinity (PA) of the site Y is defined as the negative of the enthalpy change for the gas-phase reaction

$$Y(g) + H^+(g) \rightarrow YH^+(g) \tag{5}$$

whereas the deprotonation enthalpy (DPE) of the site HX is defined as the enthalpy change for the reaction

$$HX(g) \rightarrow X^-(g) + H^+(g) \tag{6}$$

and corresponds to the gas-phase acidity. Both PA and DPE are provided at $T = 298$ K.

The most basic site in the uracil−glycine complex is the N13 atom of glycine; see Figure 1. The calculated PA of 212.3 kcal/mol for this site is in excellent agreement with the experimental results of 211.8 kcal/mol for gas-phase glycine.[65] The value of PA for the O9 site of glycine is 29 kcal/mol smaller than for the N13 site. The protonation of glycine at the O10 site leads to a barrier-free detachment of water.

The most basic site of uracil is the C5 side of the O8 atom with the calculated proton affinity of 205.6 kcal/mol, which is the same as the experimental result for gas-phase uracil.[33] The value of PA at the N3 side of the O8 atom is smaller by only 2.9 kcal/mol. The O7 site is less basic than the O8 site by 9.5 kcal/mol with the N3 and N1 sides differing by only 1.4 kcal/mol.

The most acidic site in the uracil−glycine complex is the N1H group of uracil with the calculated DPE of 332.8 kcal/mol. The N3H group of uracil is less acidic than the N1H group by 13.4 kcal/mol. The C5H group, which is considered here in the context of weak hydrogen bonds formed by CH groups, is characterized by a large value of DPE of 378.7 kcal/mol. The value of DPE for the O10H group of glycine of 340.2 kcal/mol is bracketed by the DPE values for the N1H and N3H sites of uracil. This value is in good agreement with the experimental values of 341.6 ± 2.1[66] and 336.9 ± 1.4[67] kcal/mol for the DPE of glycine. The value of DPE for the N13H site of glycine is very large and amounts to 390.4 kcal/mol.

Other monomer properties are also well reproduced at the B3LYP/6-31++G** level. For glycine, the mean differences between the calculated and observed geometrical parameters[68,69] are 0.011 Å and 1.5° for bond lengths and bond angles, respectively. Similarly for uracil, the mean differences between the calculated and observed geometrical parameters[70−72] are 0.011 Å and 0.8° for bond lengths and bond angles, respectively. The calculated rotational constants of glycine and uracil differ by less than 1.1% and 0.9% from the experimental results,[14,67,31] respectively.

The unscaled harmonic frequencies of glycine and uracil are in reasonable agreement with the experimental data. For uracil, the scaling factors between the calculated and measured frequencies[72] are in a range of 0.897−0.996. The analogous range of scaling factors for glycine of 0.947−0.998[73] is even narrower. The B3LYP/6-31++G** value of the dipole moment of uracil of 4.67 D is in surprisingly poor agreement with the experimental results of 3.86[31] and 4.16 D.[74] For glycine, the agreement between the calculated (1.21 D) and experimental dipole moment of 1.15 D[75] is excellent.

III.B. Selection of Hydrogen-Bonded Structures. Both glycine and uracil belong to the class of molecules having several proton donor and acceptor centers capable of forming hydrogen bonds of various strengths. These are O7, O8, N1, N3, and C5 for uracil and N13, O9, and O10 for glycine; see Figure 1. Here, we focused on complexes with *two* intermolecular hydrogen bonds, as formation of dimers with three strong hydrogen bonds is not favored for topological reasons. The five proton donor or acceptor sites present in uracil create four regions (see Figure 2a) capable of forming two adjacent hydrogen bonds, with one site acting as a proton donor and another as a proton acceptor. On the other hand, glycine possesses one proton acceptor (O9) and two proton donor-and-acceptor centers (O10 and N13) which can be assembled in six proton acceptor−donor pairs. Six donor−acceptor pairs of glycine times four complementary acceptor−donor pairs of uracil yields 24 initial structures of the uracil−glycine complex to be explored. These 24 structures were split into six families (see Figure 2a−f), according to the nature of the glycine's proton donor and acceptor sites involved in the hydrogen bonds with uracil.

The 24 structures were characterized initially at the PM3 level of theory. The PM3 method was selected because it had been parameterized to reproduce formation energies of hydrogen bonds.[76] The minimum energy structures identified in this initial search were further refined at the B3LYP/6-31++G** level in the course of full geometry optimizations and calculations of harmonic frequencies. Finally, the global minimum structure UG1 and another structure UG4, which is the most stable among those comprising a weak hydrogen bond with the C5H site of uracil acting as a proton donor, were also optimized at the MP2 level.

All but one among 24 initial PM3 structures proved to be minima on the B3LYP potential energy surface of the uracil−glycine complex. The one PM3 structure, which did not correspond to a B3LYP energy minimum, would have been denoted UG24 in Figure 2f and contained a weak hydrogen bond between the C5H site of uracil and O10 of glycine. This structure, however, collapsed to either UG8, or UG12, or UG16 in the course of the B3LYP geometry optimization. We have

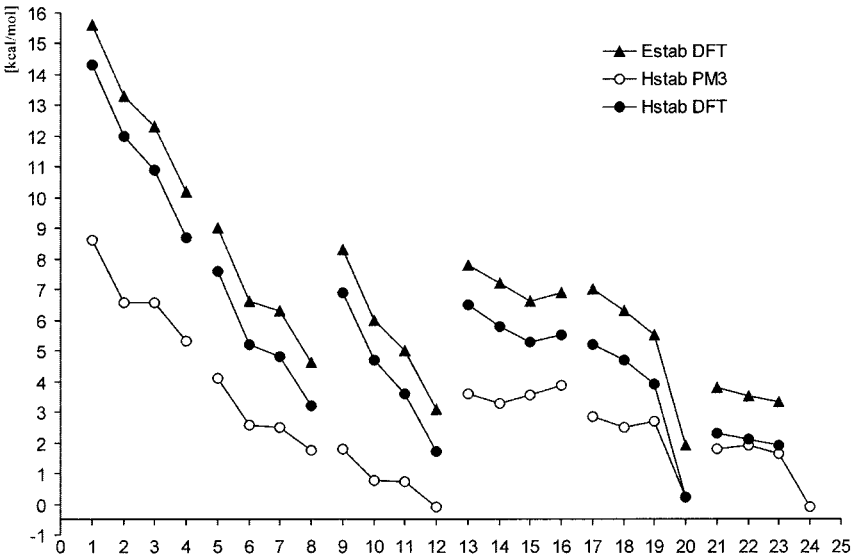

Figure 3. Energies and enthalpies of complexes obtained at the PM3 and B3LYP/6-31++G** levels of theory.

also explored structures in which uracil provides two hydrogen donor sites, N1H and C6H, whereas glycine provides two proton acceptor sites, N13 and O9. These structures, however, collapsed to UG1 in the course of the B3LYP geometry optimization.

The trends in the stabilization enthalpies determined at the B3LYP and PM3 levels for UG1-UG23 are in good agreement, see Figure 3. The PM3 values of H_{stab} are systematically smaller than the B3LYP results, but the relative ordering of structures is well reproduced. The satisfactory performance of the PM3 method is important as this method is fast enough to allow thorough scanning of potential energy surfaces for medium size systems. Moreover, recent improvements in the treatment of core−core interactions can make this method even more accurate for hydrogen bonded systems.[77]

III.C. Basis Set and Methodological Saturation. The basis set and methodological saturation tests for the dimer were performed for the UG1 and UG4 complexes, which are representative for structures bonded through two strong (UG1) and one strong and one weak (UG4) hydrogen bonds.

The basis set saturation was tested at the B3LYP level by performing additional calculations with aug-cc-pVDZ basis sets,[55] and the results are presented in square brackets in Tables 2 and 3. The values of E_{stab} obtained with the 6-31++G** and aug-cc-pVDZ sets differ by less than 0.4 kcal/mol, and the same applies to the components E_{int}^{UG}, E_{dist}^{U}, and E_{dist}^{G} (see Table 2). There is also an excellent agreement for the geometrical parameters of hydrogen bonds as the Y···H distances and Y···HX valence angles differ by less than 0.02 Å and 1.5°, respectively (see Table 3). Apparently, the B3LYP results presented in Tables 2 and 3 are basis set converged.

The methodological saturation was tested by performing MP2 calculations with the 6-31++G** basis set, and the MP2 results are presented in curly brackets in Tables 2 and 3. The discrepancies between the MP2 and B3LYP values of E_{stab} do not exceed 1.2 kcal/mol, and the geometrical parameters of hydrogen bonds differ by less than 0.03 Å for the Y···H distances and 1° for the Y···HX angles. The hydrogen bonds determined at the MP2 level are weaker (in terms of E_{stab}) and longer (in terms of the Y···H distance) than those determined at the B3LYP level.

The values of E_{int}^{UG} and E_{stab} reported in Table 2 were corrected for BSSE using the counterpoise procedure of Boys and Bernardi.[58] The values of BSSE were found to be moderate at the B3LYP/6-31++G** level as the counterpoise estimates are in a range from −0.37 to −1.42 kcal/mol. The B3LYP/aug-cc-pVDZ values of BSSE are the same to within ±0.05 kcal/mol. At the MP2/6-31++G** level, however, the values of BSSE are ca. three times larger than at the B3LYP/6-31++G** level, and the effect of the counterpoise correction is significant. Unfortunately, we could not determine the MP2

TABLE 2: Values of E_{dist}^{U}, E_{dist}^{G}, E_{int}^{UG}, E_{stab}, H_{stab}, and G_{stab} Calculated at the B3LYP/6-31++G Level[a]**

structure	E_{dist}^{U}	E_{dist}^{G}	E_{int}^{UG}	E_{stab}	H_{stab}	G_{stab}
UG1	−0.9	−1.6	18.1	15.6	14.3	3.3
	[−0.951]	[−1.64]	[18.1]	[15.5]	[14.1]	[3.0]
	(−0.7)	(−1.3)	(16.4)	(14.4)	(13.0)	(2.9)
UG2	−0.8	−1.3	15.4	13.3	12.0	1.2
UG3	−0.7	−1.0	14.0	12.3	10.9	0.3
UG4	−0.3	−0.8	11.3	10.2	8.7	−1.2
	[−0.30]	[−0.9]	[11.1]	[9.9]	[8.4]	[−1.6]
	(−0.27)	(−0.7)	(10.5)	(9.6)	(8.1)	(−0.6)
UG5	−0.3	−0.5	9.8	9.0	7.6	−2.5
UG6	−0.2	−0.2	7.0	6.6	5.2	−4.4
UG7	−0.2	−0.2	6.7	6.3	4.8	−4.7
UG8	−0.1	−0.1	4.8	4.6	3.2	−5.8
UG9	−0.6	−1.8	10.7	8.3	6.9	−2.4
UG10	−0.5	−1.6	8.2	6.0	4.7	−4.7
UG11	−0.4	−1.7	7.1	5.0	3.6	−5.7
UG12	−0.1	−1.6	4.8	3.1	1.7	−6.7
UG13	−0.2	−0.5	8.5	7.8	6.5	−2.9
UG14	−0.2	−0.5	7.9	7.2	5.8	−3.2
UG15	−0.2	−0.4	7.2	6.6	5.3	−3.6
UG16	−0.2	−0.5	7.5	6.9	5.5	−2.8
UG17	−1.2	−9.8	18.0	7.0	5.2	−6.3
UG18	−1.1	−9.1	16.5	6.3	4.7	−6.6
UG19	−1.1	−8.9	15.5	5.5	3.9	−7.3
UG20	−0.4	−9.3	11.5	1.9	0.2	−10.3
UG21	−0.1	−1.1	5.7	3.8	2.3	−7.3
UG22	−0.1	−1.0	5.2	3.5	2.1	−7.2
UG23	−0.1	−0.9	4.9	3.3	1.9	−7.2

[a] The basis set and methodological saturation was tested for selected structures only. The B3LYP/aug-cc-pVDZ results are provided in square brackets. The MP2/6-31++G** results are provided in curly brackets. All quantities in kcal/mol. Absolute values for the sum of monomers are as follows: B3LYP/6-31++G**: E_{el} = −699.304364, H = −699.124001, G = −699.197219 [au]. B3LYP/aug-cc-pVDZ: E_{el} = −699.366474, H = −699.186691, G = −699.259898 [au]. MP2/6-31++G**: E_{el} = −697.364448, H = −697.182198, G = −697.256312 [au].

TABLE 3: Selected Geometrical Characteristics of Hydrogen Bonds in the Uracil−Glycine Complexes Optimized at the B3LYP/6-311++G Level**

structure[b]	hydrogen bond type	Y···H[c] distance (Å)	Y···HX[d] valence angle (°)
UG1	O7····HO10	1.663	177.38
		[1.662]	[176.23]
		(1.694)	(176.83)
	N1H····O9	1.783	170.89
		[1.763]	[171.98]
		(1.803)	(170.83)
UG2	O8····HO10	1.666	177.22
	N3H····O9	1.836	167.90
UG3	O7····HO10	1.697	177.88
	N3H····O9	1.863	166.12
UG4	O8····HO10	1.729	171.39
		[1.727]	[169.96]
		(1.744)	(170.39)
	C5H····O9	2.236	154.41
		[2.229]	[154.82]
		(2.241)	(154.12)
UG5	N1H····O9	1.879	171.20
	O7····HN13	2.114	171.35
UG6	N3H····O9	1.934	172.33
	O8····HN13	2.114	172.79
UG7	N3H····O9	1.942	171.46
	O7····HN13	2.130	172.92
UG8	O8····HN13	2.146	177.66
	C5H····O9	2.288	164.98
UG9	N1H····N13	1.891	157.35
	O7····HN13	2.376	128.78
UG10	N3H····N13	1.934	154.63
	O8····HN13	2.305	131.18
UG11	N3H····N13	2.064	150.65
	O7····HN13	2.249	137.46
UG12	O8····HN13	2.178	155.07
	C5H····N13	1.512	136.82
UG13	O7····HO10	1.827	146.57
	N1H····O10	2.138	138.17
UG14	O8····HO10	1.808	148.47
	N3H····O10	2.198	136.36
UG15	O7····HO10	1.840	147.22
	N3H····O10	2.245	135.56
UG16	O8····HO10	1.780	160.98
	C5H····O10	2.750	122.69
UG17	O7····HO10	1.738	168.72
	N1H····N13	1.877	172.16
UG18	O8····HO10	1.772	167.40
	N3H····N13	2.462	148.38
UG19	O7····HO10	1.740	168.53
	N3H····N13	1.856	174.00
UG20	O8····HO10	1.780	167.37
	C5H····N13	2.421	161.85
UG21	N1H····O10	2.058	169.15
	O7····HN13	2.437	134.08
UG22	N3H····O10	2.096	167.53
	O8····HN13	2.473	132.79
UG23	N3H····O10	2.099	167.02
	O7····HN13	2.513	132.20

[a] The basis set and methodological saturation was tested for selected structures only. The B3LYP/aug-cc-pVDZ results are provided in square brackets. The MP2/6-31++G** results are provided in curly brackets. [b] For atomic labels, see Figure 1. [c] Y denotes proton acceptor (N or O). [d] X denotes proton donor (N, O, or C).

values of E_{stab} with more complete one-electron basis sets because of hardware limitations.

III.D. Relative Stability of the UGn (n=1−23) Complexes. The B3LYP/6-31++G** values of H_{stab} and E_{stab} for the UGn (n = 1−23) hydrogen-bonded complexes are plotted in Figure 3. The parallelism between E_{stab} and H_{stab} indicates that the contributions to H_{stab} arising from rotations and vibrations are of similar magnitude for every complex.

The values of E_{stab}, H_{stab}, and G_{stab} are collected in Table 2. The most stable complexes are UG1-UG4 with the carbonyl (O9) and hydroxyl (O10H) groups of glycine interacting with the proton donor and acceptor centers of uracil (see Figure 2a). The UG1 structure is the most stable, followed by UG2 and UG3. These three structures have two strong hydrogen bonds, and the values of E_{stab} for UG1-UG3 span a range of 15.6 to 12.3 kcal/mol, which provides ca. 8 to 6 kcal/mol per hydrogen bond. These stabilization energies are typical for dimers forming ring-like structures, such as the formic acid dimer (E_{stab}=15.2 kcal/mol[78]) or the formamide dimer (E_{stab}=14.4 kcal/mol[79]). The values of G_{stab} are positive for these three structures indicating a thermodynamic preference to form the uracil−glycine dimer. The UG4 structure, with one weak C5H···O9 hydrogen bond, is stable by only 8.7 kcal/mol in terms of E_{stab}.

The second (UG5-UG8) and third (UG9-UG12) families, in which the proton acceptor and donor sites of glycine are O9&N13H and N13&N13H, respectively (see Figure 2 parts b and c), display a similar stability. The values of H_{stab} change significantly within these families as they span a range of 7.6−1.7 kcal/mol. The fourth family (UG13-UG16), in which O10 of glycine acts as both a proton donor and acceptor (see Figure 2d), provides complexes that are only weakly bound and the values of H_{stab} span a narrow range of 6.2−5.5 kcal/mol. The fifth family (UG17-UG20), with the N13 and O10H of glycine acting as a proton acceptor and donor, respectively (see Figure 2e), is even more weakly bound with the values of H_{stab} in a range of 5.5−0.2 kcal/mol. The values of H_{stab} in the sixth family (UG21-UG23), in which the O10 and N13H of glycine act as a proton acceptor and donor, respectively (see Figure 2f), are very small and do not exceed 2.3 kcal/mol. All structures in the families two to six (UG5-UG23) are characterized by negative values of G_{stab}, and the largest value of E_{stab} of 9.0 kcal/mol is found for UG5.

The stability of structures within every family displays a striking regularity (see Table 2 and Figure 3). The most stable is always the UG($4n$ + 1) structure, n = 0−5, with N1H and O7 of uracil acting as the proton donor and acceptor, respectively. This pair of proton donor and acceptor is marked as region I in Figure 2a, and it is relevant for hydrogen bond formation by free uracil only. In RNA, however, this region is not operational as the base is covalently attached to a sugar through the N1 atom. The next most stable is the UG($4n$ + 2) structure, with N3H and O8 acting as the proton donor and acceptor, respectively (see region III in Figure 2a), followed by the UG($4n$ + 3) structures characterized by the participation of N3H and O7 of uracil in two hydrogen bonds (see region II in Figure 2a). The regions II and III are the most important for hydrogen bond formation by uracil bonded to the sugar−phosphate RNA backbone. Usually the least stable is the UG($4n$ + 4) structure (see region IV in Figure 2a), in which one hydrogen bond involves two electronegative atoms (O8 of uracil and OH or NH of glycine), but the second bond is formed between the C5H site of uracil and a proton acceptor of glycine. The second hydrogen bond is weaker, which is typical for hydrogen bonds with CH proton donors,[80] and the values of H_{stab} span a range of 8.7−0.2 kcal/mol for the UG($4n$ + 4) structures. To summarize, the strongest hydrogen bonds are formed by the region I of uracil, followed by the regions III, II, and IV. This ordering is irrespective of the nature of the proton donor and acceptor sites of glycine. The same ordering was reported for a water molecule interacting with uracil.[39]

The formation of a stable cyclic structure with two hydrogen bonds requires not only a favorable E_{int}^{UG} term but also a favorable topological match of the proton donor and acceptor

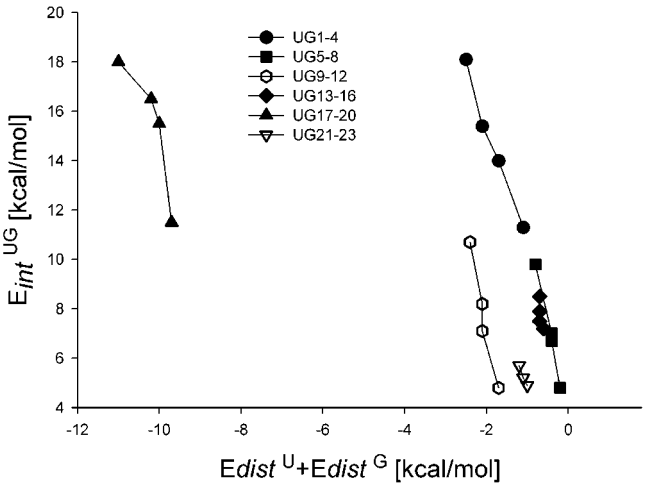

Figure 4. Interplay between $E_{\text{int}}^{\text{UG}}$ and the sum of monomer distortion terms.

sites of interacting monomers. For the uracil-glycine complex the stabilization energy E_{stab} is usually dominated by the two-body interaction energy $E_{\text{int}}^{\text{UG}}$, with the monomer distortion terms $E_{\text{dist}}^{\text{U}}$ and $E_{\text{dist}}^{\text{G}}$ playing a secondary role, see Figure 4. The only exception is the fifth family (UG17-UG20), in which glycine needs to overcome a topological mismatch with uracil and the values of $E_{\text{dist}}^{\text{G}}$ approach -10 kcal/mol. Clearly, the $E_{\text{int}}^{\text{UG}}$ term provides a driving force for overcoming the topological mismatch. Indeed, the values of $E_{\text{int}}^{\text{UG}}$ are *larger* for UG17-UG20 than for the most stable family UG1-UG4, see Figure 4. The monomer distortion terms are always more destabilizing for glycine than for uracil, and within every family, the sum of distortion terms is more destabilizing the more stabilizing is the $E_{\text{int}}^{\text{UG}}$ term, see Table 2.

III.E. Geometries and Selected Vibrational Frequencies. The geometrical features of intermolecular hydrogen bonds that are present in the UG1-UG23 structures are summarized in Table 3. The strength of a hydrogen bond is determined by the (i) charge distribution in the proton donor (XH) and acceptor (Y) fragments, (ii) the distance between H and Y, (iii) the Y···HX angle, and (iv) the monomer distortion terms E_{dist}^{X} that quantify strains acquired by the monomers when a dimer is formed. As demonstrated by the data gathered in Table 3 and Figure 2, every structure is stabilized by two hydrogen bonds that differ in length and angle.

The hydrogen bonds are the shortest and the most linear in the most stable family UG1-UG4. The favorable geometries are reflected by large values of $E_{\text{int}}^{\text{UG}}$, accompanied by nonexuberant monomer distortion terms. All four complexes in this family have a symmetry plane, i.e., the minimum energy structures are of C_s symmetry. In the second most stable family UG5-UG8, the hydrogen bonds are still quite linear and only slightly longer than in the first family. As the monomer distortion terms are comparable for the two families, a significantly lower stability of the second rather than the first family must result from different proton acceptor and donor sites involved in hydrogen bonds. These bonds remain quite linear in the UG17-UG20 family, which is characterized by the strongly attractive $E_{\text{int}}^{\text{UG}}$ terms and the strongly repulsive $E_{\text{dist}}^{\text{G}}$ terms. Finally, the bonds are quite nonlinear and sometimes quite long in the third, fourth, and sixth family (UG9-UG16 and UG21-UG23). These families are characterized by small absolute values of both $E_{\text{int}}^{\text{UG}}$ and $E_{\text{dist}}^{\text{G}}$. There is a symmetry plane for every complex in the fourth family (UG13-UG16), but the hydrogen bonds with

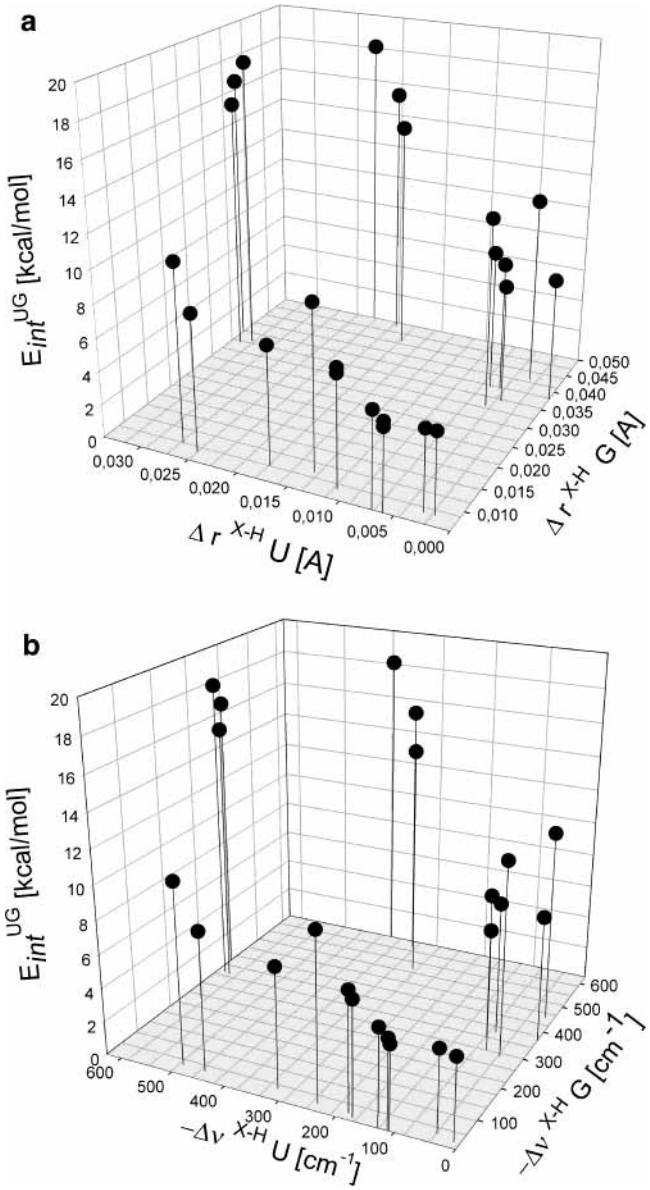

Figure 5. Interaction energy $E_{\text{int}}^{\text{UG}}$ as a function of elongation of the proton donor X−H bonds (a) and vibrational red shifts for the stretching X−H modes (b).

the hydroxyl group of glycine acting as both a proton donor and acceptor are apparently quite weak.

Formation of a hydrogen bond Y···HX is usually accompanied by an elongation of the H−X bond, $\Delta r^{\text{H-X}}$, and a red shift of the frequency for the H−X stretching mode, $\Delta\nu^{\text{H-X}}$. The tabulated values of $\Delta r^{\text{H-X}}$ and $\Delta\nu^{\text{H-X}}$ in UG1-UG23 are provided in the Supporting Information. The values of $\Delta r^{\text{H-X}}$ reach at most 0.050 Å for the OH bond of glycine and 0.034 Å for the NH bond in uracil. The values of $\Delta\nu^{\text{H-X}}$ attain at the maximum -650 cm^{-1} for the OH bond of glycine and -629 cm^{-1} for the NH bond in uracil. These perturbations of covalent H−X bonds correlate with the magnitude of the $E_{\text{int}}^{\text{UG}}$ term, see Figure 5 parts a and b, and provide insight as to the strength of individual hydrogen bonds. Both 3-D plots from Figure 5 show that the largest values of $E_{\text{int}}^{\text{UG}}$ of ca. 18 kcal/mol occur only when *both* hydrogen bonds involve significantly perturbed proton donors. Two three-point clusters in Figure 5 parts a and b with $E_{\text{int}}^{\text{UG}}$ above 14 kcal/mol are related to the first three structures of the families one and five. The small values of

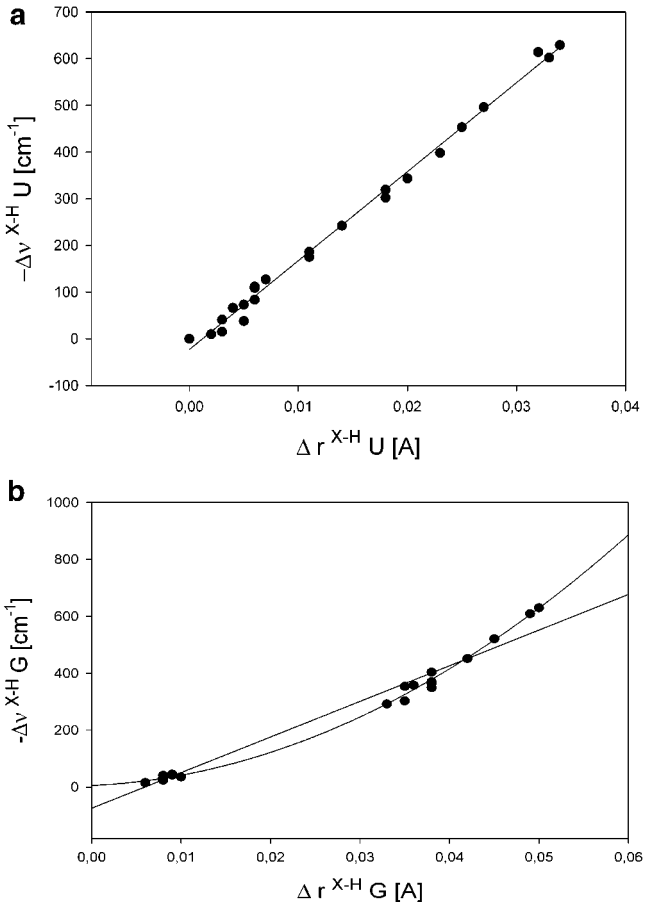

Figure 6. Correlation between elongations of the proton donor X−H bonds and vibrational red shifts for the X−H stretching modes of uracil (a) and glycine (b).

E_{int}^{UG} for other structures are reflected by small perturbations of the H−X bonds of either glycine, or uracil, or both of them.

The striking similarity between parts a and b of Figure 5 suggests a strong correlation between the values of Δr^{H-X} and $\Delta\nu^{H-X}$. This correlation is explicitly displayed in Figure 6 parts a and b, for uracil and glycine, respectively. The plot $\Delta\nu^{H-X}$ versus Δr^{H-X} is linear for uracil with a square correlation coeffcient r^2 of 0.995. For glycine, however, a departure from linearity is observed with the parabolic and linear fits providing 0.996, and 0.968 for r^2, respectively. This is in contrast with typical findings from the literature, which report linear relationships between $\Delta\nu^{H-X}$ and Δr^{H-X}.[38,81]

IV. Discussion

Hydrogen bonding can be studied in detail for small systems using highly correlated electronic structure methods and symmetry-adapted perturbation theory of intermolecular forces.[82] There is also a real need for qualitative interpretations that can be used for large systems or for a large variety of systems.[83] The most common qualitative approach is to relate the strength of a hydrogen bond to the values of PA and DPE for the proton acceptor and donor, respectively. It should be kept in mind, however, that the values of PA and DPE can provide insight into the values of E_{int}, Δr^{H-X}, $\Delta\nu^{H-X}$, etc., but they are not sufficient to predict a topological match (or mismatch!) between the proton donor and acceptor sites, and therefore, they give no hint about the magnitude of E_{dist} terms. These terms, however, may be important for topologically poorly matched sites, see UG17-UG20, and may seriously weaken the resulting hydrogen bonds.

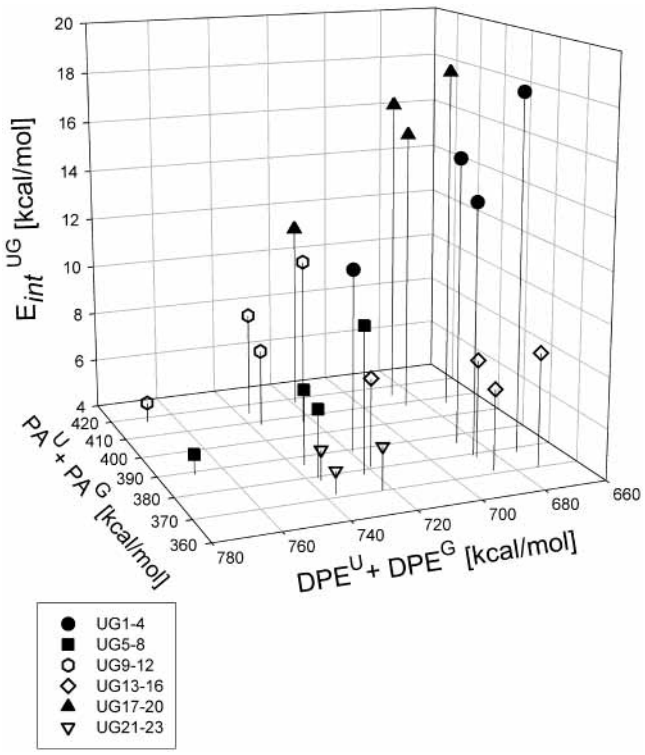

Figure 7. Dependence of E_{int}^{UG} on $PA^G + PA^U$ and $DPE^G + DPE^U$ for the 23 uracil−glycine structures from Figure 2.

Because every uracil−glycine complex in Figure 2 possesses two hydrogen bonds, we will use the values of PA and DPE for the appropriate sites, see Table 1, to define two variables. The first set of variables is given by $PA^G + PA^U$ and $DPE^G + DPE^U$ and a plot of E_{int}^{UG} versus these two variables is presented in Figure 7. As expected, large values of E_{int}^{UG} are observed for small values of $DPE^G + DPE^U$ and large values $PA^G + PA^U$. The values of E_{int}^{UG} systematically increase as the values of $DPE^G + DPE^U$ decrease. On the other hand, the dependence of E_{int}^{UG} on $PA^G + PA^U$ is less systematic. This may suggest that the values of DPE are indeed more important for the strength of a hydrogen bond than the values of PA (see also ref 38). For every family, the four participating structures form a "3+1" pattern in the $DPE^G + DPE^U$ and $PA^G + PA^U$ plane. The first three points are close to each other and correspond to the structures with three conventional hydrogen bonds. The fourth point always corresponds to the UG($4n + 4$) structure with a weak C5H···Y hydrogen bond and the large value of DPE^U.

The second set of variables is given by $DPE^G - PA^U$ and $DPE^U - PA^G$, which are suitable parameters to characterize propensity of the proton donor and acceptor pairs of uracil and glycine to form hydrogen bonds, and a plot of E_{int}^{UG} versus these two variables is presented in Figure 8. The large values of E_{int}^{UG} occur for small values of both $DPE^G - PA^U$ and $DPE^U - PA^G$. The largest values of E_{int}^{UG} are reported for the family UG17-UG20, for which the strongest basic (N13) and acidic (O10H) sites of glycine are involved in hydrogen bonds with uracil. This family is not dominant in terms of E_{stab} due to a topological mismatch between uracil and glycine and a destabilizing E_{dist}^{G} term.

Another characteristic feature of Figure 8 is that the 23 structures are arranged in two rows parallel to the $DPE^U - PA^G$ axis. These rows are determined by two accessible ranges of $DPE^G - PA^U$ at 135−145 and 185−196 kcal/mol. The two ranges are related to two proton donating sites of glycine, i.e.,

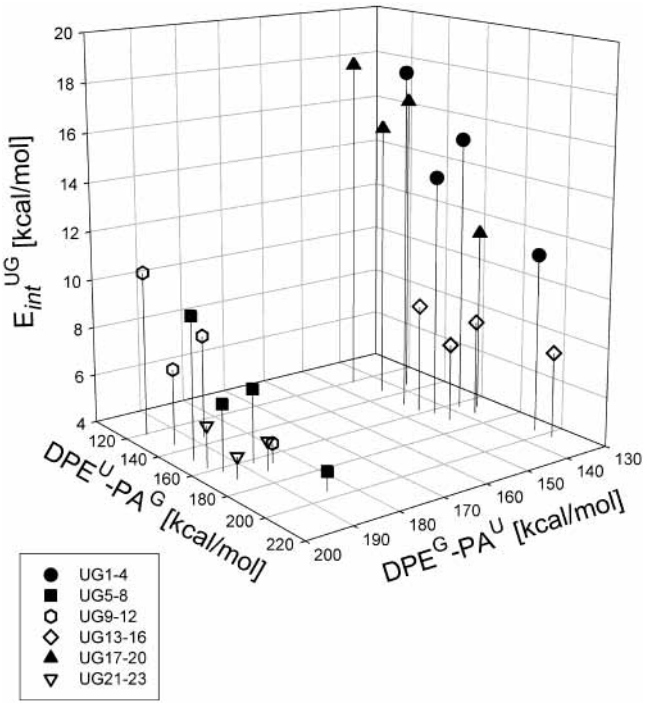

Figure 8. Dependence of E_{int}^{UG} on $DPE^G - PA^U$ and $DPE^U - PA^G$ for the 23 uracil−glycine structures from Figure 2.

O10H (first range) and N13H (second range). Clearly, the values of PA^U are responsible for only minor variations among the values of $DPE^G - PA^U$ for a fixed DPE^G. Then, the structures from one family belong to the same range of $DPE^G - PA^U$, and the related values of E_{int}^{UG} increase with a decreasing value of $DPE^U - PA^G$. This relation applies to every family presented in Figure 2. Finally, the four participating structures for every family form again a "3 + 1" pattern in the $DPE^G - PA^U$ and $DPE^U - PA^G$ plane with the single feature being again the UG-(4n+4) structure.

In view of the parallelism between Δr^{H-X} and $\Delta\nu^{H-X}$, see Figures 5 and 6, we will only analyze the dependence of the former on the values of PA and DPE. The 3-D plots of Δr^{H-X} versus $DPE^U - PA^G$ and $DPE^G - PA^U$ are displayed in Figure 9 parts a and b for uracil and glycine, respectively. The significant H−X bond elongations for uracil are observed primarily for small values of $DPE^U - PA^G$. The values of $DPE^G - PA^U$ play a secondary role, see Figure 9a, but their small values further enhance uracil's Δr^{H-X}s. This suggests a positive cooperativity for cyclic hydrogen bonds in the uracil−glycine complexes. Similarly, the H−X bond elongations for glycine are primarily controlled by the values of $DPE^G - PA^U$, see Figure 9b, but the role of the second hydrogen bond, related here to $DPE^U - PA^G$, is less transparent than is the case uracil's Δr^{H-X}'s.

For uracil's Δr^{H-X}'s, a correlation with $DPE^U - PA^G$ is observed within each family, see Figure 9a, because every family corresponds to a fixed proton donor and acceptor site of glycine and variable binding sites of uracil. For glycine's Δr^{H-X}'s, on the other hand, the structures are grouped according to k of UG($4n + k$), because every such a group corresponds to a fixed proton donor and acceptor site of uracil and variable binding sites of glycine.

Finally, we attempted to establish a quantitative relation between $DPE^G - PA^U \equiv x_1$, $DPE^U - PA^G \equiv x_2$ and E_{int}^{UG}. The second-order polynomial expansion of E_{int}^{UG} in terms of x_1 and x_2 led to a small value of 0.772 for r^2, when all 23 UG structures

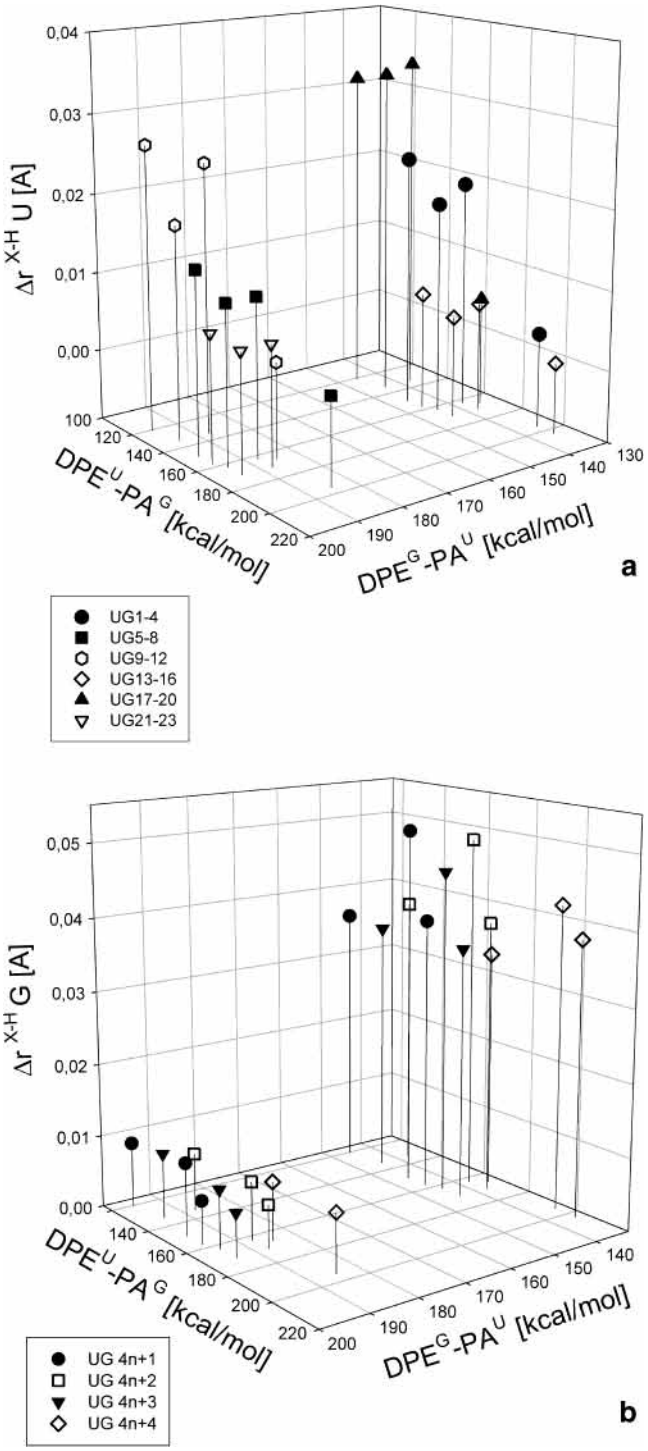

Figure 9. Dependence of Δr^{H-X} on $DPE^U - PA^G$ and $DPE^G - PA^U$ for the uracil (a) and glycine (b) H−X bonds.

were used in the fitting procedure. This failure to provide a quantitative prediction should not be discouraging. Hydrogen bonding is a complex phenomenon that cannot be fully described by the monomer properties restricted to DPE and PA. For example, it would be difficult to correlate the valence repulsion or dispersion contributions to E_{int}^{UG} with DPEs and PAs alone. Moreover, the monomer properties determined for isolated species may be different than those for the monomers deformed upon formation of hydrogen-bonded clusters. Finding robust but simple quantitative relations between monomer properties and stability of hydrogen bonds is a challenging problem for theorists interested in hydrogen bonding effects in large molecular systems.

V. Summary

We demonstrated that the most stable complexes between uracil and glycine are formed when the carboxylic group of glycine is bound through two hydrogen bonds to uracil. The largest stabilization energy of 15.6 kcal/mol was determined at the B3LYP/6-31++G** level for the UG1 structure. This structure, however, involves the N1 atom of uracil, which in RNA is covalently bonded to the sugar−phosphate backbone. Two other structures, UG2 and UG3, which involve the carboxylic group of glycine and uracil's sites available under the RNA structural constraints, are bound by 13.3 and 12.3 kcal/mol, respectively. Very similar stabilization energies are obtained at the MP2/6-31++G** level of theory. The enthalpies are smaller than energies of stabilization by ca. 1.3 kcal/mol. The free energies of stabilization favor formation of uracil−glycine complexes for UG1, UG2, and UG3 only.

The 24 UG structures analyzed in this study were grouped into families according to the nature of glycine's proton donating and accepting sites. Within every family, the stability of cyclic dimers clearly correlates with the nature of uracil's proton donating and accepting sites. The strongest binding is provided by the N1H proton donor and the O7 proton acceptor, and the weakest binding is provided by the C5H proton donor and the O8 proton acceptor.

The formation of a stable cyclic structure with two hydrogen bonds requires not only a favorable two-body interaction but also a favorable topological match of the proton donor and acceptor sites to minimize the monomer distortion terms. A family UG17-UG20 has been identified that might have excellent two-body interactions but which fails to meet the topological match requirement and proves to be only marginally stable in terms of energy and unstable in terms of free energy.

Upon formation of a uracil−glycine complex, the elongations of proton donor bonds and vibrational red shifts for proton donor stretching modes can reach 0.05 Å and 650 cm^{-1}, respectively. These perturbations of H−X bonds correlate with the magnitude of the two-body interaction energy terms and provide insight into the strength of individual hydrogen bonds. The elongations of proton donor bonds and their vibrational red shifts are well correlated with a linear relation for uracil and a parabolic relation for glycine.

The two-body interaction energy terms were correlated with the values of proton affinity and deprotonation enthalpy of the sites involved in hydrogen bonds. The quality of numerical fits was low, indicating that the diversity of proton donating and accepting sites in uracil and glycine is too large to make a robust second-order polynomial expansion of the interaction energy in terms of PA and DPE.

Scanning the topological space of hydrogen-bonded clusters of complex molecules is a computationally demanding task. The PM3 method was found very useful in identifying low energy structures for uracil−glycine complexes. The work on complexes formed by higher energy tautomers of glycine and uracil and on anionic complexes is in progress.

Acknowledgment. This work was supported by the Polish State Committee for Scientific Research (KBN) Grant No. BW/8000-5-0296-2 to J.R. M.G. was supported by the U.S. DOE, Office of Basic Energy Sciences, Chemical Sciences Division. This research was performed in part at the William R. Wiley Environmental Molecular Sciences Laboratory (EMSL) at the Pacific Northwest National Laboratory (PNNL). Operation of the EMSL is funded by the Office of Biological and Environmental Research in the U.S. Department of Energy (DOE). PNNL is operated by Battelle for the U.S. DOE under Contract DE-AC06-76RLO 1830. The computer time allocation provided by the Academic Computer Center in Gdańsk (TASK) is gratefully acknowledged. This research also used resources of the National Energy Research Scientific Computing Center (NERSC), which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Supporting Information Available: Two tables: (i) elongations of valence X−H bonds involved in hydrogen bonds and (ii) frequency shifts for stretching X−H modes. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Desfrancois, C.; Carles, S.; Schermann, J. P. *Chem. Rev.* **2000**, *100*, 3943.
(2) Sarai, A.; Saito, M. *Int. J. Quantum Chem.* **1984**, *25*, 527.
(3) Sarai, A.; Saito, M. *Int. J. Quantum Chem.* **1985**, *28*, 399.
(4) Otto, P.; Clementi, E.; Ladik, J.; Martino, F. *J. Chem. Phys.* **1984**, *80*, 5294.
(5) Pichierri, F.; Aida, A.; Gromiha, M. M.; Sarai, A. *J. Am. Chem. Soc.* **1999**, *121*, 6152.
(6) Galetich, I.; Kosevich, M.; Shelkovsky, V.; Stepanian, S. G.; Blagoi, Y. P.; Adamowicz, L. *J. Mol. Struct.* **1999**, *478*, 155.
(7) Smolyaninova, T. I.; Bruskov, V. I.; Kashparova, Ye. V. *Mol. Biology (Russian)* **1985**, *19*, 992.
(8) Helene, C.; Lancelot, G. *Prog. Biophys. Mol. Biol.* **1982**, *39*, 68.
(9) Brown, R. D.; Godfrey, P. D.; Storey, J. W.; Bassez, M. P. *J. Chem. Soc. Chem. Commun.* **1978**, *14*, 547.
(10) Suenram, R. D.; Lovas, F. J. *J. Mol. Spec.* **1978**, *72*, 372.
(11) Schafer, L.; Sellers, H. L.; Lovas, F. J.; Suenram, R. D. *J. Am. Chem. Soc.* **1980**, *102*, 6566.
(12) Locke, M. J.; MvIver, R. T., Jr. *J. Am. Chem. Soc.* **1983**, *105*, 4226.
(13) Iijima, K.; Tanaka, K.; Onuma, S. *J. Mol. Struct.* **1991**, *246*, 257.
(14) Godfrey, P. D.; Brown, R. D. *J. Am. Chem. Soc.* **1995**, *117*, 2019.
(15) Godfrey, P. D.; Brown, R. D.; Rodgers, F. M.; *J. Mol. Struct.* **1996**, *376*, 65.
(16) Wyttenbach, T.; Bushnell, J. E.; Bowers, M. T. *J. Am. Chem. Soc.* **1998**, *120*, 5098.
(17) Wyttenbach, T.; Mathias, W.; Bowers, M. T. *J. Am. Chem. Soc.* **2000**, *122*, 3458.
(18) Wyttenbach, T.; Mathias, W.; Bowers, M. T. *Int. J. Mass Spectrom.* **1999**, *183*, 243.
(19) Vishveshwara, S.; Pople, J. A. *J. Am. Chem. Soc.* **1977**, *99*, 2422.
(20) Sellers, H. L.; Schafer, *J. Am. Chem. Soc.* **1978**, *100*, 7728.
(21) Dykstra, C. E.; Chiles, R. A.; Garrett, M. D. *J. Comput. Chem.* **1981**, *2*, 266.
(22) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1991**, *113*, 7917.
(23) Frey, R. F.; Coffin, J.; Newton, S. Q.; Ramek, M.; Cheng, V. K. W.; Momany, F. A.; Schafer, L. *J. Am. Chem. Soc.* **1992**, *114*, 5368.
(24) Csaszar, A. G. *J. Am. Chem. Soc.* **1992**, *114*, 9568.
(25) Hu, C.-H.; Shen, M.; Schaefer, H. F., III. *J. Am. Chem. Soc.* **1993**, *115*, 2923.
(26) Gordon, M. S.; Jensen, J. H. *Acc. Chem. Res.* **1996**, *29*, 536.
(27) Nguyen, D. T.; Scheiner, A. C.; Andzelm, J. W.; Sirois, S.; Salahub, D. R.; Hagler, A. T. *J. Comput. Chem.* **1997**, *18*, 1609.
(28) Ruterjans, H.; Kaun, E.; Hall, W. E.; Limbach, H. H. *Nucleic Acids Res.* **1982**, *10*, 7.
(29) Kubota, M.; Kobayashi, T. *J. Electron Spectrosc. Relat. Phenom.* **1996**, *82*, 61.
(30) Chin, S.; Scot, I.; Szczepaniak, K.; Person, W. B. *J. Am. Chem. Soc.* **1984**, *106*, 3415.
(31) Brown, R. D.; Godfrey, P. D.; McNaughton, D.; Pierlot, A. P. *J. Am. Chem. Soc.* **1988**, *110*, 2329.
(32) Lowdin, P. O. *Rev. Mod. Phys.* **1963**, *35*, 724.
(33) Kryachko, E. S.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **2001**, *105*, 1288.
(34) Leszczynski, J. *J. Phys. Chem.* **1992**, *96*, 1649.
(35) Hobza, P.; Sponer J. *Chem. Rev.* **1999**, *99*, 3247.
(36) Rybak, S.; Szalewicz, K.; Jeziorski, B.; Corongiu, G. *Chem. Phys. Lett.* **1992**, *199*, 567.
(37) Smets, J.; McCarthy, W. J.; Adamowicz, L. *J. Phys. Chem.* **1996**, *100*, 14655.
(38) Chandra, A. K.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **1998**, *102*, 6010.
(39) van Mourik, T.; Price, S. L.; Clary, D. C. *J. Phys. Chem. A* **1999**, *103*, 1611.

(40) Dolgounitcheva, O.; Zakrzewski, V. G.; Ortiz, J. V.; *J. Phys. Chem. A* **1999**, *103*, 7912.
(41) van Mourik, T.; Benoit, D. M.; Price, S. L.; Clary, D. C. *Phys. Chem. Chem. Phys.* **2000**, *2*, 1281.
(42) Gutowski, M.; Dabkowska, I.; Rak, J.; Xu, S.; Nilles, J. M.; Radisic, D.; Bowen, K. H.; *Eur. Phys. J. D*, accepted for publication.
(43) Dąbkowska, I.; Gutowski, M.; Rak, J. *First Russian-Ukrainian-Polish Conference on Molecular Interactions*; Gdansk, Poland, June 10−16, 2001.
(44) Dąbkowska, I.; Rak, J.; Gutowski, M. Manuscript in preparation.
(45) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(46) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(47) Lee, C.; Yang, W.; Paar, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(48) Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724.
(49) Hehre, W. J.; Ditchfield, R.; Pople J. A. *J. Chem. Phys.* **1972**, *56*, 2257.
(50) Rak, J.; Skurski, P.; Simons, J.; Gutowski, M. *J. Am. Chem. Soc.* **2001**, *123*, 11695.
(51) Skurski, P.; Rak, J.; Simons, J.; Gutowski, M. *J. Am. Chem. Soc.* **2001**, *123*, 11073.
(52) Kryachko, E. S.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **2001**, *105*, 1934.
(53) Chandra, A. K.; Nguyen, M. T.; Uchimaru, T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **1999**, *103*, 8853.
(54) Dkhissi, A.; Adamowicz, L.; Maes, G. *J. Phys. Chem. A* **2000**, *104*, 2112.
(55) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
(56) Kendall, R. A.; Simons, J.; Gutowski, M.; Chałasiński, G. *J. Phys. Chem.* **1989**, *93*, 621.
(57) Szalewicz, K.; Jeziorski, B. *J. Chem. Phys.* **1998**, *109*, 1198.
(58) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(59) Gutowski, M.; Chałasiński, G. *J. Chem. Phys.* **1993**, *98*, 5540.
(60) Gutowski, M.; Chałasiński, G.; van Duijneveldt-van de Rijdt, J. G. C. M. *Int. J. Quantum Chem.* **1984**, *26*, 971.
(61) Chałasiński, G.; Gutowski, M. *Chem. Rev.* **1988,** *88*, 943.
(62) Stewart, J. J. P. MOPAC2000, Fujitsu Limited, 1999.
(63) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian, Inc.: Pittsburgh, PA, 1998.
(64) Harrison, R. J.; et al. *NWChem, A Computational Chemistry Package for Parallel Computers*, version 4.0.1; Pacific Northwest National Laboratory: Richland, WA, 2001.
(65) National Institute of Standards and Technology http://webbook.nist-.gov/chemistry.
(66) Locke, M. J.; McIver, R. T., Jr. *J. Am. Chem. Soc.* **1983**, *105*, 4226.
(67) Nuftakov, M. V.; Vasil'ev, Y. V.; Mazunov, V. A. *Rapid Commun. Mass. Spectrom.* **1999**, *13*, 1104.
(68) Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Rosado, M. T. S.; Duarte, M. L. T. S.; Fausto, R.; Adamowicz, L. *J. Phys. Chem. A* **1998**, *102*, 1041.
(69) Kascher, R.; Hohl, D. *J. Phys. Chem. A* **1998**, *102*, 5111.
(70) Stewart, R. F.; Jensen, L. H. *Acta Crystallogr.* **1967**, *23*, 1102.
(71) Ferenczy, G.; Harsanyi, L.; Rozsondai, B.; Hargittai, I. *J. Mol. Struct.* **1986**, *140*, 71.
(72) Gaigeot, M.-P.; Ghomi, M. *J. Phys. Chem. B* **2001**, *105*, 5007.
(73) Chaban, G. M.; Jung, J. O.; Gerber, R. B. *J. Phys. Chem. A* **2000**, *104*, 10035.
(74) Kulakowska, I.; Geller, M.; Lesyng, B.; Wierzchowski, K. L. *Biochim. Biophys. Acta* **1974**, *361*, 119.
(75) Lovas, F. J.; Kawashima, Y.; Grabow, J.-U.; Suenram, R. D.; Freser, G. T.; Hirota, E. *Astrophys. J.* **1995**, *455*, 201.
(76) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209 and 221.
(77) Bernal-Uruchurtu, M. I.; Martins-Costa, M. T. C.; Millot, C.; Ruiz-Lopez, M. F. *J. Comput. Chem.* **2000**, *21*, 572.
(78) Pecul, M.; Leszczynski, J.; Sadlej, J. *J. Chem. Phys.* **2000**, *112*, 7930.
(79) Vargas, R.; Garza, J.; Friesner, R. A.; Stern, H.; Hay, B. P.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 4963.
(80) Garza, V. R.; Dixon, D. A.; Hay, B. P. *J. Am. Chem. Soc.* **2000**, *122*, 4750.
(81) Kryachko, E. S.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **2001**, *105*, 7118.
(82) Chałasiński, G.; Szczesniak, M. M. *Chem. Rev.* **2000**, *100*, 4227.
(83) Abraham, M. H.; Platts, J. A. *J. Org. Chem.* **2001**, *66*, 3484.