Molecular Orbital Calculations on the Protonation of Hydrogen-Bonded Formamide Chains. Implications for Peptides

Sandy Moisan and J. J. Dannenberg*

Department of Chemistry, City University of New York—Hunter College and The Graduate School, 695 Park Avenue, New York, New York 10021

Received: June 24, 2003; In Final Form: September 28, 2003

We report density functional studies of the protonation of H-bonded formamide chains containing up to 10 monomeric units. These chains contain H bonds that are similar to those in peptides and proteins. All structures considered were completely geometrically optimized at the B3LYP/D95** level of calculation. The proton affinities of the chains are greatest at the terminal C═O oxygen atoms. They increase significantly with the length of the formamide chain from two to five formamides. Protonation of the terminal C═O of chains containing five or more formamides results in the transfer of the N—H proton of the terminal (protonated) monomer across the H bond to the adjacent formamide. Protonations at the NH_2 termini of the formamide chains are generally unfavorable as they result in the rupture of the proximate H bond. The proton affinities at the C═O's of formamide chains containing three or more monomers exceed that of the amino group of glycine, implying that peptides that contain such H-bonding chains might be preferentially protonated on a C═O rather than a terminal amino group, in contrast to small oligopeptides where H-bonding chains do not form. The quasi-linear relationship between H-bond strength and length clearly does not hold for the protonated chains. The implications for studies on peptides are discussed.

Cooperative hydrogen bonding plays an important role in the relative energies of the secondary structures of peptides. Other reports from our group have demonstrated the unusually large extent of the energetic and structural cooperativity as well as their effects upon the vibrational coupling in extended chains of H-bonding formamides.[1−3] Related studies have shown that cooperative H bonding influences the formation of 3_{10} helices, even in small peptides,[4,5] and the formation of α-helices in larger peptides.[4,6] These theoretical results are in qualitative agreement with Kemp's experimental reports.[7] However, these studies do not address two important related problems: (1) Can these secondary structures be more easily protonated than others less influenced by cooperative H bonding? (2) Might such protonated peptides be important intermediates in protein folding?

In this paper, we begin to address these questions by studying the proton affinities at various positions in H-bonding formamide chains and the effect of such protonation on the structure of the chains and the energies of the individual H bonds in these chains. These studies primarily address the first of the two questions posed above. They generate implications about the second that would require further study.

Methods

Molecular orbital calculations were performed using a hybrid DFT method at the B3LYP/D95(d,p) level. This method combines Becke's three-parameter functional[8] with the nonlocal correlation provided by the correlation functional of Lee, Yang, and Parr[9] using the Gaussian 98 suite of programs.[10] The geometries were completely optimized with the sole constraints that all the structures were of C_s symmetry (all atoms are coplanar). The vibrational frequencies were calculated for the planar structures, using the normal harmonic approximations employed in the Gaussian 98 program, to verify the stationary points and calculated the enthalpies of the various species.[10] All frequencies were real except for some very low frequency imaginary frequencies (less than 25 cm^{-1}) that involved out-of-plane twists between pairs of formamides in some of the longer formamide chains.

Results

Formamide Monomer. Protonation of the formamide monomer can, in principle, occur at either the nitrogen or the oxygen. Several theoretical[11−17] and experimental[15] studies indicate protonation at the C═O to be unambiguously favored in the gas phase. Our results predict the enthalpic proton affinity of formamide at the C═O to be 202.78 kcal/mol, more than 18 kcal/mol more than that obtained at the NH_2 (182.32 kcal/mol), in agreement with the previous studies.

Formamide Chains. Figure 1 shows the structure and numbering scheme of a protonated formamide chain containing five monomer units. This figure also depicts the proton transfer that occurs in chains of five or more formamides (see discussion below).

Proton Affinities of the Chains. Protonation at the terminal NH_2 of a H-bonding formamide chain generally leads to the rupture of the H bond between the protonated formamide and the rest of the chain. We were unable to geometrically optimize structures of this type as the H-bonding distance simply lengthened until the interaction between the departing protonated formamide and the rest of the chain defined a very flat part of the energy surface. Protonation at any of the nitrogens or oxygens that participate in the H-bonds (i.e., those that are not at the N-terminus of the chain) also leads to rupture of the H bond at the protonated position. Only protonation at the terminal C═O leads to a stable protonated H-bonding chain. As one might expect, the (enthalpic) proton affinity of the H-bonding

10.1021/jp035791g CCC: $25.00
Published on Web 10/29/2003

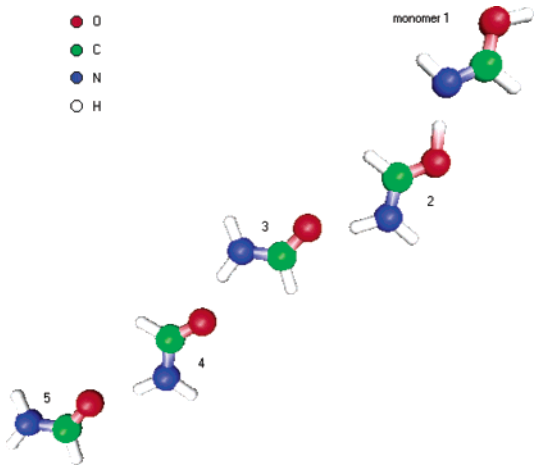

Figure 1. Structure of protonated pentamer of formamide indicating the numbering scheme and the proton transferred from the first to second formamide. Note the repeating unit in the chain is two formamides.

TABLE 1: Proton Affinities of Formamide Chains[a]

monomer	200.8	hexamer	238.7
dimer	222.9	heptamer	239.2
trimer	231.7	octamer	241.5
tetramer	234.8	nonamer	242.6
pentamer	240.8	decamer	240.8

[a] Enthalpies are in kcal/mol.

chain increases with chain length, from 222.9 kcal/mol for the dimer (20.1 more than the monomer) to more than 240 kcal/mol for the longest chains (see Table 1 and Figure 2).

One can easily obtain the H-bonding energies by simple difference in energy or enthalpy between the larger chain and the two fragments. This is analogous to the method that we previously used in our studies of neutral H-bonding chains.[1,3] The energies of the individual H bonds in each of the protonated chains are presented in Table 2 and Figure 3. As expected, these H bonds are significantly stronger than those in the neutral formamide chains, where the strongest H bond in a decamer has been reported to be 12.12 kcal/mol. [3] The H bond nearest the protonated fomamide is clearly the strongest. For each protonated chain containing five or more formamides, the H-bonding proton of the protonated monomer is transferred to the C═O of the second formamide, creating a C−OH···N H bond in place of the C═O···HN interaction in the corresponding neutral species (see Figure 1). For these species, the strongest C═O···HN interaction becomes the one at the adjacent (second position) as the second formamide is the protonated molecule in these chains. The energies of the other H bonds decrease with distance from the protonated position. One notices an oscillation in the energies for analogous H bonds in systems containing even and odd numbers of formamides. This phenomenon results from the fact that the linear chain contains a repeating unit of two monomers (see Figure 3).

With the exceptions noted below, the H-bond lengths follow the order of the energies within each chain and for the corresponding position in the different chains (i.e., the third H bond of the hexamer is stronger than the third H bond of the pentamer). The strongest H bonds (nearest the protonation site) are the shortest. H-bond lengths at each position in the chains become shorter and stronger as the chains become longer (see Figures 4 and 5 and Table 3). The quasi-linear relationship between H-bond length and energy that is observed in the neutral formamide chains[3] breaks down upon protonation. Figure 5 shows that each H-bonding position within a protonated chain

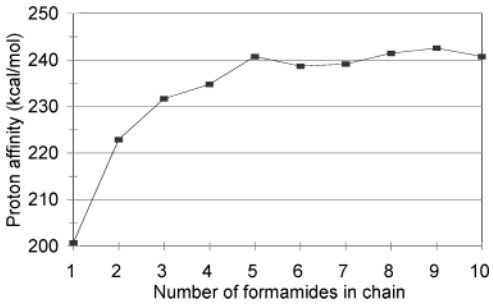

Figure 2. Proton affinities (enthalpies at 298 K in kcal/mol) of protonated formamide chains as a function of chain length.

approaches a different limiting O···H distance despite becoming appreciably stronger as the chains increase in length. Furthermore, the limiting distance for each H-bonding position does not appear to be directly related to its strength. For example, the H bond at the third position of the decamer is both stronger and longer than those in the second position of the pentamer, hexamer, and heptamer.

Formamide Dimer. The preference for protonation at the C═O in the monomer becomes amplified in the dimer. Protonation at this site leads to a consequent strengthening of the H bond, which increases from 4.49, in the neutral dimer,[1,3] to 23.20 kcal/mol. The H bond shortens upon protonation, as expected, from 1.923 to 1.534 Å. However, protonation on the terminal nitrogen is markedly disfavored. It leads to the rupture of the H bond.

Formamide Trimer and Tetramer. Protonation of the trimer or tetramer occurs preferentially at the terminal C═O, as in the dimer. The proton affinity at the terminal oxygen increases in the expected order tetramer > trimer > dimer. The differences in the proton affinities of the chains upon adding another formamide to a chain are 8.8, 3.1, and 6.0 kcal/mol for dimer, trimer, and tetramer, respectively. The differences between the proton affinities' chains containing even and odd numbers of monomers become apparent in the greater increase upon adding a fifth formamide to a tetramer as compared to a fourth to a trimer. The H bond between the first two formamides shortens and increases in strength as the chain lengthens.

Pentamer through Decamer. For the chains containing five or more monomeric units, protonation at the terminal C═O causes the H-bonding proton of the protonated formamide to migrate to the oxygen of the adjacent monomer. Thus, the protonated chain consists of the *iminol* tautomer of the terminal formamide H bonded to a chain containing one fewer formamide, which is now protonated at its terminal C═O (Figure 1). This migration effectively moves the formally positive formamide from the end of the chain (where it can be stabilized by only one H bond) into the chain (where it is stabilized by two H bonds). In principle, one can consider the possibility of further migration of the positively charged formamide into a more central position in the chain. However, each time a proton migrates, another formamide molecule is converted to its less stable iminol tautomer. The driving force for the H migrations thus abates after the first proton transfer. Upon the increase of the chain from four to five formamides, the proton affinity increases by another 6.0 kcal/mol as the proton transfer occurs between the first two formamides. Protonation of chains containing more than five formamides all have similar effects upon the first H bond (as the proton is transferred each time). The effect of increasing the chain length beyond five formamides upon proton affinity attenuates significantly. In fact, the proton affinity decreases slightly upon addition of a sixth

TABLE 2: Hydrogen-Bonding Interaction Energies (kcal/mol) in Protonated Formamide Chains[a]

position	1	2	3	4	5	6	7	8	9
dimer	−27.6893								
trimer	−41.1257	−16.6275							
tetramer	−48.1077	−25.3445	−11.9080						
pentamer		−33.0456	−21.3441	−12.6271					
hexamer		−37.8808	−26.7071	−19.7252	−10.2890				
heptamer		−40.1877	−29.4226	−22.9684	−15.2673	−8.1693			
octamer		−44.5269	−33.8279	−27.7823	−20.6090	−15.2460	−10.2677		
nonamer		−45.6846	−35.1036	−29.1241	−22.3594	−17.5242	−14.2809	−7.2042	
decamer		−50.8660	−40.2893	−34.4278	−27.7292	−23.3025	−20.5870	−15.2454	−11.2322

[a] No energies are listed for the first H-bonding position in chains with five or more formamides as proton transfer from the second to first formamide occurs for these species.

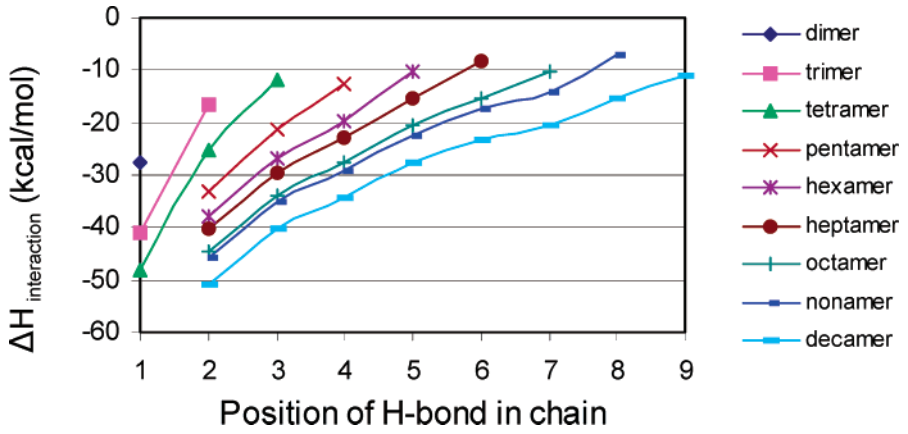

Figure 3. H-bond enthalpies at 298 K of protonated formamide chains as a function of chain length and H-bond position (kcal/mol).

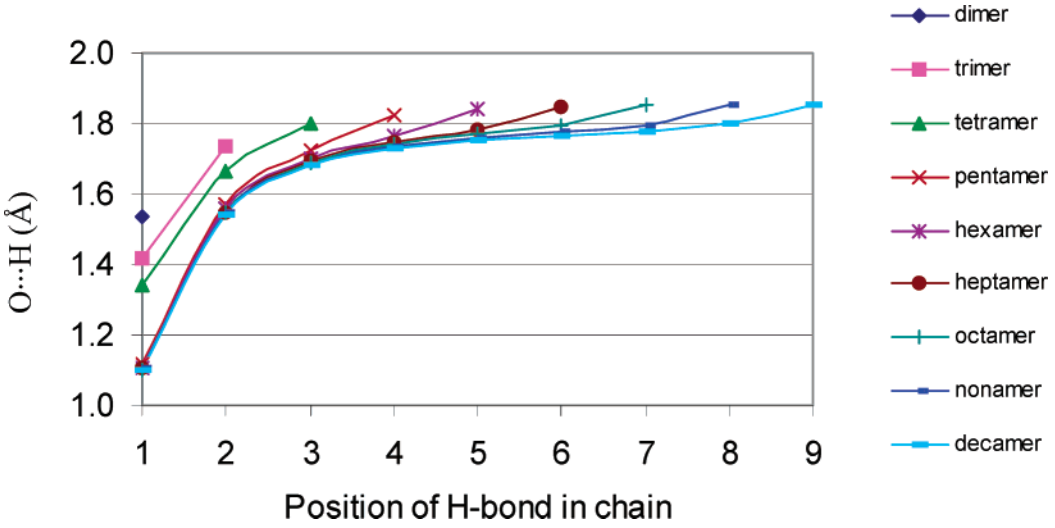

Figure 4. O···H distances for protonated formamide chains as a function of chain length and H-bond position (Å).

formamide to the chain and then slightly increases upon further addition of formamide monomers until the chain reaches nine formamides. The addition of a tenth formamide results in a slightly reduced proton affinity. This behavior might be due to the onset of a pattern of alternating proton affinities for chains containing odd and even quantities of monomers. Calculations on longer chains would be required to confirm this.

Gilli has proposed that N−H···O H bonds can be resonance assisted.[18−20] The transfer of the H-bonding proton of the terminal formamide, which is protonated to the adjacent monomer (as mentioned above), provides an interesting example of resonance-assisted hydrogen bonding (RAHB). Protonation of the terminal C═O's of dimers, trimers, and tetramers strengthens the H-bonding interaction between the terminal (protonated) formamide (the H-bond donor) and the adjacent monomer (H-bond acceptor). When the chain contains five or more formamides, the proton migration leaves an H-bond in which the terminal (protonated) formamide has become the H-bond acceptor, while the adjacent monomer has become the donor. One can consider this situation as if one resonance structure dominated the protonated chains of two to four monomers, while the other dominates the larger chains. Such resonance interactions support the importance of a strong covalent component in H bonding of this type.

Implications for Peptides. The present results have several clear implications for the protonation of peptides. First of all, the terminal C═O of a H-bonding chain will be the most basic position. Different secondary structures of peptides have different H-bonding patterns. For example, an α-helix contains three H-bonding chains while a 3_{10}-helix contains only two. A β-sheet can contain many chains. In principle, the terminal C═O of any of these chains might be protonated. Protonation at

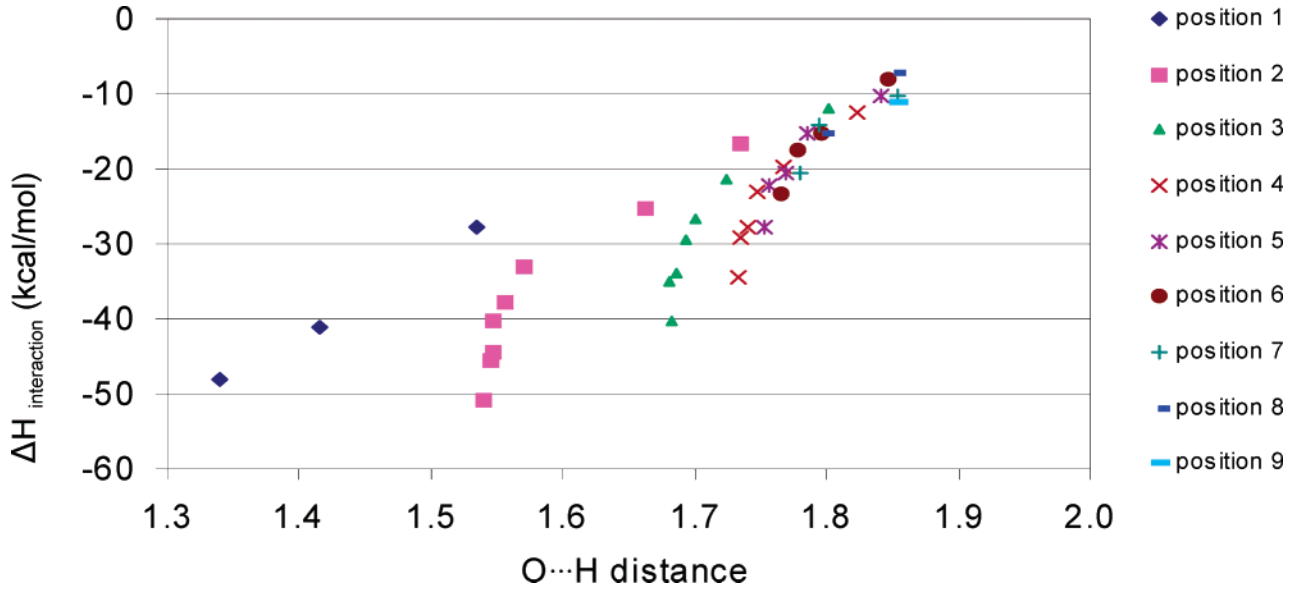

Figure 5. H-bond enthalpies at 298 K as a function of O···H distances in angstroms for protonated formamide chains.

TABLE 3: O···H Distances (Å) as a Function of Position in Protonated Chains of Formamides

	position of H-bond in chain								
O···H	1	2	3	4	5	6	7	8	9
dimer	1.5338								
trimer	1.4163	1.7349							
tetramer	1.3393	1.6627	1.8015						
pentamer	1.1189	1.5708	1.7234	1.8236					
hexamer	1.1070	1.5568	1.7011	1.7673	1.8412				
heptamer	1.1055	1.5468	1.6934	1.7478	1.7851	1.8467			
octamer	1.1020	1.5480	1.6865	1.7397	1.7689	1.7953	1.8541		
nonamer	1.1034	1.5449	1.6801	1.7355	1.7569	1.7788	1.7935	1.8521	
decamer	1.0996	1.5405	1.6829	1.7323	1.7526	1.7658	1.7792	1.7987	1.8544

[a] Positions are defined as in previous table.

any of the other (H-bonding) C═O positions should be energetically disfavored and lead to destabilization of the helical structure. While protonation at either the oxygen or the nitrogen of one of the H-bonds in the formamide chains leads to rupture of the H-bonding structure, the remaining H-bonding chains in helical structures may be sufficiently strong to maintain the helix upon protonation of one of the interior H bonds. However, protonation of one of the amidic nitrogens at the terminal H bond would certainly cause an unraveling at the end of the helix.

The terminal NH_2 of a peptide is an amine (rather than part of an amide). Thus, the present results cannot properly be used as a model for protonation of this position. However, the proton affinity of glycine at the NH_2, calculated using the same methods, is 229.3 kcal/mol, which lies between the proton affinities of the formamide dimer and trimer. Thus, one might expect that peptides containing H-bonding chains with more than three H bonds tend to be favorably protonated on a terminal C═O of the H-bonding chain. While the proton affinity at C═O for H-bonding chains significantly increases with chain length, one would not expect similar increases for protonation at the N-terminus of a peptide, which is not directly involved in the H-bonding chains (although it will form other H bonds). Since the present results suggest protonation to be more facile at a C═O than the N-terminus of peptides with amidic H-bonding chains, the traditional zwitterionic structure that prevails in aqueous solution for amino acids and small peptides might not be appropriate for many larger peptides. While further study will be needed to clarify this point, one should be cautious in assuming the classical zwitterionic structure to be generally appropriate of all peptides.

NMR methods, particularly $^{13}C-^{15}N$ trans H-bond coupling,[21] are becoming methods of choice for measuring the H-bonding distances in peptides.[22−25] These studies are sometimes performed under acidic conditions, where the peptide H-bonding chains might be protonated at one or more of the terminal C═O's. To the extent that this might occur in a sample, the H-bonding distances may need to be interpreted differently from the simple model of an unprotonated peptide. The H-bonding distances would change. Furthermore, a proton might migrate from the NH on the protonated amino acid to the oxygen of the amino acid to which it forms its H bond if the H-bonding chain be sufficiently long. This would be analogous to the behavior reported here for the longer formamide chains. NMR studies of large helical peptides might provide evidence for his behavior if it exists.

Since the proton affinity of the H-bonding chains becomes greater as the chain becomes longer, the protonation of small helices might lead to expansion of the helical structure. Similarly, helical formation might be acid catalyzed via protonation of a C═O in some prehelical conformation.

Conclusions

Unlike amino acids or small peptides without internal H bonds, chains of H-bonding formamides have their greatest proton affinities at the terminal C═O. Protonation at the terminal NH_2 destabilizes the H-bonding chain, causing a rupture of the H bond between the protonated formamide and its neighbor. The protons of the H bonds in the first position of all chains containing five or more formamides are transferred to the second

formamide. Within each chain, the H bonds decrease in energy and increase in length with distance from the protonated C═O. However, the H-bond lengths approach different limiting values for each corresponding H-bond position in chains of increasing length. Thus, the quasi-linear relationship between H-bond strength and length clearly does not hold for the protonated chains.

These observations strongly imply that protonation at C═O may provide a simple means of acid catalysis for formation of secondary structural units that contain H-bonding chains such as α-helices and β-sheets. Because of the high proton affinities of the terminal C═O's in H-bonding chains, peptides that contain such H-bonding motifs might easily exist in protonated states even in vivo.

Acknowledgment. This work was supported in part by Grants from the National Institutes of Health (S06GM60654) and from PSC-CUNY. Some calculations were performed on the CUNY Graduate School Research Computing Cluster.

References and Notes

(1) Kobko, N.; Paraskevas, L.; del Rio, E.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2001**, *123*, 4348.
(2) Kobko, N.; Dannenberg, J. J. *J. Phys. Chem. A* **2003**, *107*, 6688.
(3) Kobko, N.; Dannenberg, J. J. *J. Phys. Chem. B* **2003**, *107*, in press.
(4) Wu, Y.-D.; Zhao, Y.-L. *J. Am. Chem. Soc.* **2001**, *123*, 5313.
(5) Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2003**, *125*, in press.
(6) Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2003**, *125*, 8124.
(7) Kennedy, R. J.; Tsang, K.-Y.; Kemp, D. S. *J. Am. Chem. Soc.* **2002**, *124*, 934.
(8) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(9) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(10) 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*, revision A.10; Gaussian, Inc.: Pittsburgh, PA, 1998.
(11) Ou, M.-C.; Chu, S.-Y. *J. Phys. Chem.* **1995**, *99*, 556.
(12) Tortajada, J.; Leon, E.; Morizur, J. P.; Luna, A.; Mo, O.; Yanez, M. *J. Phys. Chem.* **1995**, *99*, 13890.
(13) Bagno, A.; Scorrano, G. *J. Phys. Chem.* **1996**, *100*, 1536.
(14) Rablen, P. R.; Lockman, J. W.; Jorgensen, W. L. *J. Phys. Chem. A* **1998**, *102*, 3782.
(15) Herreros, M.; Gal, J.-F.; Maria, P.-C.; Decouzon, M. *Eur. Mass Spectrom.* **1999**, *5*, 259.
(16) Kim, N.-K.; Lee, H.-J.; Choi, K.-H.; Yu, J.-A.; Yoon, C.-J.; Park, J.; Choi, Y.-S. *J. Phys. Chem. A* **2000**, *104*, 5572.
(17) Pliego, J. R., Jr.; Riveros, J. M. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1622.
(18) Gilli, P.; Bertolasi, V.; Pretto, L.; Lycka, A.; Gilli, G. *J. Am. Chem. Soc.* **2002**, *124*, 13554.
(19) Gilli, P.; Bertolasi, V.; Ferretti, V.; Gilli, G. *J. Am. Chem. Soc.* **2000**, *122*, 10405.
(20) Bertolasi, V.; Nanni, L.; Gilli, P.; Ferretti, V.; Gilli, G.; Issa, Y. M.; Sherif, O. E. *New J. Chem.* **1994**, *18*, 251.
(21) Grzesiek, S.; Cordier, F.; Dingley, A. J. *Methods Enzymol.* **2001**, *338*, 111.
(22) Jaravine, V. A.; Alexandrescu, A. T.; Grzesiek, S. *Protein Sci.* **2001**, *10*, 943.
(23) Cornilescu, G.; Ramirez, B. E.; Frank, M. K.; Clore, G. M.; Gronenborn, A. M.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 6275.
(24) Juranic, N.; Moncrieffe, M. C.; Likic, V. A.; Pendergast, F. G.; Macura, S. *J. Am. Chem. Soc.* **2002**, *124*, 14221.
(25) Barfield, M. *J. Am. Chem. Soc.* **2002**, *124*, 4158.