Relative Strengths of NH··O and CH··O Hydrogen Bonds between Polypeptide Chain Segments

Steve Scheiner*

Department of Chemistry & Biochemistry, Utah State University, Logan, Utah 84322-0300

Received: June 23, 2005

Correlated ab initio calculations are used to compare the energetics when the CH and NH groups of the model dipeptide $CHONHCH_2CONH_2$ are each allowed to form a H-bond with the proton acceptor O of a peptide group. When the dipeptide is in its C_7 conformation, the NH··O H-bond energy is found to be 7.4 kcal/mol, as compared to only 2.8 kcal/mol for the CH··O interaction. On the other hand, the situation reverses, and the CH··O H-bond becomes stronger than NH··O, when the dipeptide adopts a C_5 structure. This reversal is important as C_5 is nearly equal in stability to C_7 for the dipeptide, and is representative of the commonly observed β-sheet structure in a protein. Immersing the dipeptide−peptide pair in a model solvent weakens both sorts of H-bonds, and in a fairly uniform manner. Consequently, the trends observed in the in vacuo situation retain their validity in either aqueous solution or the protein interior. Likewise, the desolvation penalty, suffered by removing a H-bonded complex from water and placing it in the less polar interior of a protein, is quite similar for the NH··O and CH··O bonds.

Introduction

Although discounted for some time in the past, a particular type of weak H-bond has undergone a resurgence of attention. The ability of a CH group to act as proton donor was originally thought limited to strong acids such as HCN, but a plethora of work over the past decade[1−7] has convincingly demonstrated that even nominally weaker proton donors can form CH··X H-bonds (X = O,N). The variety of systems containing CH··X bonds includes biological macromolecules of every stripe.[8−15]

In addition to geometric data that have pointed to the existence of these CH··X bonds, there is growing evidence that they play an active structural role[16−18] in biomolecules. For example, these bonds are at least partly responsible for the observed registration of α-helices and β-sheets,[18−20] and they can determine the direction of protein chain folding.[21] CH··O bonds have been advanced as the driving force for ligand selectivity[22] into certain hydrophobic pockets. These bonds help stabilize a chain reversal motif[21,23] and affect the rotation angles of Trp residues in proteins.[24]

While it is now apparent that CH··X bonds have a direct influence upon protein structure, their energetic contribution remains an open question. Crystal studies alone cannot provide definitive information "because of significant structural interference from many other interactions[25]". Kallenbach et al.[26] have commented that "there are few experimental studies of the energetic contribution from any type of C−H··O H-bonds in protein or polypeptide systems", and another group[27] went so far as to claim that at this time "it is not possible to know whether they are attractive or repulsive". An alternate form of this question has been posed:[28] what is the penalty in binding affinity for replacing a traditional H-bond with a CH··O H-bond? It was especially intriguing that these authors found that replacement of NH donor by CH has little effect on binding affinities, and they went on to suggest that "if traditional H-bonds and CH··O H-bonds are indeed roughly interchangeable, the impact on ligand design should be tremendous."

Recent attempts to address this crucial issue provide only rough and contradictory estimates. Empirical potentials[29] estimate that the collection of CH··O H-bonds, as a group, may contribute anywhere between 17% and 50% to the total interaction energy of protein−protein interfaces. A combined CD and NMR study of a model polypeptide found the binding energy may be 0.5 kcal/mol, or zero, depending upon the direction of the bond.[26] A more recent examination of a particular CH··O H-bond within a protein[30] estimated its energy as 0.9 kcal/mol, while another work in that same year[31] found no evidence of an energetic contribution at all.

It is within the context of conflicting experimental estimates that modern theoretical calculations have shown their worth. Unlike crystal structures where the relatively weak CH··O interaction is typically a secondary factor in the overall geometry, so even its geometric preferences are obscured by other forces and consequently difficult to elucidate, calculations can be performed on systems designed specifically to examine the CH··O bond in isolation from other forces. In such a system, the strength of the latter bond can be unambiguously determined, because other contributors to the interaction are absent. Much of this work has been reviewed recently.[32]

However, nearly all of the work to this point has dealt with small molecules such as CH_4, ethylene, and cyclopropane,[33−43] poor models for biomolecules. Moreover, most of the prior calculations have addressed the in vacuo situation, hardly compatible with placement within a protein or solution. There has been scant attention paid to the placement of even these small models within a realistic environment: a dimer involving one solvent molecule[44,45] simply cannot be considered an adequate model of solvation.

The primary goal of the present work is an assessment of the energetic contribution that CH··O H-bonds make to protein structure, using realistic models, in relevant surroundings, and all within the framework of reliably accurate ab initio theory.

* Corresponding author. E-mail: scheiner@cc.usu.edu.

10.1021/jp053416d CCC: $30.25
Published on Web 07/30/2005

The most common CH··O bonds occurring in proteins involve $C^{\alpha}H$ as donor;[8] the polypeptide proton donor is consequently modeled here by $CHONHC^{\alpha}H_2CONH_2$, wherein the central $C^{\alpha}H_2$ proton donor is surrounded on both sides by peptide groups, as would be the case in a real polypeptide chain. This particular model contains also a peptide NH group, which can participate in the classical NH··O H-bond so common in protein structures. The location of both CH and NH donors within the context of the same molecule provides a direct means by which to compare them to one another, which is, after all, the ultimate goal. A second advantage of this strategy is that many of the errors that may be associated with a particular level of theory, imperfect modeling of environment, etc., will cancel one another as they are present in both NH··O and CH··O in very nearly equal measure.

The proton acceptor groups considered to date have generally consisted of hydroxyl groups situated on small molecules, such as H_2O or CH_3OH. Even when a carbonyl oxygen was considered, it has typically been located on a simple aldehyde or ketone, as opposed to the O atom of a peptide. The proton acceptor here is taken as the model peptide, formamide; however, other calculations are performed with a larger model to ensure the results are not compromised by the size of the proton acceptor model. Another problem with the vast majority of earlier ab initio calculations of CH··O bonds has been their limitation to the gas phase, quite unlike the actual protein environment. The entire dipeptide−proton acceptor system is thus immersed in a model solution environment and examined over a range of different solvent polarities. (The $CHONHC^{\alpha}H_2CONH_2$ is termed a dipeptide to reflect the presence of two peptide groups, flanking the central CH_2.)

Method of Calculations

Ab initio calculations were carried out using the Gaussian 03 code[46] with double-ζ quality basis sets: 6-31+G** contains polarization functions on all atoms, augmented by diffuse functions on heavy atoms; 6-31G* is somewhat smaller and lacks the diffuse functions. The B3LYP variant of density functional theory (DFT)[47,48] was used as an intermediate-level means of including electron correlation, as it has had good success in the past[49−52] in dealing with H-bonded systems. Second-order Møller−Plesset (MP2) was used as a more complete means of including correlation.

Solvent was treated as a continuous medium, with a dielectric constant ϵ, via the self-consistent reaction field (SCRF) approach.[53−55] The polarizable continuum method (PCM)[56−58] embeds the solute in a cavity that reproduces the shape of the molecule by a series of overlapping spheres. The particular variant of this method used here is the conductor polarized continuum model (CPCM),[59] wherein the apparent charges distributed on the cavity surface are such that the total electrostatic potential cancels on the surface. Recent calculations[60] had shown that the CPCM variant provides results that are in good agreement with other approaches, notably PCM and SCIPCM, in treating CH··O as well as conventional H-bonds. The UAHF (United Atom model for Hartree-Fock/6-31G*) definition[61] was used for the construction of the solute cavity.

Results

Full geometry optimization of the glycine dipeptide $CHONHCH_2CONH_2$ yielded the structure illustrated in Figure 1. This sort of geometry has been recently identified[62,63] for short polypeptides in the gas phase and is consistent with a wealth of prior computations at various levels of theory[64−67] that establish it as the global minimum. Both amide units are very nearly planar, and the (φ,ψ) pair of angles is equal to (−83°,-74°) at the MP2/6-31+G** level. (φ is defined as the backbone CNCC dihedral angle, and ψ refers to NCCN.) A dominant characteristic of this structure is the internal H-bond between the terminal O and NH_2 groups. As indicated in the figure, θ(NH··O) = 142° and θ(CO··H) = 104°, making for a moderately distorted H-bond, with a distance of 2.098 Å separating the H and O. Because the intramolecular H-bond involves a total of seven atoms in a cyclic arrangement, this structure is commonly referred to as C_7. This H-bond is partly responsible for a certain amount of nonplanarity within the two peptide units, visible in the view of this system on the right side of Figure 1. The largest deviation involves the bridging H atom, which is twisted such that φ(HNCC) = −10°. The terminal NH_2 is puckered by 18°, with φ(HNCH) = 162°. The peptide group containing the proton-accepting O is twisted by 5° so as to strengthen the H-bond. The specific structure presented in Figure 1 arises from MP2/6-31+G** optimization, but much the same geometry was obtained by optimization at the HF/6-31G*, HF/6-31+G**, and B3LYP/6-31+G** levels.

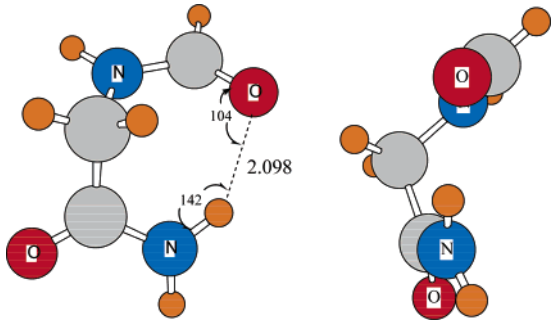

Figure 1. Two views of the Gly dipeptide C_7 structure optimized at the MP2/6-31+G** level. Interatomic distance in angstroms, angles in deg.

The formamide molecule was taken as the simplest model system that contains a peptide-like amide unit. A molecule of $HCONH_2$ was hence placed in a position where its O atom could accept a proton from the NH of the dipeptide. It was hoped that an optimization would lead to a single NH··O bond, and thus a means of estimating the energy of this interaction. However, upon full optimization, this configuration collapsed to a structure like that illustrated in Figure 2a, which contains a pair of NH··O interactions. The second of these (the upper one in Figure 2a) is undesired, first because it involves the NH hydrogen atom which is cis to the O of formamide, and which would normally be replaced in a second peptide by a C atom. Second, this additional H-bond causes a severe angular distortion of the first, desired bond (in the lower part of Figure 2a), bending θ(NH··O) a full 30° from its preferred value of 180°.

To prevent the formation of the upper H-bond, a second optimization was attempted in which the NH··O angle of the lower bond was set equal to, and frozen at, 180°. While this strategy did indeed prevent the formation of the upper NH··O bond, the formamide molecule nevertheless twisted around to form a second, undesired H-bond. The secondary H-bond involved the O atom of the dipeptide as acceptor, but the donor NH_2 of formamide was replaced by the aldehydic H atom. While the upper H-bond may in fact be weaker than the lower, the total interaction energy would nonetheless provide a poor estimate of the strength of the NH··O bond, due to this contamination.

The situation is worse in terms of assessing the strength of a CH··O bond. When the O atom of the formamide was positioned

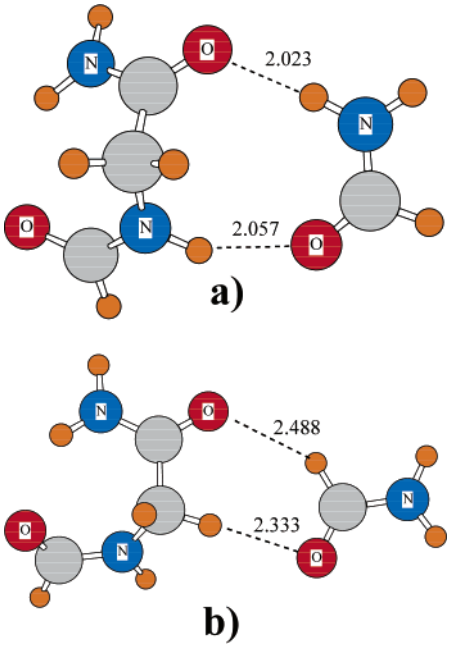

Figure 2. (a) Result of full optimization of Gly dipeptide + formamide. (b) Optimization of CH··O interaction, with θ(CH··O) frozen at 180°. Interatomic distances in angstroms.

such that θ(CH··O) was frozen at 180°, the aldehydic proton of the formamide swung into position to form a secondary H-bond with the peptide O, as illustrated in Figure 2b. It is clear then that a single restriction of either θ(NH··O) or θ(CH··O) to 180° is insufficient to isolate a single H-bond.

It was thought that the situation might perhaps be alleviated by removing the "hot" aldehydic proton from the formamide. This was done by replacing each of three H atoms of formamide by a methyl group, resulting in *N*,*N*-dimethylacetamide (DMA) as the proton acceptor. While this substitution resulted in some improvement, it was still not fully satisfactory. When the O atom of DMA was placed in proximity to the NH proton of the dipeptide (similar to Figure 2a), the H atoms of the methyl group of DMA were found to interact with the dipeptide O atom, in much the same way as does the NH proton in Figure 2a, forming a secondary CH··O H-bond. When the DMA O atom was placed near one of the CH protons of the dipeptide, again a methyl H formed a secondary H-bond with the dipeptide O.

Because the imposition of a single geometry restriction (θ(NH··O) or θ(CH··O) = 180°) was insufficient to limit the interaction to a single H-bond, a second restriction was added. To prevent the formamide proton acceptor molecule from "pivoting" around its O atom, the C═O···H angle was fixed at 180°.[68] This double restriction indeed had the desired effect. The geometries optimized under this dual restriction at the MP2/6-31+G** level are illustrated in Figure 3, along with the distances between the bridging H and the proton-accepting O atom. Not surprisingly, the NH··O interaction in Figure 3a is the shortest, with R(H··O) = 1.925 Å. There are two $C^{\alpha}H$ protons that can form a H-bond. The H that lies approximately in the plane of the remainder of the dipeptide molecule is labeled H^1, while H^2 refers to the other CH, which is pointed out of that plane to a large extent. These two CH groups form H-bonds with formamide, as pictured in Figure 3b and c. Comparison illustrates that the H-bond involving H^2 is slightly shorter than the other by 0.02 Å, but both CH··O bonds are some 0.3 Å longer than the NH··O bond in (a).

As in the earlier situations, the results depicted in Figure 3 are associated with MP2/6-31+G** calculations; results at lower

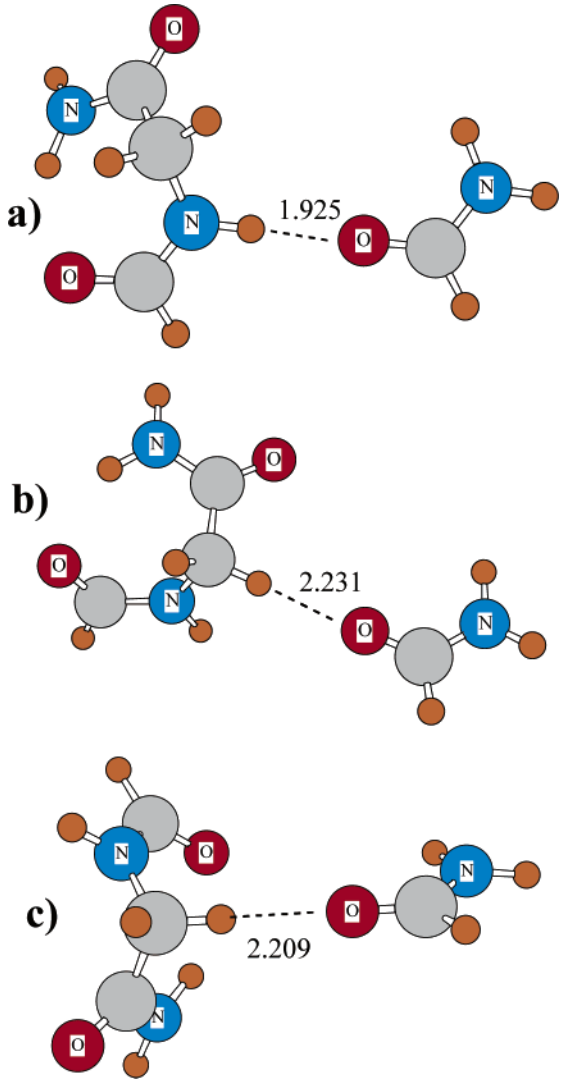

Figure 3. MP2/6-31+G** optimized geometries of the (a) NH··O and (b) and (c) two CH··O interactions formed between the Gly dipeptide and formamide. Double linearity was imposed on each structure, with θ(XH··O) and θ(H··OC) both held at 180°. Interatomic distance in angstroms, angles in deg.

TABLE 1: Intermolecular H-Bond Lengths and Interaction Energies

	NH··O	CH^1··O	CH^2··O
	R(H··O), Å		
MP2/6-31+G**	1.925	2.231	2.209
B3LYP/6-31+G**	1.941	2.309	2.276
HF/6-31+G**	2.030	2.386	2.376
HF/6-31G*	2.003	2.374	2.342
	ΔE,[a] kcal/mol		
MP2/6-31+G**	−7.44	−2.81	−2.25
B3LYP/6-31+G**	−7.17	−2.35	−1.72
HF/6-31+G**	−7.15	−2.21	−1.61
HF/6-31G*	−7.00	−1.99	−1.34

[a] Including counterpoise correction.

levels follow identical trends, as illustrated in Table 1. In general, the H-bonds tend to be somewhat shortened by the inclusion of electron correlation, more so with MP2 than with B3LYP, and the bonds are elongated by the extra basis functions of 6-31+G** as compared to 6-31G*. The interaction energies, after suitable counterpoise correction of basis set superposition error, are reported in the lower half of Table 1. The NH··O interaction energy is slightly in excess of 7 kcal/mol, as

compared to the two CH··O bonds, which are between 2 and 3 kcal/mol, with the highest level of theory. (Negative values of ΔE correspond to attractive interactions.) Comparison of the various rows in Table 1 indicates that enlargement of basis set slightly increases each H-bond energy, as does inclusion of correlation; again, MP2 has a greater effect than does B3LYP.

The formation of the H-bond has only very limited effects upon the preferred conformation of the dipeptide. For example, the NH··O bond of Figure 3a changes φ and ψ by 0.3° and 2.4°, respectively, from their optimized values in the absence of a proton acceptor. These angles remain constant to within 1.5° and 1.2° for the two CH··O H-bonds. The internal H-bond within the C_7 structure remains largely unchanged as well, undergoing a slight contraction. The NH··O interaction shrinks this internal bond from 2.098 Å in the uncomplexed dipeptide to 2.055 Å, and the H··O distance is reduced even less for the two different CH··O interactions: 2.074 and 2.078 Å. These internal H-bond contractions are indicative of a certain degree of cooperativity.

Immersion in Solvent. The next issue considered dealt with the placement of each of these complexes within the context of a surrounding medium. The first medium chosen was water ($\epsilon = 78.4$), due to its obvious prevalence in biological systems. The second medium considered was diethyl ether, primarily because its dielectric constant of 4.3 closely mimics the value generally considered to most closely reproduce the environment within a protein.[67,69−73] The geometry of each subunit was reoptimized within each solvent. The same was true of each complex, subject once again to the same two restraints imposing linearity upon the X−H···O═C linkage, as were employed in the gas-phase calculations above.

When placed in the two solvents, the dipeptide retains its same basic C_7 geometry, with only small changes. For example, at the B3LYP/6-31+G** level, the length of the internal H-bond that completes the cycle, that is, r(O··H), is equal to 2.096 Å in the gas phase, shortens to 2.063 Å in ether, and lengthens to 2.087 Å in water. Without correlation effects included, these changes are slightly larger in magnitude. (In fact, in one case, the 6-31+G** basis set (but not 6-31G*), the lengthening of the internal H-bond is just strong enough to actually break this H-bond entirely.) The φ angle remains constant at −81° throughout, but ψ diminishes from 65° in a vacuum, to 60° in ether, and then to 55° in water. The geometry of the other subunit, formamide, is much less sensitive to the polarity of the solvent.

Results for the complexes are reported at the correlated B3LYP/6-31+G** level in Table 2. The upper set of data indicates a distinction between the NH··O and CH··O bonds. Whereas the increasing polarity of the solvent has only a small effect upon the length of the former, the latter sort of bond elongates very quickly as ϵ rises. In fact, in water the complex falls apart, as the CH··O bond is not strong enough to hold it together in this environment. The energetics in the second portion of Table 2 reinforce this behavior. The increase in dielectric constant rapidly weakens both types of H-bonds. Whereas the NH··O interaction remains stabilizing ($\Delta E < 0$) when $\epsilon = 4.3$, the two CH··O H-bonds have a positive value of ΔE, indicating the solvated complex is less stable than the pair of individually solvated subunits. When ϵ has risen to 78, both NH··O and CH··O alike are destabilizing.

In terms of geometry, the NH··O interaction produces only very minor changes into the preferred structure of the dipeptide. Regardless of whether in vacuo, or in one of the two solvents, the (φ,ψ) angles are changed by no more than 2° when the formamide proton acceptor approaches. The intramolecular H-bond is shortened by a very small amount, which varies from 0.04 Å in vacuo to 0.01 Å in water. The CH··O interaction, too, induces only minimal changes in the dipeptide geometry. The internal H-bond is contracted by 0.04 Å in vacuo, but only 0.004 Å in ether; (φ,ψ) remains constant within 1°.

TABLE 2: Intermolecular H-Bond Lengths and Interaction Energies Computed at the B3LYP/6-31+G Level for Various Dielectric Constants**

ϵ	NH··O	CH^1··O	CH^2··O
		R(H··O), Å	
1	1.941	2.309	2.276
4.3	1.908	2.384	2.310
78	1.911	dissoc[a]	dissoc[a]
		ΔE,[b] kcal/mol	
1	−7.50	−2.55	−1.99
4.3	−2.01	+1.81	+2.75
78	+1.52		−[c]
		ΔE_{unopt},[d] kcal/mol	
1	−7.50	−2.55	−1.99
4.3	−1.93	+1.98	+2.89
78	+1.62	+4.14	+5.39

[a] Complex dissociates. [b] Geometry optimized in solvent; no counterpoise correction. [c] No optimization. [d] Gas-phase geometry; no counterpoise correction.

TABLE 3: Solvation Energies, kcal/mol, of Separated Subunits, and Each Complex, for Two Dielectric Constants

ϵ	separated subunits	NH··O	CH^1··O	CH^2··O
		B3LYP/6-31+G**		
4.3	−14.2	−8.7	−9.7	−9.4
78	−27.7	−18.6	−21.0	−20.3
		HF/6-31+G**		
4.3	−15.7	−9.7	−11.0	−10.4
78	−29.4	−20.0	−22.9	−22.0
		HF/6-31G*		
4.3	−13.8	−8.2	−9.5	−8.9
78	−25.8	−17.1	−19.9	−19.0

The lowest section of data in Table 2 reports the interaction energies, but using the geometries of the gas-phase species, that is, unoptimized within the solvents. Comparison with the fully optimized data in the preceding section indicates that the reoptimization within each solvent produces only a marginal effect on H-bond energies and that the positive values coincide with dissociation whether the geometry is reoptimized or not.

The source of the H-bond weakening is apparent in Table 3, which lists the solvation energies of each complex (all negative, and therefore stabilizing), along with that computed for the separate subunits in the first column of data. One may note first that the solvation energy is not highly susceptible to the details of the basis set, nor to inclusion of electron correlation. The solvation energy of any species grows quickly, in fact approximately doubles, as the dielectric constant rises from 4 to 78. This increase is in fact quantitatively consonant with the rise in the Onsager function[53,54] $F_O = (\epsilon - 1)/(\epsilon + 2)$, which is itself magnified by a factor of 1.8 when ϵ increases from 4.3 to 78. Probably more importantly, the solvation energy of any of the H-bonded complexes is substantially less than the combined value of the pair of separated subunits. It is this lesser stabilization of the complex that results in the diminished H-bond energies so evident in Table 2.

It may be noted finally that the solvation energies of the two CH··O complexes are more negative by 1 or 2 kcal/mol than that of the NH··O system. The most likely source of this

discrepancy has to do with the H-bonding groups that are available to interact with solvent. The CH··O complexes leave the polar NH group of the dipeptide unoccupied, and hence available to interact with the polar medium, whereas this group is already tied up in the NH··O complex, leaving only the less polar $C^{\alpha}H$ bonds.

In summary, both NH··O and CH··O sorts of bonds are rapidly destabilized by growing solvent dielectric, as was noted in our earlier work dealing with simpler fluoromethanes rather than a dipeptide model.[60] Also, as found in the earlier calculations, the conventional sort of H-bond remains more stabilizing, regardless of the particular solvent chosen.

Our earlier set of computations[60] of fluoromethanes had suggested that the CH··O bonds that they form can be comparable in strength to the conventional OH··O bond of a water dimer in a particular context. One may consider the energetic contribution of a given H-bond to the protein folding process to begin with the pair of subunits separated from one another, both immersed in aqueous solvent. As the protein folds, the H-bond forms between these two subunits, and they are removed from water and placed within the interior of the protein, with a much lower dielectric constant. This process may be characterized as

$$\mathrm{AH(aq) + B(aq) \rightarrow AH\cdots B(protein)} \qquad (1)$$

This process can be broken down conceptually into two consecutive steps as follows:

$$\mathrm{AH(aq) + B(aq) \rightarrow AH\cdots B(aq)} \qquad (2)$$

$$\mathrm{AH\cdots B(aq) \rightarrow AH\cdots B(protein)} \qquad (3)$$

Reaction 2 refers to the formation of the H-bond in aqueous medium, which is followed by removal of the H-bonded complex from water and into the much less polar interior of the protein. Indeed, reaction 3 is the source of the so-called "desolvation penalty", which causes H-bonds to play a weaker energetic role in protein folding than might be expected based solely on their strength in less polar media. In the case of the fluoromethanes, the calculations[60] had indicated that while their CH··O bonds with a water proton−acceptor molecule are weaker than the conventional OH··O bond of a water dimer within the context of aqueous medium (reaction 2), the desolvation penalty (reaction 3) is quite a bit smaller for the former bonds than for the latter. Consequently, the combination of these two processes, reaction 1, can be more favorable for CH··O than for OH··O.

A principal source of concern about the prior results had to do with the modeling of solvation of not only the H-bonded complexes, but of the isolated subunits as well. That is, implicit in evaluation of the energetics of reaction 2 is an assessment of the hydration energy of each of the subunits AH and B, as well as the AH··B complex. The subunits are fundamentally different for the CH··O and OH··O cases: CF_nH_{4-n} and OH_2 in one case, and HOH and OH_2 in the other. The energetics of reaction 2 could hence be thrown off if the modeling of solvation of CF_nH_{4-n} is of a different level of accuracy than is the modeling of HOH. This error would persist when reaction 2 is combined with reaction 3, to yield the final result.

There is no such error in the current model that compares CH··O with a conventional NH··O bond, for the simple reason that the isolated monomers are identical in either case: $CHONHCH_2CONH_2 + HCONH_2$. Any errors in the solvation model of these two molecules largely cancel when comparing their two different modes of combining with one another, either CH··O or NH··O. The situation is improved also with regard to the H-bonded complexes. Whereas the $CF_nH_{4-n}\cdots OH_2$ complex is in some ways fundamentally distinct from a $HOH\cdots OH_2$ complex, and the two may hence be treated differently by a solvation model, the CH··O and NH··O complexes that are under consideration here are very similar indeed, differing only in minor aspects of their geometry (see Figure 3). It is hence concluded that the findings in the present models of H-bond formation in different media are far less subject to inaccuracies introduced by the particular means of including solvation phenomena.

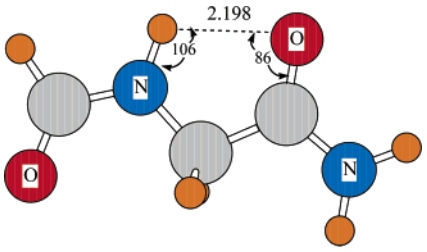

Figure 4. MP2/6-31+G** geometry of C_5 structure, with $\varphi = -170$, $\psi = -172$. These two angles are both 180° at lower levels of theory. Interatomic distance in angstroms, angles in deg.

TABLE 4: Energies, kcal/mol, Computed for Reaction 3, To Bring Each Complex from Water ($\epsilon = 78$) into Protein Interior ($\epsilon = 4$)

	NH··O	CH^1··O	CH^2··O
B3LYP/6-31+G**	9.9	11.3	11.0
HF/6-31+G**	10.4	12.0	11.6
HF/6-31G*	8.9	10.4	10.1

The values computed for the "solvation penalty", reaction 3, are reported in Table 4 at various levels of theory. There is fair agreement from one level of theory to the next, and one can note that enlargement of basis set raises this quantity, which is then lowered when correlation is included. Yet most importantly, all levels of theory support the idea that the desolvation penalty is somewhat larger for the CH··O complexes than for NH··O, opposite from the conclusions reached earlier for the substituted methanes. Hence, one cannot expect a weaker H-bond in any particular medium, be it aqueous or a vacuum, to be compensated by the desolvation penalty. Indeed, the latter property acts to reinforce the weaker nature of the CH··O bond, as compared to NH··O, such that the former will contribute less energetically to protein folding than will the latter.

Alternate Geometries of Dipeptide. As reported above, the C_7 type structure, reminiscent of a γ turn in a polypeptide, is the most stable for the dipeptide. However, there is an alternate structure, containing an intramolecular H-bond within the context of a five-membered cycle, and thus denoted C_5, that is competitive in energy,[65−67,74−77] and may even be preferred[64] in solution. In fact, this geometry has recently been witnessed in the gas phase[62] as a part of the β-sheet structure adopted[78] by certain small polypeptides. As may be seen in Figure 4, this structure is rather extended, with $(\varphi,\psi) = (-170, -172)$, at the MP2/6-31+G** level. The internal H-bond is distorted from its optimum parameters, with a θ(NH··O) angle of 106°, and with θ(CO··H) = 86°, and so must be considered as rather weak. Indeed, at this same level of theory, this C_5 structure is higher in energy than the global C_7 minimum by 1.3 kcal/mol. At lower levels, the two configurations are closer in energy: C_7 is preferred by only 0.1 kcal/mol at B3LYP/6-31+G**. The order reverses if correlation is not included with C_5 becoming the more stable by 0.6 kcal/mol with the 6-31+G** basis set, and by 0.5 kcal/mol with the smaller 6-31G*.

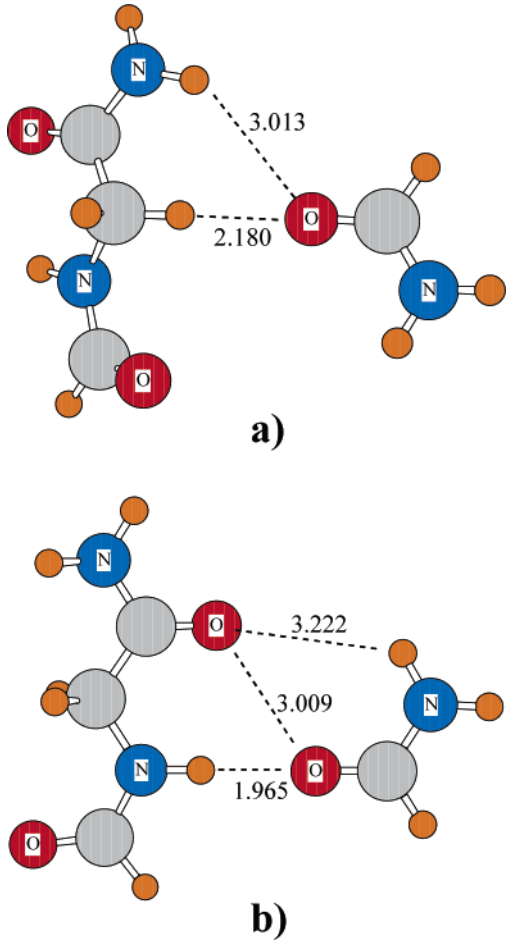

Figure 5. MP2/6-31+G** geometry of complexes formed between C_5 structure of dipeptide and formamide. CH··O and NH··O geometries are illustrated in (a) and (b), respectively.

Other structures were sought that might represent minima. An alternate C_5 type of geometry was considered, differing from that pictured in Figure 4 by a 180° rotation around the N−C bond on the left side. However, this structure was more than 4 kcal/mol higher in energy than the standard C_5 in Figure 4. Another starting point considered begins with a ∼90° rotation around the C−C bond, that is, $\psi = -73$. However, this configuration collapsed to the C_7 structure upon optimization. An attempt to locate a α_R structure began an optimization from ($\varphi = -63, \psi = -46$), but this collapsed to the C_7 geometry, as did a search for a potential α_L local minimum that began with ($\varphi = 57, \psi = 34$). Searches for β and β_2 structures also failed: A starting point of ($\varphi = -63, \psi = 134$) collapsed to C_7 as did ($\varphi = -120, \psi = 20$). The identification of only these two particular minima, C_7 and C_5, on the potential energy surface of a glycine dipeptide is fully consistent with prior[75] computations.

A central question thus arises as to how the energetics of H-bond formation reported above for the C_7 structure might be altered in the case of C_5 geometry, given that the latter appears to be competitive in energy with the former, and also represents a fair approximation to the β-sheet conformation. For the sake of consistency, a formamide molecule was again brought up toward the relevant NH and CH protons and held in the linear arrangement used before, that is, θ(XH··O) = θ(H··OC) = 180°. This sort of prescription led to the CH··O bond shown in Figure 5a. There is a secondary H-bond formed with the upper NH_2 group, but this is rather long (3.013 Å) and bent, with θ(NH··O) = 139°, and so is apt to be vanishingly weak. (Nearly identical results were obtained when considering the other CH proton to serve as the bridge, due to the near planar symmetry of the C_5 configuration.) Despite the rather weak nature of this CH··O bond, it does induce certain conformational changes within the dipeptide. For example, the (φ,ψ) angles of (−170,−172) in the isolated dipeptide are altered to (+177,−157) in the complex, changes in the 13°−17° range. Another apparent change has to do with the internal H-bond: the O··H distance elongates from 2.198 to 2.276 Å.

TABLE 5: Intermolecular Energies,[a] kcal/mol, for Complexes Containing the C_5 Structure of the Dipeptide; H-Bond Geometries Held Linear[b]

	NH··O	CH^1··O	CH^2··O
MP2/6-31+G**	−2.5	−3.8	−3.8
B3LYP/6-31+G**	−1.6	−2.9	−2.9
HF/6-31+G**	−1.5	−2.8	−2.8
HF/6-31G*	−1.3	−2.5	−2.5

[a] Including counterpoise correction. [b] θ(XH··O) = θ(H··OC) = 180°.

When this same set of angular constraints, that is, θ(XH··O) = θ(H··OC) = 180°, was imposed on the complex, the approach of the formamide toward the NH group led to an apparently stronger perturbation, sufficient in fact to induce a transformation of the dipeptide from its C_5 structure back to the more stable C_7 arrangement. To guarantee the dipeptide remain in its C_5 configuration, while introducing a minimum of geometric restraints, the φ and ψ angles were held in the values that obtain in the optimized C_5 monomer (and the terminal OCNC angle was held equal to 180°). The optimized structure contains the NH··O bond illustrated in Figure 5b. In addition to the desired NH··O bond, there is a secondary one present, but again this bond is much longer at 3.222 Å and suffers from angular distortion with θ(NH··O) 53° from linearity. In either the CH··O or the NH··O case, the intermolecular complex is apt to represent a realistic assessment of the XH··O bond of interest.

Comparison of the three columns of data in Table 5 reveals the very interesting conclusion that the CH··O bonds are stronger than the NH··O bonds. That is, the expected order of NH··O being the stronger of the two bonds, exemplified by the C_7 geometries, is reversed when the bonds are formed to the same dipeptide, but in its C_5 configuration. There are several possibilities for the origin of this reversal. First is the presence of a secondary NH··O bond in Figure 5a, which, although rather long, may offer additional stabilization to the system. The NH··O system in Figure 5b may be destabilized to some extent by the fairly close approach of the two O atoms (3.009 Å at the MP2 level). This destabilizing effect is counterbalanced by the secondary NH··O bond, which is rather long and hence likely weak, at R(H··O) = 3.222 Å, but may offer some stabilization nonetheless. (An alternate geometry was considered, related to Figure 5b by a 180° rotation of the formamide around the intermolecular NH··O axis, which would rotate the formamide NH_2 group down, away from the C═O. However, this structure is slightly higher in energy than that pictured in Figure 5b, which would amount to a further weakening of the NH··O H-bond. Moreover, upon optimization, this configuration relaxes back to (b).) As still another consideration, the formation of the intermolecular NH··O bond in Figure 5b competes to some extent with the intramolecular NH··O of the C_5 dipeptide ring. The H··O distance elongates from 2.20 Å in the uncomplexed C_5 dipeptide to 2.49 Å in Figure 5b. On the other hand, given the strong angular distortions of this internal H-bond (see Figure 4), its stretch is not anticipated to cost much energy. In any case, these various factors are not particular to the models

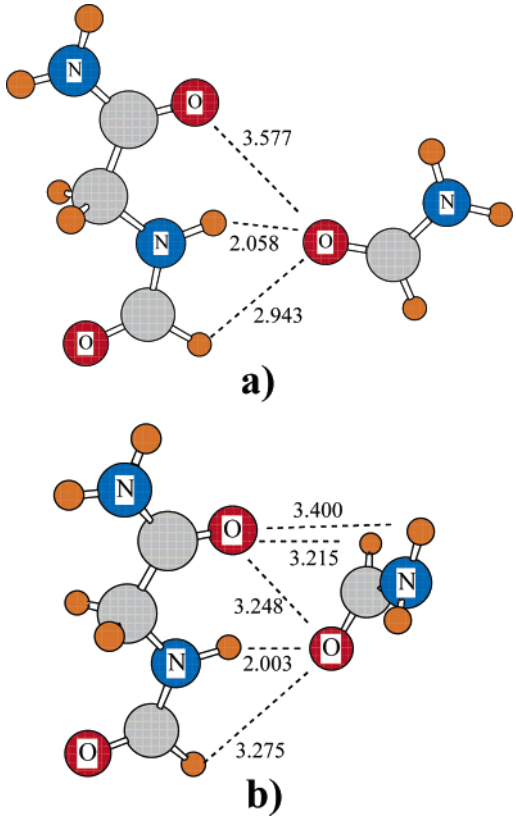

Figure 6. MP2/6-31+G** geometry of complex formed between C_5 geometry of dipeptide and formamide, (a) with removal of NH··O linearity restriction, and (b) with φ(NCO··O) = 90°.

TABLE 6: Intermolecular Energies,[a] kcal/mol, for Complexes Containing the C_5 Structure of the Dipeptide, with Restrictions from Linearity Lifted

	NH··O		CH··O
	θ(NH··O) relaxed	φ(NCO··O) = 90°	θ(CO··H) relaxed
MP2/6-31+G**	−3.9	−5.1	−4.3
B3LYP/6-31+G**	−3.3	−4.1	−2.6
HF/6-31+G**	−3.7	−4.7	−2.4
HF/6-31G*	−3.5	−4.5	−2.3

[a] Including counterpoise correction.

chosen, but would be present in the full polypeptides themselves, and so ought to be considered as real considerations.

Nonetheless, to try to reduce any O··O repulsion that might be present, which may be destabilizing the NH··O configuration, the two oxygens in Figure 5b were permitted to move away from one another by removing the restriction that θ(NH··O) must be 180°. Indeed, this new freedom allowed the formamide O to move down, away from the carbonyl O of the dipeptide, and toward the CH. As illustrated in Figure 6a, this motion did in fact elongate the O··O distance by some 0.57 Å, but only stretched the H··O distance of the H-bond by 0.09 Å. The θ-(NH··O) angle was lowered from 180° to 148°, another factor that might slightly weaken this H-bond. The motion does bring the formamide's O atom within about 2.9 Å of the aldehydic CH, which may introduce a secondary stabilizing force. The first column of Table 6 illustrates that this additional freedom brings the NH··O interaction energy up closer to the CH··O value, surpassing the latter to a small extent.

Still another means of assessing the strength of the NH··O interaction lies in removal of the second constraint on the H-bond, that is, lifting of the restriction that the CO··H angle

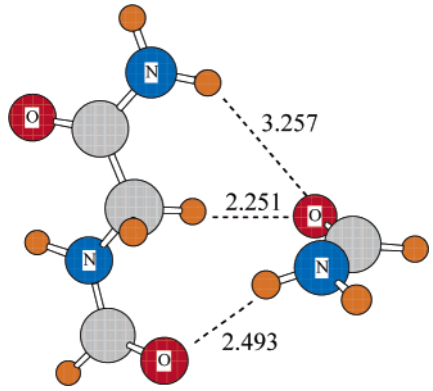

Figure 7. MP2/6-31+G** geometry of CH··O complex formed between C_5 geometry of dipeptide and formamide with removal of CO··H linearity restriction, but keeping formamide plane coincident with the $C^{\alpha}H_2$ plane of the dipeptide.

be maintained at 180°. However, it is clear that this would enable the NH_2 group of the formamide to rotate in toward the carbonyl O of the dipeptide to form an undesirable second NH··O bond. This possibility was excluded by insisting that the formamide plane (characterized by its N,C,O atoms) be held orthogonal to the dipeptide O atom, that is, φ(NCO··O) = 90°. The optimized configuration is presented in Figure 6b, again at the MP2/6-31+G** level. This structure weakens the NH··O bond in the sense that the bridging H atom lies well out of the carbonyl O atom's plane, φ(H··O=CN) = 87°. There is a set of potentially stabilizing interactions, which may act to inflate the interaction energy, although all of these separate the carbonyl O from the bridging H atoms by more than 3 Å. As indicated by the second column of Table 6, this alternate arrangement adds roughly 1 kcal/mol to the interaction energies exemplified by Figure 6a, wherein the C=O···H arrangement was held to linearity.

Efforts were made to allow analogous sorts of relaxation into the CH··O C_5 structure pictured in Figure 5a. The CH··O angle was held at 180° because it was obvious that removal of this restriction would permit the carbonyl O of the formamide to move away from the CH group and up toward the dipeptide NH_2 group. Likewise, if the C=O··H angle were relaxed, the NH_2 group of the formamide would easily pivot in toward the dipeptide O atom, forming a second strong NH··O H-bond. To hinder such reorientation, the C=O··H angle was unfrozen, but the formamide plane was forced to remain perpendicular to the dipeptide O atom, that is, φ(NCO··O) = 90°. Even with this restriction, the formamide NH_2 shifted so as to form a secondary H-bond, albeit a geometrically distorted one. The same was true if the dihedral angle were held to −90°. Constraining this dihedral angle to 180° did prevent formation of the undesired NH··O bond, but the structure formed a secondary H-bond nonetheless: in this case, it was the aldehydic H of the formamide that approached the carbonyl O of the dipeptide.

A more successful approach indeed lifted the linearity restriction on the CO··H configuration, but did so by insisting that the plane of the formamide proton acceptor molecule be coincident with the $C^{\alpha}H_2$ group plane, that is, that the dihedral angle of the N of formamide and the H of the $C^{\alpha}H_2$ group not involved in the H-bond, around the formamide C=O bond, be 0°. The geometry optimized at the MP2/6-31+G** level, illustrated in Figure 7, exhibits a θ(CO··H) angle of 105°. It also contains two NH··O bonds, albeit rather weak ones. The NH··O distance of the shorter one is 2.493 Å; however, the bond is expected to be weakened by a high degree of nonlinearity, θ(NH··O) = 148°. The longer of the two (3.257 Å) is even more distorted, with θ(NH··O) = 135°.

Comparison of the last columns of Tables 5 and 6 indicates that this configuration strengthens the intermolecular interaction energy, but only by a little. At the MP2/6-31+G** level, ΔE is −4.3 kcal/mol, as compared to −3.8 kcal/mol for the fully linear structure in Figure 5a. Nonetheless, comparison of Tables 5 and 6 provides some estimates as to how sensitive the energetics are, for either NH··O or CH··O, to differing geometric restrictions.

Summary and Discussion

This work has considered the ability of the CH and NH groups of a polypeptide chain to engage in H-bonds with an acceptor atom. The models represent realistic reproductions of the actual systems that would participate within a polypeptide: The CH and NH donors are surrounded by full peptide groups, and the O acceptor is part of an amide function. It is clear that both NH and CH can indeed participate in stabilizing H-bonds, but the question as to which is a stronger bond is not as simple as one might hope. The relative strengths appear to depend first upon the particular geometry adopted by the dipeptide donor. A second factor has to do with the particular orientation of the approaching acceptor, and what other atoms it may be able to interact with.

The two primary minima in the potential energy surface of a dipeptide may be categorized by the number of atoms that constitute a ring so as to enable formation of an internal, albeit weak, H-bond. The C_7 geometry is slightly more stable than C_5, although this issue can be reversed at certain levels of theory. Hence, both of these structures were considered in the analysis of intermolecular H-bonding to an external proton acceptor. In the case of the C_7 structure, more stable at the highest level of theory, the NH··O bond formed with an approaching carbonyl is clearly stronger than its parallel CH··O. The interaction energy in the gas phase is computed to be 7.4 kcal/mol, as compared to 2.8 kcal/mol for the CH··O analogue. These values are surprisingly stable to variations in level of theory, lowering slightly if electron correlation is neglected or if the basis set is made somewhat smaller. Along with its greater energetic strength, the NH··O bond is a shorter one, with R(H··O) = 1.93 Å, as compared to 2.23 Å for CH··O.

These same systems were immersed in a series of model solvents, each characterized by a different dielectric constant. Aqueous solvation was simulated by $\epsilon = 78$, and the interior of a protein by $\epsilon = 4$. The rising dielectric constant causes a weakening of both the NH··O and the CH··O bonds, but in a uniform manner such that the former remains stronger than the latter regardless of ϵ. The weaker nature of the CH··O bond causes it to rupture when ϵ reaches 78, whereas NH··O remains connected, albeit only marginally so. This H-bond weakening can be attributed to the lesser solvation energy of each complex, as compared to the same quantity for the pair of uncomplexed subunits.

An earlier work[60] had considered the contribution of any H-bond to protein folding as the energy of a process that begins with the proton donor and acceptor, separated in aqueous solvent, and takes them into the protein interior where they form a H-bond. The calculations had suggested that, within this framework, a CH··O bond may contribute more toward protein stability than a conventional H-bond, due in large measure to the greater desolvation penalty of the latter. However, the crudeness of the models, using di- and trifluoromethane as a CH donor rather than a peptide, precluded forming a definitive conclusion. The present work provides a more authoritative answer to this question, in particular because the model is far more appropriate, using a real dipeptide as donor. Equally important, because the isolated partners are the same in both the NH··O and the CH··O cases, that is, dipeptide + formamide, any errors incurred in computing the solvation energies of the isolated subunits completely cancel. In this more accurate assessment of energetics of protein folding, it was found that the desolvation penalties of the NH··O and CH··O systems are quite similar; indeed, this quantity is slightly larger for the CH··O bond. As a consequence, the stronger nature of the NH··O bond, as encountered not only in the gas phase, but also in each of several solvents, remains true also even if one considers energetic contributions to protein folding.

Whereas one may safely conclude that the NH donor of a dipeptide in its C_7 geometry will engage in a stronger H-bond than will its CH donor, this conclusion may not be extended to the C_5 structure. Not only is C_5 competitive in energy to C_7, but the former is very commonly observed within the context of the β-sheets of proteins, and so is of great importance. Indeed, the situation is reversed here: the CH··O H-bond energy is increased from its C_7 value of 2.8 to 3.8 kcal/mol for C_5, while the NH··O bond is weakened from 7.5 to 2.5 kcal/mol, making the latter weaker than the former.

This very substantial bond weakening may be due in part to the approach, to within about 3.0 Å of one another, of two carbonyl O atoms, that of the proton acceptor, and that of the dipeptide. When this distance was permitted to elongate, by allowing nonlinearity of the NH··O configuration, this interaction energy increased slightly to almost equalize with the CH··O energy. On the other hand, this new geometry includes a second, weaker CH··O bond, which contributes to this greater stability. An alternate means of retaining the NH··O bond but at the same time permitting the two carbonyl oxygens to move away from one another resulted in a further enhancement of the stabilization energy. However, this configuration may be contaminated by two or three other H-bonds, which inflate this quantity. Analogous relaxation of certain geometry restrictions in the CH··O case leads to a moderate strengthening of this interaction as well.

In summary, it is clear that the NH··O H-bond is stronger than CH··O when the dipeptide is in its C_7 configuration, regardless of medium. However, the situation is less clear-cut in the C_5, or extended chain structure, wherein the CH··O appears to be slightly more stabilizing than is NH··O.

This analysis has modeled the protein interior by a dielectric constant of 4, a value chosen on the basis of a number of protein studies[69−73] over the years. This particular value is not accepted by all, with some arguing that a larger ϵ might be more appropriate, at least for certain regions of proteins.[79,80] Fortunately, our earlier calculations had demonstrated that various solvation properties are related in a very nearly fashion to the Onsager function, $F_O = (\epsilon - 1)/(\epsilon + 2)$, that in turn relates ϵ to the properties of the solvent.[53,54] It is therefore possible to estimate rather closely what the solvation energies would be in any arbitrary solvent with dielectric constant ϵ.

As an example, if ϵ were set equal to 10 instead of 4, the solvation energies for the separated subunits, in the B3LYP/6-31+G** context, would be approximately −21.2 kcal/mol. Likewise, the solvation energies of the NH··O and CH··O complexes would be respectively −13.8 and −15.5 kcal/mol. When these values replace the quantities estimated previously with a protein interior of $\epsilon = 4$, the B3LYP/6-31+G** values for the energy required to bring each of the two complexes from water into protein (reaction 3) would change from 9.9 and 11.3 kcal/mol, respectively, in the first row of Table 4, to 4.8 and

5.5 kcal/mol. This quantity, representing the desolvation penalty, remains smaller for NH··O than for CH··O, as it was for $\epsilon = 4$. In a similar vein, if the dielectric constant of the protein is taken as the even higher value of 20, the desolvation penalty of the NH··O complex (2.2 kcal/mol) remains smaller than for CH··O (2.5 kcal), with the difference now equal to 0.3 kcal/mol. In summary, raising the dielectric constant of the protein reduces the difference in desolvation penalty between NH··O and CH··O, but clearly does not reverse the conclusion that the former is smaller than is the latter, regardless of ϵ.

Of course, a homogeneous continuum serves as an inexact model for a solvent or protein, with its discrete solvent molecules or neighboring groups, and the lack of explicit dynamic time dependence in these methods. Nonetheless, continuum models have achieved what is now a long history of successful applications to a range of chemical problems,[81−83] and even specifically to CH··X bonds[84−86] and to peptides.[67,87,88] A group has recently found evidence that many of the same principles governing conformation that apply in the gas phase are applicable also in solution,[89] and advanced the argument that "the gas phase should be viewed as an environment that represents an extreme on the hydrophobicity/dielectric constant scale, rather than an environment that is completely alien to chemical biology." Moreover, our own earlier work[60] had considered the explicit inclusion of a first hydration cage, and we found only small changes from the results derived from a pure continuum. In any case, it is reiterated that the direct comparison of the CH··O interaction to the conventional NH··O bond, within the same molecular framework, and using the exact same solvation models for both, is expected to greatly reduce any errors that might be incurred by any theoretical imprecision.

It was mentioned above that the approach of the proton acceptor toward the NH group of the C_5 configuration of the dipeptide tended to push this configuration toward the C_7 geometry. This finding has interesting implications for protein folding, suggesting as it does that the incipient formation of an interpeptide H-bond may be capable of inducing a rather major structural change within the polypeptide backbone. After all, the $C_5 \rightarrow C_7$ transition represents a change of roughly 100° in both φ and ψ.

Earlier calculations[90] had considered the interaction energy for the CH··O H-bond connecting a Gly amino acid with a proton acceptor. This model system differs from the current work first in that the Gly amino acid surrounds the CH_2 group of interest by the small COOH and NH_2 groups, rather than the full peptide groups in the current calculations, which are more representative of a polypeptide chain. A second difference resides in the nature of the proton acceptor: the earlier OH_2 molecule has been replaced by the carbonyl O of a full amide group, again a better model for interpeptide H-bonding. The value computed in the more primitive model for the CH··O H-bond energy in the gas phase at the MP2/6-31+G** level[90] was 2.5 kcal/mol, which may be compared with the gas-phase quantities reported herein at the same level, which vary from 2.3 and 2.8 kcal/mol for the two CH protons of the C_7 conformer of the dipeptide, up to 3.8 kcal/mol for the C_5 geometry.

There have been earlier attempts to estimate the CH··O bond energy, which may offer additional corroboration of the quantitative aspects of our results although they involved less representative models of a polypeptide. Extraction of the CH··O bond energy was unfortunately complicated by its occurrence as a secondary H-bond, angularly distorted, and weaker than the stronger NH··O typically present in the same dimer. Because the optimized geometries contained multiple H-bonds, as many as four in some cases, the extraction of the energy of any single H-bond was further clouded by issues of cooperativity. Nor did this earlier work take account of the surroundings, concentrating on a vacuum environment. In any event, in one study of pairs of small amides,[91] the CH donors were not of the same C^α sort that occur in polypeptides, but rather the proton was bonded directly to the amide C, a group which does not occur in polypeptides. This work led to estimates of the CH··O H-bond energy in the range of 1.9−2.6 kcal/mol. Other calculations[92] modeled the polypeptide chain by a pair of dimethylformamide molecules, in which the CH donor is bordered on only one side by an amide function, arriving at an estimate in the range of 2.1−2.7 kcal/mol. An attempt to extrapolate their result to a more realistic model of a polypeptide led to a rough guess of 3.5−4.5 kcal/mol. Another work was more relevant in that the donor was an aliphatic CH_2 group,[93] but this group was again bordered by only one amide instead of two. Moreover, each geometry of the parent molecule contained various H-bonds, each of different type and each with different geometry (all angularly distorted), preventing any sort of unambiguous determination of the CH··O interaction energy.

The dipeptide investigated here is representative of glycine, with its pair of $C^\alpha H$ hydrogens. Earlier computations[60] had compared Gly with a number of other amino acids and found the strength of its $C^\alpha H$··O bond to water to be similar to those of Ala, Val, Ser, and Cys. It is hence anticipated that the results reported here are not limited to Gly alone, but are applicable to the full range of amino acids, albeit with slight modifications.

References and Notes

(1) Desiraju, G. R.; Steiner, T. *The Weak Hydrogen Bond in Structural Chemistry and Biology*; Oxford: New York, 1999.
(2) Scheiner, S. CH···O Hydrogen Bonding. In *Advances in Molecular Structure Research*; Hargittai, M., Hargittai, I., Eds.; JAI Press: Stamford, CT, 2000; Vol. 6, pp 159−207.
(3) Steiner, T. *Angew. Chem., Int. Ed.* **2002**, *41*, 48−76.
(4) Hobza, P.; Havlas, Z. *Theor. Chem. Acc.* **2002**, *108*, 325−334.
(5) Delanoye, S. N.; Herrebout, W. A.; van der Veken, B. J. *J. Am. Chem. Soc.* **2002**, *124*, 11854−11855.
(6) Vishweshwar, P.; Thaimattam, R.; Jaskolski, M.; Desiraju, G. R. *Chem. Commun.* **2002**, 1830−1831.
(7) Barnes, A. J. *J. Mol. Struct.* **2004**, *704*, 3−9.
(8) Derewenda, Z. S.; Lee, L.; Derewenda, U. *J. Mol. Biol.* **1995**, *252*, 248−262.
(9) Leonard, G. A.; McAuley-Hecht, K.; Brown, T.; Hunter, W. N. *Acta Crystallogr.* **1995**, *D51*, 136−139.
(10) Wahl, M. C.; Sundaralingam, M. *Trends Biochem. Sci.* **1997**, *22*, 97−102.
(11) Mandel-Gutfreund, Y.; Margalit, H.; Jernigan, R. L.; Zhurkin, V. B. *J. Mol. Biol.* **1998**, *277*, 1129−1140.
(12) Weiss, M. S.; Brandl, M.; Sühnel, J.; Pal, D.; Hilgenfeld, R. *Trends Biochem. Sci.* **2001**, *26*, 521−523.
(13) Loll, B.; Raszewski, G.; Saenger, W.; Biesiadka, J. *J. Mol. Biol.* **2003**, *328*, 737−747.
(14) Cordier, F.; Barfield, M.; Grzesiek, S. *J. Am. Chem. Soc.* **2003**, *125*, 15750−15751.
(15) Kang, B. S.; Devedjiev, Y.; Derewenda, U.; Derewenda, Z. S. *J. Mol. Biol.* **2004**, *338*, 483−493.
(16) Lee, K. M.; Chang, H.-C.; Jiang, J.-C.; Chen, J. C. C.; Kao, H.-E.; Lin, S. H.; Lin, I. J. B. *J. Am. Chem.* Soc. **2003**, *125*, 12358−12364.
(17) Manikandan, K.; Ramakumar, S. *Proteins: Struct., Funct., Genet.* **2004**, *56*, 768−781.
(18) Singh, S. K.; Babu, M. M.; Balaram, P. *Proteins: Struct., Funct., Genet.* **2003**, *51*, 167−171.
(19) Senes, A.; Ubarretxena-Belandia, I.; Engelman, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 9056−9061.
(20) Babu, M. M.; Singh, S. K.; Balaram, P. *J. Mol. Biol.* **2002**, *322*, 871−880.
(21) Aravinda, S.; Shamala, N.; Bandyopadhyay, A.; Balaram, P. *J. Am. Chem. Soc.* **2003**, *125*, 15065−15075.
(22) Klaholz, B. P.; Moras, D. *Structure* **2002**, *10*, 1197−1204.
(23) Ramagopal, U. A.; Ramakumar, S.; Sahal, D.; Chauhan, V. S. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 870−874.

(24) Petrella, R. J.; Karplus, M. *Proteins: Struct., Funct., Genet.* **2004**, *54*, 716−726.
(25) Sarkhel, S.; Desiraju, G. R. *Proteins* **2004**, *54*, 247−259.
(26) Shi, Z.; Olson, C. A.; Bell, A. J.; Kallenbach, N. R. *Biophys. Chem.* **2002**, *101*−*102*, 267−279.
(27) Baures, P. W.; Beatty, A. M.; Dhanasekaran, M.; Helfrich, B. A.; Perez-Segarra, W.; Desper, J. *J. Am. Chem. Soc.* **2002**, *124*, 11315−11323.
(28) Pierce, A. C.; Sandretto, K. L.; Bemis, G. W. *Proteins: Struct., Funct., Genet.* **2002**, *49*, 567−576.
(29) Jiang, L.; Lai, L. *J. Biol. Chem.* **2002**, *277*, 37732−37740.
(30) Arbely, E.; Arkin, I. T. *J. Am. Chem. Soc.* **2004**, *126*, 5362−5363.
(31) Yohannan, S.; Faham, S.; Yang, D.; Grosfeld, D.; Chamberlain, A. K.; Bowie, J. U. *J. Am. Chem. Soc.* **2004**, *126*, 2284−2285.
(32) Scheiner, S. The CH··O Hydrogen Bond. A Historical Account. In *Theory and Applications of Computational Chemistry: The First 40 Years*; Dykstra, C., Frenking, G., Kim, K., Scuseria, G., Eds., in press.
(33) Novoa, J. J.; Lafuente, P.; Mota, F. *Chem. Phys. Lett.* **1998**, *290*, 519−525.
(34) Alkorta, I.; Elguero, J. *J. Phys. Chem. A* **1999**, *103*, 272−279.
(35) Cubero, E.; Orozco, M.; Luque, F. J. *Chem. Phys. Lett.* **1999**, *310*, 445−450.
(36) Sponer, J.; Hobza, P. *J. Phys. Chem. A* **2000**, *104*, 4592−4597.
(37) Hartmann, M.; Wetmore, S. D.; Radom, L. *J. Phys. Chem. A* **2001**, *105*, 4470−4479.
(38) van der Veken, B.; Herrebout, W. A.; Szostak, R.; Shchepkin, D. N.; Havlas, Z.; Hobza, P. *J. Am. Chem. Soc.* **2001**, *123*, 12290−12293.
(39) Hermansson, K. *J. Phys. Chem. A* **2002**, *106*, 4695−4702.
(40) Futami, Y.; Kudoh, S.; Takayanagi, M.; Nakata, M. *Chem. Phys. Lett.* **2002**, *357*, 209−216.
(41) Cabaleiro-Lago, E. M.; Otero, J. R. *J. Chem. Phys.* **2002**, *117*, 1621−1632.
(42) Xu, Z.; Li, H.; Wang, C.; Wu, T.; Han, S. *Chem. Phys. Lett.* **2004**, *394*, 405−409.
(43) Herrebout, W. A.; Delanoye, S. N.; Veken, B. J. v. d. *J. Phys. Chem. A* **2004**, *108*, 6059−6064.
(44) Raveendran, P.; Wallen, S. L. *J. Am. Chem. Soc.* **2002**, *124*, 12590−12599.
(45) Wang, B.; Hinton, J. F.; Pulay, P. *J. Phys. Chem. A* **2003**, *107*, 4683−4687.
(46) 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.; Baboul, A. G.; 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.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 03*; Gaussian, Inc.: Pittsburgh, PA, 2003.
(47) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(48) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(49) Latajka, Z.; Bouteiller, Y. *J. Chem. Phys.* **1994**, *101*, 9793−9799.
(50) Kim, K.; Jordan, K. D. *J. Phys. Chem.* **1994**, *98*, 10089−10094.
(51) Del Bene, J. E.; Person, W. B.; Szczepaniak, K. *J. Phys. Chem.* **1995**, *99*, 10705−10707.
(52) Rablen, P. R.; Lockman, J. W.; Jorgensen, W. L. *J. Phys. Chem. A* **1998**, *102*, 3782−3797.
(53) Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486−1493.
(54) Wong, M. W.; Frisch, M. J.; Wiberg, K. B. *J. Am. Chem. Soc.* **1991**, *113*, 4776−4782.
(55) Wong, M. W.; Wiberg, K. B.; Frisch, M. *J. Chem. Phys.* **1991**, *95*, 8991−8998.
(56) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117−129.
(57) Miertus, S.; Tomasi, J. *Chem. Phys.* **1982**, *65*, 239−245.
(58) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151−5198.
(59) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995−2001.
(60) Scheiner, S.; Kar, T. *J. Phys. Chem. B* **2005**, *109*, 3681−3689.
(61) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210−3221.
(62) Dian, B. C.; Longarte, A.; Mercier, S.; Evans, D. A.; Wale, D. J.; Zwier, T. S. *J. Chem. Phys.* **2002**, *117*, 10688−10700.
(63) Chin, W.; Dognon, J.-P.; Canuel, C.; Piuzzi, F.; Dimicoli, I.; Mons, M.; Compagnon, I.; Helden, G. v.; Meijer, G. *J. Chem. Phys.* **2005**, *122*, 054317.
(64) Gould, I. R.; Cornell, W. D.; Hillier, I. H. *J. Am. Chem. Soc.* **1994**, *116*, 9250−9256.
(65) Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. B* **1999**, *103*, 4730−4737.
(66) Vargas, R.; Garza, J.; Hay, B. P.; Dixon, D. A. *J. Phys. Chem. A* **2002**, *106*, 3213−3218.
(67) Wang, Z.-X.; Duan, Y. *J. Comput. Chem.* **2004**, *25*, 1699−1716.
(68) C═O···H angles of 180° are generally slightly less stable than smaller angles, but this difference tends to be rather small.
(69) Sharp, K. A.; Honig, B. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 301−335.
(70) Simonson, T.; Perahia, D.; Brunger, A. T. *Biophys. J.* **1991**, *59*, 670−690.
(71) Smith, P. E.; Brunne, R. M.; Mark, A. E.; van Gunsteren, W. F. *J. Phys. Chem.* **1993**, *97*, 2009−2014.
(72) Simonson, T.; Perahia, D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1082−1086.
(73) Dwyer, J. J.; Gittis, A. G.; Karp, D. A.; Lattman, E. E.; Spencer, D. S.; Stites, W. E.; Garcia-Moreno, B. *Biophys. J.* **2000**, *79*, 1610−1620.
(74) Scarsdale, J. N.; Alsenoy, C. V.; Klimkowski, V. J.; Schaefer, L.; Momany, F. A. *J. Am. Chem. Soc.* **1983**, *105*, 3438−3445.
(75) Head-Gordon, T.; Head-Gordon, M.; Frisch, M. J.; Brooks, C. L.; Pople, J. A. *J. Am. Chem. Soc.* **1991**, *113*, 5989−5997.
(76) Frey, R. F.; Coffin, J.; Newton, S. Q.; Ramek, M.; Cheng, V. K. W.; Momany, F. A.; Schaefer, L. *J. Am. Chem. Soc.* **1992**, *114*, 5369−5377.
(77) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 5908−5920.
(78) Dugourd, P.; Antoine, R.; Breaux, G.; Broyer, M.; Jarrold, M. F. *J. Am. Chem. Soc.* **2005**, *127*, 4675−4679.
(79) Harvey, S. C. *Proteins: Struct., Funct., Genet.* **1989**, *5*, 78−92.
(80) Schutz, C. N.; Warshel, A. *Proteins: Struct., Funct., Genet.* **2001**, *44*, 400−417.
(81) Tomasi, J. *Theor. Chem. Acc.* **2004**, *112*, 184−203.
(82) Winget, P.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.* **2004**, *112*, 217−227.
(83) Moreau, Y.; Loos, P.-F.; Assfeld, X. *Theor. Chem. Acc.* **2004**, *112*, 228−239.
(84) Cannizzaro, C. E.; Houk, K. N. *J. Am. Chem. Soc.* **2002**, *124*, 7163−7169.
(85) Ribeiro-Claro, P. J. A.; Vaz, P. D. *Chem. Phys. Lett.* **2004**, *390*, 358−361.
(86) Melikova, S. M.; Rutkowski, K. S.; Rodziewicz, P.; Koll, A. *J. Mol. Struct.* **2004**, *705*, 49−61.
(87) Zhang, H.; Zhou, Z.; Shi, Y. *J. Phys. Chem. A* **2004**, *108*, 6735−6743.
(88) Gong, X.; Zhou, Z.; Du, D.; Dong, X.; Liu, S. *Int. J. Quantum Chem.* **2005**, *103*, 105−117.
(89) Ruotolo, B. T.; Russell, D. H. *J. Phys. Chem. B* **2004**, *108*, 15321−15331.
(90) Scheiner, S.; Kar, T.; Gu, Y. *J. Biol. Chem.* **2001**, *276*, 9832−9837.
(91) Vargas, R.; Garza, J.; Friesner, R. A.; Stern, H.; Hay, B. P.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 4963−4968.
(92) Vargas, R.; Garza, J.; Dixon, D. A.; Hay, B. P. *J. Am. Chem. Soc.* **2000**, *122*, 4750−4755.
(93) Vargas, R.; Garza, J.; Dixon, D. A.; Hay, B. P. *J. Phys. Chem. A* **2000**, *104*, 5115−5121.