

## How Strong Is the C<sup>α</sup>–H···O=C Hydrogen Bond?

Rubicelia Vargas,<sup>†</sup> Jorge Garza,<sup>†</sup> David A. Dixon,\* and Benjamin P. Hay\*

Contribution from the Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, P.O. Box 999, Richland, Washington 99352

Received October 6, 1999. Revised Manuscript Received March 14, 2000

**Abstract:** Although the existence of C<sup>α</sup>–H···O=C hydrogen bonds in protein structures recently has been established, little is known about their strength and, therefore, the relative importance of these interactions. We have discovered that similar interactions occur in *N,N*-dimethylformamide dimers. High level ab initio calculations (MP2/aug-cc-pTZV) yield electronic association energies ( $D_e$ ) and association enthalpies ( $\Delta H^{298}$ ) for four dimer geometries. These data provide a lower limit of  $D_e = -2.1$  kcal mol<sup>-1</sup> for the C<sup>α</sup>–H···O=C hydrogen bond. A linear correlation between C–H···O bond energies and gas-phase proton affinities is reported. The gas-phase anion proton affinity of a peptide C<sup>α</sup>–H hydrogen was calculated (355 kcal mol<sup>-1</sup>) and used to estimate values of  $D_e = -4.0 \pm 0.5$  kcal mol<sup>-1</sup> and  $\Delta H^{298} = -3.0 \pm 0.5$  kcal mol<sup>-1</sup> for the C<sup>α</sup>–H···O=C hydrogen bond. The magnitude of this interaction, roughly one-half the strength of the N–H···O=C hydrogen bond, suggests that C<sup>α</sup>–H···O=C hydrogen bonding interactions represent a hitherto unrecognized, significant contribution in the determination of protein conformation.

### Introduction

It is increasingly recognized that C–H···O hydrogen bonds play an important role in determining molecular conformation and crystal packing,<sup>1</sup> in molecular recognition processes,<sup>2</sup> in the stabilization of inclusion complexes,<sup>3</sup> and possibly in the activity of biological macromolecules.<sup>4</sup> Much of the evidence for C–H···O hydrogen bonding comes from the observation of close C···O contacts in crystal structures. Although such interactions have been observed with a variety of possible C–H donors and oxygen acceptors,<sup>5–7</sup> including amino acids,<sup>8</sup> it is

only recently that C–H···O hydrogen bonds have been observed in the crystal structures of proteins. Surveys of high-resolution protein structures reveal the widespread occurrence of close C–H···O contacts.<sup>9–13</sup> The majority of the reported contacts involve hydrogen atoms attached to the  $\alpha$ -carbons in the peptide backbone. This is consistent with the polarized character of the C<sup>α</sup>–H bond, adjacent to electron withdrawing N–H and C=O groups. The most frequently observed hydrogen bond acceptors are the carbonyl oxygens in adjacent peptide strands of  $\beta$ -sheets (Figure 1).<sup>10–12</sup> In these cases, one oxygen lone pair interacts with a C<sup>α</sup>–H proton and the other oxygen lone pair interacts with an N–H proton. This bifurcated hydrogen bonding pattern also has been observed in helical protein structures.<sup>11</sup> A different bifurcated hydrogen bonding topology has been observed in a collagen triple helix in which both oxygen lone pairs contact C<sup>α</sup>–H protons and one of the hydrogen atoms is shared between two oxygens (Figure 1).<sup>11</sup>

Although it has been established that C<sup>α</sup>–H···O=C interactions exist in protein structures, their role in determining structure is unclear. Nothing is known, either experimentally or theoretically, about the strength of this interaction in proteins. It is generally assumed the C<sup>α</sup>–H···O=C hydrogen bond is much weaker than other interactions, e.g., N–H···O=C hydrogen bonds, between peptide strands. It has been noted that C<sup>α</sup>–H···O=C hydrogen bonds may be responsible for alleviating the unsaturated hydrogen bonding potential of buried polar groups with the suggestion that the total contribution of the C–H···O bonds may, at least in some proteins, play a significant role in stabilizing the folded structure.<sup>10</sup> It also has been suggested that C–H···O bonds probably should not be considered as a determinant of any particular folded structure, but that

<sup>†</sup> Permanent address: Departamento de Química, División de Ciencias Básicas e Ingeniería Universidad Autónoma Metropolitana-Iztapalapa, Apartado Postal 55-534, México Distrito Federal 09340, México.

(1) (a) Green, R. D. *Hydrogen Bonding by C–H Groups*; Macmillan: London, 1974. (b) Z. Berkovich-Yellin, Z.; Leiserowitz, L. *Acta Crystallogr., Sect. B* **1984**, *40*, 159. (c) Desiraju, G. R. *Acc. Chem. Res.* **1996**, *29*, 441. (d) Steiner, T. *Crystallogr. Rev.* **1996**, *6*, 1. (e) Allen, F. J. P.; Lommerse, M.; Hoy, V. J.; Howard, J. A. K.; Desiraju, G. R. *Acta Crystallogr., Sect. B* **1996**, *52*, 734. (e) Steiner, T. *Chem. Commun.* **1997**, 727. (f) Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press: Oxford, 1997. (g) Desiraju, G. R.; Steiner, T. *The Weak Hydrogen Bond in Structural Chemistry and Biology*; Oxford University Press: Oxford, 1999.

(2) (a) Shimon, L. J. W.; Vaida, M.; Addadi, L.; Lahav, M.; Leiserowitz, L. *J. Am. Chem. Soc.* **1990**, *112*, 6215. (b) Keegstra, E. M. D.; Spek, A. L.; Zwicker, J. W.; Jenneskens, L. W. *J. Chem. Soc., Chem. Commun.* **1994**, 1633. (c) Desiraju, G. R. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 2311.

(3) (a) Steiner, T.; Saenger, W. *J. Chem. Soc., Chem. Commun.* **1995**, 2087. (b) Braga, D.; Grepioni, F.; Byrne, J. J.; Wolf, A. *J. Chem. Soc., Chem. Commun.* **1995**, 125. (c) Behrens, P.; van de Goo, G.; Freyhardt, C. *C. Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 2895. (d) Houk, K. N.; Menzer, S.; Newton, S. P.; Raymo, R. M.; Stoddart, J. F.; Williams, D. J. *J. Am. Chem. Soc.* **1999**, *121*, 1479.

(4) (a) Egli, M.; Gessner, R. V. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *189*, 92. (b) Leonard, G. A.; McAuly-Hecht, K.; Brown, T.; Hunter, W. N. *Acta Crystallogr., Sect. D* **1995**, *51*, 136. (c) Berger, I.; Egli, M.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *264*, 734. (d) Musah, R. A.; Jensen, G. M.; Rosenfeld, R. J.; McRee, D. E.; Goodin, D. B. *J. Am. Chem. Soc.* **1997**, *119*, 9083.

(5) (a) Sutor, D. *J. Nature* **1962**, *193*, 68. (b) Desiraju, G. R. *Acc. Chem. Res.* **1991**, *290*, 24. (c) Krishnamohan, C. V.; Desiraju, G. R. *J. Chem. Soc., Perkin Trans. 2* **1994**, 2345.

(6) Taylor, R.; Kennard, O. *J. Am. Chem. Soc.* **1982**, *104*, 5063.

(7) Steiner, T.; Saenger, W. *J. Am. Chem. Soc.* **1992**, *114*, 10146.

(8) Steiner, T. *J. Chem. Soc., Perkin Trans. 2* **1995**, 1315.

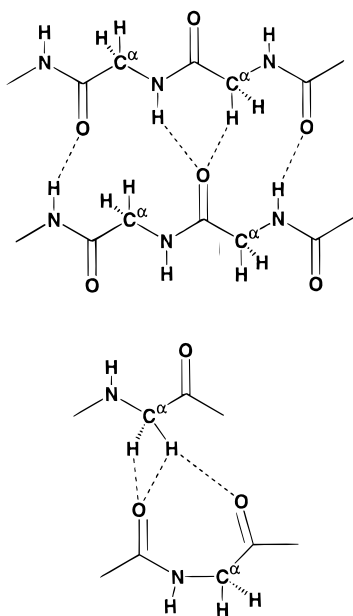
(9) Derewenda, Z. S.; Derewenda, U.; Kobos, P. M. *J. Mol. Biol.* **1994**, *241*, 83.

(10) Derewenda, Z. S.; Lee, L.; Derewenda, U. *J. Mol. Biol.* **1995**, *252*, 248.

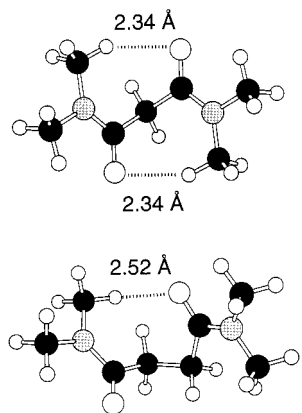
(11) Bella, J.; Berman, H. M. *J. Mol. Biol.* **1996**, *264*, 734.

(12) Fabiola, G. F.; Krishnaswamy, S.; Nagarajan, V.; Pattabhi, V. *Acta Crystallogr., Sect. D* **1997**, *53*, 316.

(13) Chakrabarti, P.; Chakrabarti, S. *J. Mol. Biol.* **1998**, *284*, 867.



**Figure 1.** The  $C^{\alpha}-H\cdots O=C$  interactions found in protein  $\beta$ -sheets (top) and in the collagen triple helix (bottom). Dashed lines represent hydrogen bonds.



**Figure 2.** Lowest energy (MP2/DZVP) conformations for  $N,N,N',N'$ -tetramethylmalonamide (top) and  $N,N,N',N'$ -tetramethylsuccinamide (bottom). As in proteins, these structures exhibit  $C-H\cdots O$  hydrogen bonds between an amide carbonyl acceptor and a  $C-H$  hydrogen donor attached to an amide nitrogen.

they more likely act to “fine-tune” local structure after the protein has adopted its final folded state.<sup>11</sup>

We now present computational evidence suggesting that  $C^{\alpha}-H\cdots O=C$  interactions are of sufficient strength that they are capable of exerting a significant influence on protein structure. As part of our work on the ab initio design of ligands for radionuclide separations, we have been studying structures of amides and their metal complexes.<sup>14–16</sup> Conformational analyses of  $N,N,N',N'$ -tetramethylmalonamide<sup>15</sup> and  $N,N,N',N'$ -tetramethylsuccinamide<sup>16</sup> yield structures that contain internal  $C-H\cdots O$  hydrogen bonds analogous to those seen in proteins (Figure 2). To investigate the nature of these interactions, we examined  $C-H\cdots O$  hydrogen bonding in a model system, the

(14) (a) Clement, O.; Rapko, B. M.; Hay, B. P. *Coord. Chem. Rev.* **1998**, *170*, 203. (b) Hay, B. P.; Clement, O.; Sandrone, G.; Dixon, D. A. *Inorg. Chem.* **1998**, *37*, 5887. (c) Sandrone, G.; Dixon, D. A.; Hay, B. P. *J. Phys. Chem. A* **1999**, *103*, 993.

(15) Sandrone, G.; Dixon, D. A.; Hay, B. P. *J. Phys. Chem. A* **1999**, *103*, 3554.

(16) Vargas, R.; Garza, J.; Dixon, D. A.; Hay, B. P. *J. Phys. Chem. A*. In press.

dimers of  $N,N$ -dimethylformamide (DMF), with high-level ab initio molecular orbital calculations. We chose this amide because it is small enough to allow the application of accurate theoretical methods and, with dialkyl substitution at nitrogen, it is unable to engage in  $N-H\cdots O=C$  hydrogen bonding. Here we report the structures and binding energies of the DMF dimers, and use these data to provide the first reliable estimate for the strength of the  $C^{\alpha}-H\cdots O=C$  interaction.

## Methods

All calculations were done with the program systems Gaussian 98<sup>17</sup> and DGauss.<sup>18</sup> Initial scouting calculations to locate stable DMF dimer geometries (**1–4**) were done with density functional theory (DFT) at the local level (VWN)<sup>19</sup> and the gradient-corrected level (BPW91)<sup>20</sup> with a polarized triple- $\zeta$  basis set (TZVP),<sup>21</sup> and at the gradient-corrected level (B3LYP)<sup>22</sup> with an augmented polarized double- $\zeta$  basis set (DZP+diff).<sup>23</sup> Optimized geometries from these calculations were used as input for geometry optimizations at the second-order Moller–Plesset Perturbation Theory (MP2) level<sup>24</sup> with an augmented polarized double- $\zeta$  basis set (DZP+diff).<sup>23</sup> Electronic association energies,  $D_e$ , were computed at the VWN/TZVP, BPW91/TZVP, B3LYP/DZP+diff, and MP2/DZP+diff levels, and with single point energy calculations at the MP2/aug-cc-pVTZ level<sup>25,26</sup> using the MP2/DZP+diff level geometries. Basis set superposition error (BSSE) was estimated for the MP2 results by the counterpoise method.<sup>27</sup> Thermodynamic values  $\Delta E^{\circ}$ ,  $\Delta E^{298}$ , and  $\Delta H^{298}$  were calculated only at the highest level of theory.<sup>28</sup>

There has been significant interest in the use of DFT to tackle biochemical problems, notably those dealing with hydrogen bond interactions.<sup>29</sup> From such studies, it has been concluded that DFT gradient corrections are required to obtain good hydrogen bond energies. However, we note that the results obtained here do not support this conclusion. Binding energies,  $D_e$  (not corrected for BSSE), for **1–4** at

(17) MP2 and B3LYP calculations were performed with *Gaussian 98, Revision A.4*; Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T. A.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Head-Gordon, M.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc.: Pittsburgh, PA, 1998.

(18) VWN and BPW91 calculations were performed with DGauss, a program that is part of UniChem and is available from Oxford Molecular: Andzelm, J.; Wimmer, E.; Salahub, D. R. In *The Challenge of d and f Electrons: Theory and Computations*; Salahub, D. R., Zerner, M. C., Eds.; ACS Symp. Ser., No. 394; American Chemical Society: Washington, DC, 1989; p 228. (b) Andzelm, J. In *Density Functional Methods in Chemistry*; Labanowski, J., Andzelm, J., Eds.; Springer-Verlag: New York, 1991. (c) Andzelm, J.; Wimmer, E. *J. Chem. Phys.* **1992**, *96*, 1280.

(19) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.

(20) (a) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098. (b) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.

(21) Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. *Can. J. Chem.* **1992**, *70*, 560.

(22) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (b) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372. (c) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev.* **1988**, *B37*, 785.

(23) (a) Dunning, T. H., Jr. *J. Chem. Phys.* **1970**, *53*, 2823. (b) Dunning, T. H., Jr.; Hay, P. J. In *Methods of Electronic Structure Theory*, Vol. 3; Schaefer, H. F., III, Ed.; Plenum Press: New York, 1977.

(24) (a) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618. (b) Pople, J. A.; Binkley, J. S.; Seeger, R. *Int. J. Quantum Chem. Symp.* **1976**, *10*, 1.

(25) We used the augmented correlation-consistent basis set that has been shown to be very successful at accurately predicting the strength of other hydrogen bonds: (a) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007. (b) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6769.

(26) Feyereisen, M. W.; Feller, D.; Dixon, D. A. *J. Phys. Chem.* **1996**, *100*, 2993.

(27) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1990**, *19*, 553.

(28) Frequencies needed to estimate the zero-point energies and thermal corrections to the energies were calculated at the DFT level (B3LYP/DZP+diff). This DFT functional gave geometries for the three dimers that were in the best overall agreement with the MP2/DZP+diff results.

**Table 1.** Binding energies ( $D_e$ ) in kcal mol<sup>-1</sup> for 1–4 at Different Computational Levels

struct.	VWN/ TZVP	BPW91/ TZVP	B3LYP/ DZP+diff	MP2/ DZP+diff	MP2/ aug-cc-pVTZ
1	-8.69 (-8.65) <sup>a</sup>	-2.42	-4.46	-6.61	-6.95
2	-6.71 (-6.97) <sup>a</sup>	-0.92	-3.06	-5.51	-5.82
3	-8.71 (-9.15) <sup>a</sup>	-0.58	-3.54	-10.32	-11.41
4	-11.59 (-12.13) <sup>a</sup>	-1.35	-3.84	-10.98	-12.11

<sup>a</sup> Values in parentheses obtained with exact Coulomb.

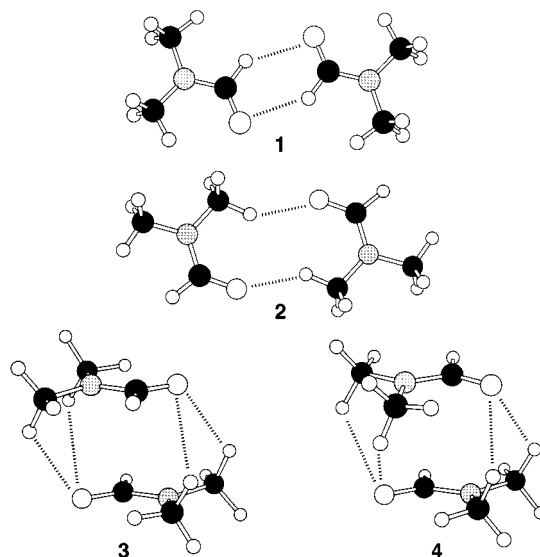
different DFT levels are compared with two sets of MP2 values in Table 1. In comparison to the MP2 values, the local DFT calculations give a much better binding energy than the gradient-corrected values. Both gradient-corrected methods give binding energies that are far too low and that fail to reproduce the correct energy ordering for 1–4. In comparison with the MP2 geometries, the local DFT method gives hydrogen bond distances that are too small by ~0.2 Å whereas the gradient-corrected DFT methods give hydrogen bond distances that are generally too large by ~0.1 Å (see the Supporting Information). However, we note that in one case, 4, a gradient-corrected DFT method yields H···O distances in error by up to 0.4 Å.

Similar methods were used in the calculation of gas-phase proton affinities for DMF and the peptide model compound, 5. Geometries were optimized initially at the VWN/TZVP level and frequencies for the thermodynamic and zero point energy corrections were calculated at this level. Geometries were subsequently optimized at the MP2/DZP+diff level.

The Cambridge Structural Database<sup>30</sup> was screened for ordered and error free structures containing two or more DMF molecules. This subset of structures was examined for intermolecular contacts between the oxygen of one DMF and a carbon atom of a second DMF in which the C···O distance was between 2.5 and 4.0 Å. Visual inspection of the final subset of structures revealed over 100 examples of DMF dimers exhibiting C–H···O hydrogen bonds. In the majority of cases, the dimers contain only one hydrogen bond. However, attempts to calculate single hydrogen bonded species outside the crystal environment failed as electronic structure calculations using initial coordinates from crystal structures yielded only multiple hydrogen bonded dimers, i.e., 1–4. Inspection of the final subset of structures also revealed at least one example for each of the multiple hydrogen bonded dimers 1–4. To allow comparison of calculated and experimental geometries, hydrogen atoms coordinates were generated using the rectify feature of the Chem3D program.<sup>31</sup> A list of reference codes for each type of DMF dimer is provided as Supporting Information.

## Results and Discussion

We obtained four stable geometries, 1–4 (Figure 3), at the MP2 level. Each dimer exhibits two or more C–H···O hydrogen bonds. In dimer 1, each oxygen is hydrogen bonded to a formyl hydrogen. An analogous geometry has been reported for the formamide dimer.<sup>32</sup> In dimer 2 each oxygen is hydrogen bonded to one *N*-methyl hydrogen. Dimer 3 and 4 exhibit a bifurcated pattern in which each oxygen atom is hydrogen bonded to two *N*-methyl hydrogens. Prior ab initio calculations with minimal basis sets yielded a DMF dimer with one C–H···O hydrogen

**Figure 3.** Stable geometries of the DMF dimer.

bond to a formyl hydrogen.<sup>33</sup> More recently, force field calculations yielded a DMF dimer with the same C–H···O hydrogen bonds as dimer 1 except that the two monomers were in a face-to-face, rather than coplanar, orientation.<sup>34</sup> However, we find neither of these alternate dimer structures to be minima at the MP2/DZP+diff level of theory. Moreover, although there are examples of 1–4 in crystal structures,<sup>35</sup> we did not locate any examples for the previously proposed dimers. A recent crystal structure of neat DMF has been described as sheets of cyclic tetramers held together by C–H···O interactions involving two formyl hydrogens and two methyl hydrogens.<sup>35e</sup> We note that this structure provides another example of dimer 2.

The calculated geometric parameters associated with the C–H···O interactions are presented in Table 2 where they can be compared with “ideal” values, crystallographic examples, and average values for C<sup>α</sup>–H···O=C hydrogen bonds that have been observed in proteins. First we discuss the idealized geometry. Although an ideal H···O distance is not defined, a distance of less than the van der Waals contact distance is often taken as evidence for a C–H···O interaction. Using van der Waals radii taken from Taylor and Kennard,<sup>6</sup> this corresponds to a distance less than 2.7 Å. Given a C–H bond length on the order of 1.1 Å, the H···O distance translates to a C···O distance of less than 3.8 Å. It should be noted, however, that the weak C–H···O interaction often exhibits distances in excess of 2.7 Å.<sup>1g</sup> Thus, while a short contact is indicative of C–H···O hydrogen bonding, one cannot rule out a C–H···O interaction at longer distances. The C–H···O interactions in all crystal structures of DMF dimers that we examined reveal a C···O distance range from 3.1 to 3.9 Å with an average value of 3.6 Å.

Examination of Table 2 reveals that the calculated geometries all exhibit C···O distances and, with the exception of 3, H···O

(29) For example: (a) Sim, F.; Amant, St. A.; Papai, I.; Salahub, D. R. *J. Am. Chem. Soc.* **1992**, *114*, 4391. (b) Hill, R. A.; Labanowski, J. K.; Heisterberg, D. J.; Miller, D. D. In *Density Functional Theory in Chemistry*; Labanowski, J., Andzelm, J., Eds.; Springer-Verlag: New York, 1991; p 357. (c) Proynov, E. I.; Ruiz, E.; Vela, A.; Salahub, D. R. *Int. J. Quantum Chem.* **1995**, *89*, 61. (d) Proynov, E. I.; Sirois, S.; Salahub, D. R. *Int. J. Quantum Chem.* **1997**, *64*, 427.

(30) *Cambridge Structural Database System, Version V5.18, October 1999 Release*; Allen, F. H.; Bellard, S. A.; Brice, B. A.; Cartwright, A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B. G.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. *Acta Crystallogr., Sect. B* **1979**, *35*, 2331.

(31) *Chem3D*; CambridgeSoft: 100 Cambridge Park Drive, Cambridge, MA 02140.

(32) Neuheuser, T.; Hess, B. A.; Reutel, C.; Weber, E. *J. Phys. Chem.* **1994**, *98*, 6459.

(33) Ohtaki, H.; Itoh, S.; Yamaguchi, T.; Ishiguro, S.; Rode, B. M. *Bull. Chem. Soc. Jpn.* **1983**, *56*, 3406.

(34) Desfrancois, C.; Periquet, V.; Carles, S.; Schermann, J. P.; Adamowicz, L. *Chem. Phys.* **1998**, *239*, 475.

(35) Examples of each dimer: (a) 1: Ray, M.; Golombek, A. P.; Hendrich, M. P.; Young, V. G.; Borovik, A. S. *J. Am. Chem. Soc.* **1996**, *118*, 6084. (b) 2: Raper, E. S.; Britton, A. M.; Creighton, J. R.; Clegg, W.; Hooper, M.; Kubiak, M. *Acta Crystallogr., Sect. C* **1987**, *43*, 1538. (c) 3: Wenkert, E.; Moeller, P. D. R.; Piettre, S. R.; McPhail, A. T. *J. Org. Chem.* **1987**, *52*, 3404. (d) 4: Oliver, J. E.; Waters, R. M.; Lusby, W. R.; Flippen-Anderson, J. L. *J. Heterocycl. Chem.* **1991**, *28*, 1569. (e) Structure of neat DMF: Borrmann, H.; Persson, I.; Sandström, M.; Stalhandske, C. M. V. *J. Chem. Soc., Perkin Trans. 2* **2000**, 393.



**Table 2.** Geometric Parameters for the C-H···O Interaction<sup>a</sup>

structure	H···O	C···O	CH···O	C=O···H	Θ
ideal	≤2.7	≤3.8	180	120	0
<b>1</b> calcd	2.36	3.32	146	92	0
<b>1</b> obsd	2.81	3.40	112	113	0
<b>2</b> calcd	2.36	3.29	143	145	3
<b>2</b> obsd	2.69	3.61	141	141	5
<b>3</b> calcd	2.52	3.25	123	125	42
	2.87	3.48	115	79	57
<b>3</b> obsd	3.20	3.72	109	125	53
	3.30	3.78	107	89	51
<b>4</b> calcd	2.59	3.32	124	103	43
	2.55	3.32	127	111	32
<b>4</b> obsd	3.18	3.77	114	107	49
	2.83	3.47	117	110	48
parallelβ-sheet	2.28	3.22	145	131	5
antiparallelβ-sheet	2.32	3.16	134	143	7
collagen triple helix	2.79	3.15	91	109	na
	2.63	3.15	110	100	na

<sup>a</sup> Distances are given in Å. Angles are given in deg. Θ, the elevation angle,<sup>10</sup> is the angle between the C-H bond and the amide plane. Observed values for **1**–**4** are taken from crystal structures<sup>35</sup> after using Chem3D software to add hydrogen atoms. Values for the protein structures are taken from the literature: β-sheets (average values from Table 2 in ref 10), collagen triple helix as shown in Figure 1.<sup>11</sup>

distances that are less than the estimated van der Waals contacts. In comparison the crystal structure examples of **1**–**4** generally exhibit longer distances and, in the case of H···O, these distances often exceed the 2.7 Å van der Waals cutoff. The protein structures show short H···O distances in the β-sheets and longer H···O distances in the collagen triple helix.

In the ideal case it is assumed that the C-H bond would point toward the oxygen giving a C-H···O angle of 180°, the hydrogen would align with the lone pair of the oxygen giving an H···O=C angle of 120°, and the C-H bond would lie in the plane of the amide moiety giving an elevation angle of 0°.<sup>10</sup> Although observed directionality is often presented as evidence for the presence of a C-H···O hydrogen bond, examination of crystal structures reveals large deviations from the perceived ideal case are common. For example, C-H···O bond angles range from 90° to 180° with no observable preference for linearity.<sup>5–7</sup> Such behavior is fully consistent for weak interactions and a recent application of Bader's atoms-in-molecules methodology<sup>36</sup> has confirmed that C-H···O hydrogen bonds are present even when there are large deviations from the optimal 180° angle.<sup>37</sup> Similarly, a survey of the structural features in C-H···O hydrogen bonds between carbonyl oxygen acceptors and acidic C-H donors (ethyne, chloroform, and methylene chloride) reveals only a weak tendency to exhibit the expected oxygen directionality.<sup>38</sup> In this case, C=O···H angles ranged from 100° to 170° with an average value of 135° and the elevation angle ranged from 0° to 90°, with an average value of 33°. In the case of the DMF dimers (see Table 2), both calculated and observed geometries show significant deviations from the idealized C-H···O and C=O···H angles. As noted above, however, such deviations are fully consistent with the known behavior of C-H···O hydrogen bonds.

Comparison of the DMF dimer geometries to those observed in proteins reveals a striking similarity between the C-H···O hydrogen bond in dimer **2** and those observed in protein β-sheets. Similarly, there is a close correspondence between the C-H···O hydrogen bonds in dimer **4** and the bifurcated

hydrogen bonding seen in the collagen triple helix (Figure 1). These structural similarities suggest that the intermolecular interactions in dimers **2** and **4** are representative of those found in the proteins.

Interaction energies and thermodynamic properties of the dimers are given in Table 3. The results reveal that **3** and **4**, with four C-H···O interactions each, are significantly more stable than **1** and **2**, with two C-H···O interactions each, suggesting that **3** and/or **4** may be the structure of the DMF dimer observed in solution by NMR spectroscopy.<sup>39,40</sup> Our survey of the CSD revealed **3** (30 examples) to be more populated than **4** (1 example) in a crystalline environment. The gas-phase Δ*H*<sup>298</sup> values for **1**, -3.86 kcal mol<sup>-1</sup>, and **4**, -7.42 kcal mol<sup>-1</sup>, bracket the measured enthalpy for DMF dimerization in CCl<sub>4</sub>, -5.9 kcal mol<sup>-1</sup>.<sup>40</sup>

Dividing the BSSE-corrected *D<sub>e</sub>* values by the number of C-H···O interactions yields an estimate of the strength of the C-H···O hydrogen bonds in the DMF dimers at 0 K uncorrected for zero-point energies. The resulting values are -2.68 kcal mol<sup>-1</sup> for **1**, -2.07 kcal mol<sup>-1</sup> for **2**, -2.09 kcal mol<sup>-1</sup> for **3**, and -2.23 kcal mol<sup>-1</sup> for **4**. These values can be compared directly with BSSE-corrected *D<sub>e</sub>* values calculated at comparable levels of theory for other types of C-H···O hydrogen bonds. When the acceptor is the sp<sup>3</sup> oxygen of water, *D<sub>e</sub>* values (kcal mol<sup>-1</sup>) are -0.49,<sup>41</sup> -0.43,<sup>42</sup> and -0.30<sup>43</sup> with methane, -0.38 with ethane,<sup>43</sup> -1.23<sup>42</sup> and -0.74<sup>43</sup> with fluoromethane, -1.01 with ethylene,<sup>41</sup> -2.24<sup>42</sup> and -1.48<sup>43</sup> with difluoromethane, -2.41 with trifluoromethane,<sup>43</sup> -2.64<sup>41</sup> and -2.73<sup>44</sup> with acetylene, and -4.95 with hydrogen cyanide.<sup>44</sup> When the acceptor is the sp<sup>2</sup> oxygen of formaldehyde, *D<sub>e</sub>* values (kcal mol<sup>-1</sup>) are -0.46 with methane,<sup>42</sup> -1.20 with fluoromethane,<sup>42</sup> -2.04 with difluoromethane,<sup>42</sup> -2.26 with acetylene,<sup>44</sup> and -3.90 with hydrogen cyanide.<sup>44</sup> When the acceptor is the sp<sup>2</sup> oxygen of acetamide, *D<sub>e</sub>* values (kcal mol<sup>-1</sup>) are -0.76 with methane, -1.37 with benzene, and -2.83 with the formyl hydrogen of acetal.<sup>45</sup> Thus, we note that the -2.07 to -2.68 kcal mol<sup>-1</sup> values calculated for the C-H···O hydrogen bonds in the DMF dimers fall in the middle of the range of values, -0.30 to -4.95 kcal mol<sup>-1</sup>, calculated for other C-H···O hydrogen bonds.<sup>46</sup> In addition, -2.68 kcal mol<sup>-1</sup> for the formyl hydrogen donor in dimer **1** is close to the value of -2.43 kcal mol<sup>-1</sup> reported for the C-H···O hydrogen bonds in the analogous formamide dimer<sup>32</sup> and to the value of -2.83 kcal mol<sup>-1</sup> for the formyl hydrogen donor in acetal.<sup>45</sup> The average hydrogen bond energy in dimers **2**–**4**, -2.1 kcal mol<sup>-1</sup>, provides the first estimate for the strength of the C=O interaction with C-H groups attached to an amide nitrogen. As we will demonstrate below, this value provides a lower limit for the strength of the C<sup>α</sup>-H···O=C hydrogen bonds that occur in proteins.

(39) Rabinovitz, M.; Pines, A. *J. Chem. Soc. (B)* **1968**, 1110.

(40) Bittrich, H.-J.; Kietz, E. *Z. Phys. Chem.* **1980**, 261, 17.

(41) Values obtained at the MP2/6-31++G(2d, 2p) level with BSSE correction: Rovira, M. C.; Novoa, J. J.; Whangbo, M.-H.; Williams, J. M. *Chem. Phys.* **1995**, 200, 319.

(42) Values obtained at the MP2/aug-cc-pVDZ level with BSSE correction: Gu, Y.; Kar, T.; Scheiner, S. *J. Am. Chem. Soc.* **1999**, 121, 9411.

(43) Values obtained at the MP2/NHFL(3d,2p) level with BSSE correction: Novoa, J. J.; Mota, F. *Chem. Phys. Lett.* **1997**, 266, 23.

(44) Values obtained at the MP2/D95++(d,p) level with BSSE correction: Turi, L.; Dannenberg, J. J. *J. Phys. Chem.* **1993**, 97, 7899.

(45) Values obtained at the PS-LMP2/cc-pVTZ(-fd) level with BSSE correction: Kim, K.; Friesner, R. A. *J. Am. Chem. Soc.* **1997**, 119, 12952.

(46) Values for C-H···O=C hydrogen bonds obtained at lower levels of theory also fall within this range: (a) Turi, L.; Dannenberg, J. J. *J. Phys. Chem.* **1993**, 97, 12197. (b) Turi, L.; Dannenberg, J. J. *J. Phys. Chem.* **1996**, 100, 9638. (c) Cardenas-Jiron, G. I.; Masunov, A.; Dannenberg, J. J. *J. Phys. Chem.* **1999**, 103, 7042.

(36) Bader, R. F. W. *Atoms in Molecules. A Quantum Theory*; Clarendon Press: Oxford, 1981.

(37) Novoa, J. J.; Lafuente, P.; Mota, F. *Chem. Phys. Lett.* **1998**, 290, 519.

(38) Steiner, T.; Kanters, J. A.; Kroon, J. *Chem. Commun.* **1996**, 1277.

**Table 3.** Interaction Energies and Thermodynamic Properties for the DMF Dimers 1–4<sup>a</sup>

dimer	MP2/DZP+diff		MP2/aug-cc-pVTZ		$\Delta E^\circ$	$\Delta E^{298}$	$\Delta H^{298}$
	$D_e$ w/o BSSE	$D_e$ w/BSSE	$D_e$ w/o BSSE	$D_e$ w/BSSE			
1	-6.61	-4.07	-6.95	-5.35	-4.59	-3.26	-3.86
2	-5.51	-3.31	-5.82	-4.14	-3.58	-2.26	-2.86
3	-10.32	-5.86	-11.41	-8.34	-7.59	-6.30	-6.90
4	-10.98	-6.36	-12.11	-8.90	-8.02	-6.82	-7.42

<sup>a</sup> All quantities are in kcal mol<sup>-1</sup>.  $D_e$  ( $\Delta E_{\text{elec}}$ ) are the electronic association energies without and with basis set superposition energy (BSSE) corrections.<sup>27</sup>  $\Delta E^\circ$  are the energy differences at 0 K after incorporating zero-point energies.  $\Delta E^{298}$  are the energy differences incorporating both zero-point energies and thermal corrections.<sup>28</sup>  $\Delta H^{298} = \Delta E^{298} + \Delta nRT$ .

There is evidence to suggest that C–H···O hydrogen bond strength depends on the acidity of the hydrogen donor. Thus, the more acidic hydrogens found in hydrogen cyanide, acetylene, and chloroform form stronger bonds than those in alkanes.<sup>1,38</sup> A linear correlation ( $r = 0.95$ ) has been observed between C···O distance and the  $pK_a$  (DMSO) of the C–H donor consistent with the concept that more acidic hydrogen donors form stronger, hence shorter, hydrogen bonds.<sup>47</sup> Moreover, linear correlations of C–H···O bond strength with both H···O distance ( $r = 0.97$ ) and proton acidities ( $r = 0.99$ ) have been obtained from AM1 calculations on the interactions of para-substituted phenylacetylenes with water.<sup>48</sup> It is expected that the C–H hydrogens of the DMF methyl groups should be less acidic than the C <sup>$\alpha$</sup> –H hydrogens of proteins because, whereas the former bonds are polarized by the amide nitrogen, the latter are polarized by both the amide nitrogen and the carbonyl group. Therefore, the C <sup>$\alpha$</sup> –H···O=C hydrogen bonds should have a  $D_e$  value of  $\geq 2.1$  kcal mol<sup>-1</sup>.

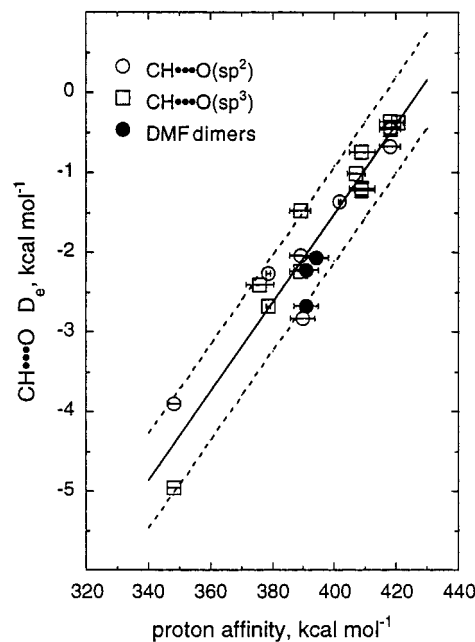
To obtain an estimate for the strength of the C <sup>$\alpha$</sup> –H···O=C hydrogen bond, we examined possible correlations between the  $D_e$  values for C–H···O interactions and parameters related to the acidity of the C–H proton. We discovered a linear correlation (Figure 4) between the  $D_e$  values<sup>41–45</sup> and experimental gas-phase proton affinities<sup>49</sup> for the conjugate anions of the various hydrogen donors where the proton affinity of the anion is defined as the enthalpy associated with the reaction  $AH \rightarrow A^- + H^+$ . Data for both sp<sup>2</sup> (formaldehyde and acetamide) and sp<sup>3</sup> (water) oxygen acceptors plot on the same line revealing that the C–H···O bond strength within this set of data is relatively independent of the type of oxygen acceptor. This correlation can be used to predict the  $D_e$  value for the C–H···O hydrogen bond from the corresponding proton affinity of the anion to within  $\pm 0.5$  kcal mol<sup>-1</sup>.

Since the experimental gas-phase proton affinities of the anions are not available for the C <sup>$\alpha$</sup> –H hydrogen atoms in proteins or for the three types of hydrogen atom in DMF, we calculated these values at the MP2/DZP+diff level of theory. The results are presented in Figure 5. When combined with their corresponding  $D_e$  values, the calculated DMF proton affinities

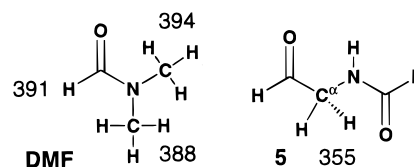
(47) Pedireddi, V. R.; Desiraju, G. R. *J. Chem. Soc., Chem. Commun.* **1992**, 988.

(48) Masunov, A.; Dannenberg, J. J. *J. Mol. Struct. THEOCHEM* **1996**, 371, 17.

(49) Experimental gas-phase proton affinities of the corresponding anion (kcal mol<sup>-1</sup>): (a) Methane  $418.0 \pm 3.5$ , fluoromethane  $409.0 \pm 4.0$ , difluoromethane  $389 \pm 3.5$ , trifluoromethane  $376.0 \pm 4.5$ , ethylene  $407.0 \pm 3.0$ : Graul, S. T.; Squires, R. R. *J. Am. Chem. Soc.* **1990**, 112, 2517. (b) Ethane  $420.1 \pm 2.0$ : DePuy, C. H.; Gronert, S.; Barlow, S. E.; Bierbaum, V. M.; Damrauer, R. *J. Am. Chem. Soc.* **1990**, 112, 2517. (c) Benzene  $401.7 \pm 0.5$ : Davico, G. E.; Bierbaum, V. M.; DePuy, C. H.; Ellison, G. B.; Squires, R. R. *J. Am. Chem. Soc.* **1995**, 117, 2590. (d) Acetylene  $378.0 \pm 0.7$ : Ervin, K. M.; Gronert, S.; Barlow, S. E.; Gilles, M. K.; Harrison, A. G.; Bierbaum, V. M.; DePuy, C. H.; Lin, W. C. *J. Am. Chem. Soc.* **1990**, 112, 5750. (e) Acetal  $389.8 \pm 0.9$ : Nimlos, M. R.; Soderquist, J. A.; Ellison, G. B. *J. Am. Chem. Soc.* **1989**, 111, 7675. (f) Hydrogen cyanide  $348.3 \pm 2.0$ : Bradforth, S. E.; Kim, E. H.; Arnold, D. W.; Neumark, D. M. *J. Chem. Phys.* **1993**, 98, 800.



**Figure 4.** Plot of gas-phase proton affinities of the corresponding anions versus BSSE-corrected  $D_e$  values. This linear relationship (slope = 0.0551, intercept = -23.6,  $r = 0.96$ ) can be used to estimate  $D_e$  values for C–H···O=C hydrogen bonds from C–H proton acidities. Hashed lines are placed  $\pm 0.5$  kcal mol<sup>-1</sup> from the line.



**Figure 5.** Gas-phase proton affinity of the anion (kcal mol<sup>-1</sup>) for the three types of hydrogen in DMF and the C <sup>$\alpha$</sup> –H hydrogen in protein model 5.

obey the relationship established in Figure 4. Calculations on the model compound, **5**, yielded a proton affinity of the anion of 355 kcal mol<sup>-1</sup> for the C <sup>$\alpha$</sup> –H hydrogen. As anticipated, this hydrogen has a significantly higher proton affinity of the anion than the *N*-methyl hydrogens in DMF (388 and 394 kcal mol<sup>-1</sup>). The result is fully consistent with known strengths of other carbon acids in that it predicts the C <sup>$\alpha$</sup> –H group (adjacent to a carbonyl and amide nitrogen substituent) to be a stronger acid than acetone (adjacent to one carbonyl substituent, proton affinity of the anion = 369 kcal mol<sup>-1</sup>) and a weaker acid than acetylacetonate (adjacent to two carbonyl substituents, proton affinity of the anion = 344 kcal mol<sup>-1</sup>).<sup>50</sup>

(50) Experimental gas-phase proton affinities of the corresponding anion (kcal mol<sup>-1</sup>): (a) Acetone  $369.1 \pm 2.1$  from: Bartmess, J. E.; Scott, J. A.; McIver, R. T., Jr. *J. Am. Chem. Soc.* **1979**, 101, 6047. (b) Acetylacetonate  $343.8 \pm 2.1$  from: Taft, R. W.; Bordwell, F. G. *Acc. Chem. Res.* **1988**, 21, 463.

From the predicted proton affinity of **5** and the relationship in Figure 4, we obtain an estimate of  $4.0 \pm 0.5$  kcal mol<sup>-1</sup> for the BSSE-corrected  $D_e$  value for the  $C^\alpha-H\cdots O=C$  hydrogen bond. This result can be compared with  $D_e$  values calculated, again at comparable levels of theory, for other types of hydrogen bonds traditionally associated with protein structure. The  $N-H\cdots O=C$  hydrogen bond in the *N*-methylacetamide dimer has a  $D_e$  value of 7.2 kcal mol<sup>-1</sup>,<sup>51</sup> the  $N-H\cdots O$  hydrogen bond in the *N*-methylacetamide–water dimer has a  $D_e$  value of 7.0 kcal mol<sup>-1</sup>,<sup>51</sup> and the  $OH\cdots O$  hydrogen bond in the water dimer has a  $D_e$  value of 5.0 kcal mol<sup>-1</sup>.<sup>26</sup> This comparison reveals that the  $C^\alpha-H\cdots O=C$  hydrogen bond, with an interaction energy greater than half of a  $N-H\cdots O=C$  hydrogen bond, is stronger than previously recognized.

Although the  $D_e$  values provide a rigorous point of comparison, they overestimate the strength of the interaction and  $\Delta H^{298}$  values should be used for comparison to experimental data. The  $\Delta H^{298}$  values corresponding to the  $D_e$  values above are -5.6 kcal mol<sup>-1</sup> for the *N*-methylacetamide dimer,<sup>51</sup> -5.3 kcal mol<sup>-1</sup> for the *N*-methylacetamide–water,<sup>51</sup> and -3.4 kcal mol<sup>-1</sup> for the water dimer.<sup>26</sup> Comparing these  $\Delta H^{298}$  values with the corresponding  $D_e$  values, we observe that zero-point energy and thermal corrections decrease the  $D_e$  value by 20 to 30% (similar decreases occur on applying the corrections to the three DMF dimers, Table 3). Assuming a decrease of 25%, we estimate a  $\Delta H^{298}$  value of  $-3.0 \pm 0.5$  kcal mol<sup>-1</sup> for the  $C^\alpha-H\cdots O=C$  hydrogen bond.

With this estimate of the  $C^\alpha-H\cdots O=C$  hydrogen bond strength, we can now assess the influence of the interactions observed in crystal structures of proteins. In these cases (Figure 1), both lone pairs of the oxygen atom are accepting a hydrogen donor. To a first approximation, we can assume that the energies of the two hydrogen bonds formed are additive, an assumption that is supported by the behavior of DMF dimers **2–4**. Thus, when coupled with the  $N-H\cdots O=C$  hydrogen bond in a  $\beta$ -sheet, the  $C^\alpha-H\cdots O=C$  hydrogen bond would strengthen the

(51) Dixon, D. A.; Dobbs, K. D.; Valentini, J. J. *J. Phys. Chem.* **1994**, *98*, 13435.

interaction from 5.3 to 8.3 kcal mol<sup>-1</sup>, an increase of 60%. Similarly, the hydrogen bonding pattern observed in the collagen triple helix, which involves the formation of three  $C^\alpha-H\cdots O=C$  hydrogen bonds, could be worth as much as 9 kcal mol<sup>-1</sup>! In conclusion, the present analysis suggests that  $C^\alpha-H\cdots O=C$  hydrogen bonds are of sufficient strength to make a significant contribution toward the stabilization of secondary and tertiary protein structure. In some cases, such as the collagen triple helix, the formation of  $C^\alpha-H\cdots O=C$  hydrogen bonds may actually play a critical role in determining the folded structure.

**Acknowledgment.** We thank a referee for suggesting a comparison to the solid state DMF dimer structures and we thank Prof. M. Sandström for providing Cartesian coordinates for the crystal structure of neat DMF. This research was supported under Grant No. 54679, Environmental Management Science Program, Office of Science and Technology, Office of Environmental Management, U.S. Department of Energy (DOE). Financial support for J. Garza was partially provided by CONACYT, México. The research described in this manuscript was performed at the W. R. Wiley Environmental Molecular Sciences Laboratory (EMSL), a national scientific user facility sponsored by the DOE Office of Biological and Environmental Research and located at Pacific Northwest National Laboratory. Pacific Northwest National Laboratory is operated for the DOE by Battelle. Some of the calculations were performed on the IBM SP computer in the Molecular Science Computing Facility of the EMSL.

**Supporting Information Available:** Cartesian coordinates for the MP2/DZP+diff optimized structures of **1–4**, comparison of the  $H\cdots O$  distances in **1–4** obtained at different computational levels, and a list of Cambridge Structural Database reference codes for each type of DMF dimer (ASCII). This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA993600A