J|A|C|S
ARTICLES

Published on Web 06/18/2009

Solvation and Hydrogen Bonding in Alanine- and Glycine-Containing Dipeptides Probed Using Solution- and Solid-State NMR Spectroscopy

Manasi P. Bhate, Jaie C. Woodard, and Manish A. Mehta*

Department of Chemistry and Biochemistry, Oberlin College, Oberlin, Ohio 44074

Received April 12, 2009; E-mail: manish.mehta@oberlin.edu

Abstract: The NMR chemical shift is a sensitive reporter of peptide secondary structure and its solvation environment, and it is potentially rich with information about both backbone dihedral angles and hydrogen bonding. We report results from solution- and solid-state ^{13}C and ^{15}N NMR studies of four zwitterionic model dipeptides, L-alanyl-L-alanine, L-alanyl-glycine, glycyl-L-alanine, and glycyl-glycine, in which we attempt to isolate structural and environmental contributions to the chemical shift. We have mapped hydrogen-bonding patterns in the crystalline states of these dipeptides using the published crystal structures and correlated them with ^{13}C and ^{15}N magic angle spinning chemical shift data. To aid in the interpretation of the solvated chemical shifts, we performed *ab initio* quantum chemical calculations to determine the low-energy conformers and their chemical shifts. Assuming low energy barriers to interconversion between thermally accessible conformers, we compare the Boltzmann-averaged chemical shifts with the experimentally determined solvated-state shifts. The results allow us to correlate the observed differences in chemical shifts between the crystalline and solvated states to changes in conformation and hydrogen bonding that occur upon solvation.

Introduction

For their size, small molecules still present many challenges. In biophysics, much of the current research effort is focused on macromolecules, and the experimental direction defined by grand challenges such as the protein-folding problem. Small molecules play an important role in these larger efforts, as they provide a fertile testing ground for experimental techniques and computational models, and as such, their investigations are more relevant than ever. Small molecules also pose intrinsically interesting questions. In solvated environments, their surface-area-to-volume ratio on a per-residue basis is disproportionately large, and thus, they provide an excellent vehicle for studying solvent interactions. For amino acids with small side chains, the reduced steric hindrance can allow multiple, thermally accessible conformations under ambient conditions. These features make for challenging spectroscopy. On the other hand, small molecules are amenable to high-level *ab initio* calculations that are computationally intractable for larger systems.[1−5] The interplay between theory and experiment enables the careful isolation and study of different factors that affect the observed conformational distribution, such as the solvent environment.

An oft-neglected, but relevant area of biophysics is molecular crowding.[6−9] Many studies of biomolecular structure either ignore or fail to take into account the thermodynamically nonideal environment of the cell interior, in which the solvent is far scarcer than in an ideal aqueous solution. In a highly heterogeneous medium such as the cell cytoplasm, excluded volume effects can cause activities of solute species to deviate significantly from their idealized values in dilute solutions.[10] These deviations have been shown to affect energetics, kinetics, equilibria, and transport.[11] When available volumes are reduced, the influence of the solvent on structure may be amplified to the point where there may be dramatic changes in the potential energy landscape as well as spectroscopic observables. In one study, it was shown that, in the presence of crowding agents under physiological conditions, the amount of secondary structure is increased in α-helical VlsE, and α/β- flavodoxin.[12] Other studies have shown equally large effects on the structure and function of nucleic acids in crowded conditions.[13] If investigations of small molecules are to be meaningful, the role of the solvent in helping to determine secondary structure should not only be better contextualized, but it should be embraced as

(1) Havlin, R.; Laws, D. D.; Bitter, H.; Sanders, L.; Wemmer, E.; Pines, A.; Oldfield, E. *J. Am. Chem. Soc.* **2001**, *123*, 10362–10369.
(2) Sun, H.; Oldfield, E. *J. Am. Chem. Soc.* **2004**, *126*, 4726–4734.
(3) Cheng, F.; Sun, H.; Zhang, Y.; Mukkamala, D.; Oldfield, E. *J. Am. Chem. Soc.* **2005**, *127*, 12544–12554.
(4) Wi, S.; Sun, H.; Oldfield, E.; Hong, M. *J. Am. Chem. Soc.* **2005**, *127*, 6451–6458.
(5) Mukkamala, D.; Zhang, Y.; Oldfield, E. *J. Am. Chem. Soc.* **2007**, *129*, 7385–7392.
(6) Zimmerman, S. B.; Minton, A. P. *Annu. Rev. Biophys. Biomolec. Struct.* **1993**, *22*, 27–65.
(7) Ellis, R. J. *Trends Biochem. Sci.* **2001**, *26*, 597–604.
(8) Ellis, R. J. *Curr. Opin. Struct. Biol.* **2001**, *11*, 114–119.
(9) Minton, A. P. *Curr. Biol.* **2006**, *16*, R269–R271.
(10) van den Berg, B.; Wain, R.; Dobson, C. M.; Ellis, R. J. *EMBO J.* **2000**, *19*, 3870–3875.
(11) Minton, A. P. *J. Biol. Chem.* **2001**, *276*, 10577–10580.
(12) Perham, M.; Stagg, L.; Wittung-Stafshede, P. *FEBS Lett.* **2007**, *581*, 5065–5069.
(13) Zimmerman, S. B.; Trach, S. O. *Nucleic Acids Res.* **1988**, *16*, 6309–6326.

a primary determinant of spectroscopic observables. In the case of NMR, that observable is the chemical shift.

Our current understanding of the connection between sequence identity, secondary structure, and chemical shift is now well-documented for globular proteins in aqueous solution.[14] With the widespread use of higher-field NMR spectrometers and careful attention to experimental details, such as referencing, chemical shift measurements are being mined for more information than ever. The past year-and-half has seen dramatic developments in the ability to generate blind protein structures from chemical shifts alone. Combining chemical shifts and molecular dynamics, Cavalli et al. successfully determined the structures of 11 proteins, containing up to 123 residues, with a resolution of 2 Å or better.[15] Vila et al. used a strategy focused entirely on $^{13}C_\alpha$ chemical shifts to predict the structure of the 20-residue all-β peptide BS2.[16] Their method deploys a combination of experimentally observed and quantum chemically calculated $^{13}C_\alpha$ shifts to derive torsional restraints for all the residues. Shen et al. have presented a yet new protocol for correlating six commonly observed chemical shifts with local structure.[17] Their procedure uses an empirically optimized approach to select structurally similar fragments from the Protein Data Bank together with ROSETTA assembly methods. The efficacy of their method was successfully demonstrated, in a blind manner, on 9 protein targets, as large as 15.4 kDa. These exciting developments hold the promise of speedier structure determination, but for all their simplicity, they demand a more thorough understanding of dependences of the chemical shift on important factors, such as hydrogen bonding, solvation, and structure.

In recent years, the term "microsolvation" has been much discussed in the literature.[18−20] It refers to the chemical environment in which a solute molecule is surrounded by a defined number of water molecules in a very specific configuration, typically held together by transient hydrogen bonds. In essence, microsolvation is the site-specific *and event-specific* solvation at the atomic level. While the ideas of site-specific solvation have been around for a long time, the conjunction with event specificity makes microsolvation a particularly cogent and timely phenomenon for study using modern spectroscopic and computational tools. For example, water molecules have been found to play a key role in bacteriorhodopsin's function.[21] Chen et al. have demonstrated that the specific geometrical arrangement of hydrogen-bonded water molecule networks around the active site of heme oxygenase plays a role in determining both the enzyme's rate of catalysis and the catalytic pathway.[22] Individual water molecules have been implicated in the mechanisms of other enzymes, including carbonic anhydrase and horseradish peroxidase.[23,24] Most recently, Born et al. reported the ability to detect the onset of collective network motions in model peptides, as a function of solvation, using terahertz spectroscopy.[25] Outside the biological realm, studies of the hydration environments around small organic molecules and ions show that hydration can alter molecular properties such as electron affinity and chemical reactivity.[26] There is some evidence to suggest that solvent interactions can alter the sterics around a molecule and affect the reaction mechanisms and rates in some S_N2 type reactions.[27] The role of solvent in deciding conformational preferences of organic molecules is now well-known and attributable to microsolvation.[28] It has also been shown that optimally positioned water molecules can act as catalysts for radical reactions in the atmosphere.[29] These experimental and computational realizations of microsolvation argue for deeper investigations of solvation mechanisms at the atomic level. The transient, dynamic nature of microsolvation makes its spectroscopic detection particularly challenging. Small molecules, and peptides in particular, provide an excellent point of entry.

Alanine, for example, is known to induce helices, and yet, in short peptides, it exhibits a far richer geometrical palette.[30−36] This intermediate regime, from 2−5 residues, poses many questions about the conformational preferences in different environments. Detailed studies of di- and tripeptides are beginning to add significant insights to structure, solvation, and spectroscopy, yet much remains unknown. Bour et al.'s combined computational and solution-state NMR study of L-alanyl-L-alanine presented evidence that the dipeptide occupies a single conformation in solution.[37] Siegrist et al. have shown, using high-resolution terahertz spectroscopy, the ability to detect weak intramolecular hydrogen bonds in the three different crystalline polymorphs of trialanine, while others have linked T_1-relaxation behavior with structural features.[38] There has been an equal amount of attention devoted to the behavior of small peptides in the gas phase.

We report here the results of a combined solution-state and solid-state NMR study, supported by *ab initio* calculations and X-ray diffraction, of four zwitterionic dipeptides, α-glycyl-glycine, glycyl-L-alanine, L-alanyl-glycine, and L-alanyl-L-alanine. We chose this system of dipeptides for three reasons. First, the crystalline form of each of the four dipeptides has

(14) Wishart, D. *J. Biomol. NMR* **2003**, *25*, 173–195.
(15) Cavalli, A.; Salvatella, X.; Dobson, C. M.; Vendruscolo, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9615–9620.
(16) Vila, J. A.; Aramini, J. M.; Rossi, P.; Kuzin, A.; Su, M.; Seetharaman, J.; Xiao, R.; Tong, L.; Montelione, G. T.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 14389–14394.
(17) Shen, Y.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 4685–4690.
(18) Aikens, C.; Gordon, M. *J. Am. Chem. Soc.* **2006**, *128*, 12835–12850.
(19) Blanco, S.; Lopez, J.; Alberto, L.; Alonso, J. *J. Am. Chem. Soc.* **2006**, *128*, 12111–12121.
(20) Bachrach, S. M. *J. Phys. Chem. A* **2008**, *112*, 3722–3730.
(21) Garczarek, F.; Gerwert, K. *Nature* **2006**, *439*, 109–111.
(22) Chen, H.; Moreau, Y.; Derat, E.; Shaik, S. *J. Am. Chem. Soc.* **2008**, *130*, 1953–1965.
(23) Isaev, A.; Scheiner, S. *J. Phys. Chem. B* **2001**, *105*, 6420–6426.
(24) Derat, E.; Shaik, S.; Rovira, C.; Vidossich, P.; Alfonso-Prito, M. *J. Am. Chem. Soc.* **2007**, *129*, 6346–6347.
(25) Born, B.; Weingartner, H.; Brundermann, E.; Havenith, M. *J. Am. Chem. Soc.* **2009**, *131*, 3752–3755.
(26) Yang, X.; Fu, Y. J.; Wang, X. B.; Slavicek, P.; Mucha, M.; Jungwirth, P.; Wang, L. S. *J. Am. Chem. Soc.* **2004**, *126*, 876–883.
(27) Mohamed, A. A.; Jensen, F. *J. Phys. Chem. A* **2001**, *105*, 3259–3268.
(28) Eliel, E. L.; Wilen, S. H.; Mander, L. N. *Sterochemistry of Organic Compounds*; Wiley-Interscience: New York, 1994.
(29) Vohringer-Martinez, E.; Hansmann, B.; Hernandez, H.; Francisco, J. S.; Troe, J.; Abel, B. *Science* **2007**, *315*, 497–501.
(30) Weir, a. F.; Lowrey, A. H.; Williams, R. W. *Biopolymers* **2001**, *58*, 577–591.
(31) Mu, Y. G.; Stock, G. *J. Phys. Chem. B* **2002**, *106*, 5294–5301.
(32) Wildman, K. A. H.; Lee, D.-K.; Ramamoorthy, A. *Biopolymers* **2002**, *64*, 246–254.
(33) Kennedy, R. J.; Walker, S. M.; Kemp, D. S. *J. Am. Chem. Soc.* **2005**, *127*, 16961–16968.
(34) Asakura, T.; Okonogi, M.; Nakazawa, Y.; Yamauchi, K. *J. Am. Chem. Soc.* **2006**, *128*, 6231–6238.
(35) Yamauchi, K. *J. Am. Chem. Soc.* **2006**, *128*, 6231–6238.
(36) Graf, J.; Nguyen, P. H.; Stock, G.; Schwalbe, H. *J. Am. Chem. Soc.* **2007**, *129*, 1179–1189.
(37) Bour, P.; Budesinsky, M.; Spirko, V.; Kapitan, J.; Sebestik, J.; Sychrovsky, V. *J. Am. Chem. Soc.* **2005**, *127*, 17079–17089.
(38) Siegrist, K.; Bucher, C. R.; Mandelbaum, I.; Walker, A. R. H.; Balu, R.; Gregurick, S. K.; Plusquellic, D. F. *J. Am. Chem. Soc.* **2006**, *128*, 5764–5775.

been characterized via single-crystal X-ray diffraction.[39−42] The crystal structures provide independent structural data against which we may study hydrogen bonding in the crystalline state, and they provide a basis of calibration of the solid-state NMR results. Second, these four dipeptides are small enough to be amenable to high-level *ab initio* calculations, which allow us to control and isolate factors that are not otherwise accessible via experiment. Finally, the systematic appearance of a side chain methyl group enables a lateral comparison of structural and spectroscopic trends across the four molecules. Comparison of chemical shifts across phases allows us to correlate structural features with solvation.

Experimental Methods

Sample Preparation. The four dipeptides, L-alanyl-L-alanine, L-alanyl-glycine, glycyl-L-alanine, and glycyl-glycine, were purchased from BaChem and Sigma, and used without further purification. Each dipeptide was dissolved in water and allowed to recrystallize by slow evaporation at room temperature. The crystals were ground into a fine powder and characterized by X-ray powder diffraction. Powder diffraction measurements were made on each of the crystalline NMR samples using a continuous scan from 5° to 50° in increments of 0.03° in 2θ using a Philips MPD-3040 X-ray powder diffractometer with a Cu-α source. X-ray structures for each of the four dipeptides were obtained from the Cambridge Crystallographic Structural Database. The software packages *CrystalMaker* and *CrystalDiffract* were used to compare the collected powder pattern for each dipeptide to that predicted from the known crystal structure in order to confirm that the correct crystalline polymorph had been grown for each dipeptide (See Supporting Information).

Following the powder diffraction check, the polycrystalline powder was packed in a 4 mm silicon nitride rotor for the solids MAS ^{13}C and ^{15}N NMR spectroscopy. We have observed a sizable and reproducible dependence of the ^{13}C and ^{15}N chemical shifts on concentration; thus, we have taken some care to work with solutions of known concentrations, sufficiently dilute to compare with *ab initio* calculations, yet with enough signal to enable the collection of spectra at natural abundance (especially for ^{15}N). A 10 mM aqueous solution (D_2O) of each dipeptide was prepared for solution-state NMR spectroscopy. A separate solution of 0.5% DSS solution in D_2O was prepared to serve as an external reference. We have also observed a temperature dependence of the chemical shifts; thus, all NMR spectra were collected with the sample temperature regulated at 25 °C.

NMR Spectroscopy. The spectra were collected on a system consisting of a 14.1 T magnet (600.377 MHz for ^1H; 150.987 MHz for ^{13}C; 60.84 MHz for ^{15}N; Magnex Scientific, Inc.), a 39-channel room temperature shim system (Resonance Research, Inc.), and a three-channel console (TecMag, Inc.). A triple-resonance (H/X/Y) narrow-bore XC4 MAS probe (Doty Scientific, Inc.), a 5 mm double-resonance direct-detect broadband probe (Nalorac-Varian, Inc.), and a 5 mm triple-resonance indirect-detect probe (Venus NMR, Livermore, CA) were used to collect the spectra. The MAS rotor spin rate was controlled to ±1 Hz using a home-built controller. ^{13}C Cross-Polarization Magic Angle spinning (CP/MAS) spectra were acquired at spin rates of 3, 4, and 5 kHz to determine the carbonyl chemical shift anisotropy (CSA). C_α CSAs required slower spinning; thus, another set of ^{13}C spectra were collected at spin speeds between 1190 and 1300 Hz. ^{13}C CP/MAS spectra were collected using a 2.5 ms cross-polarization contact time, a 20% linear power ramp on the carbon channel, and a recycle time of 5 s. ^{15}N CP/MAS spectra were collected at a spin speed of 1 kHz with a 3.5 ms contact time, 20% power ramp on the nitrogen channel, and a recycle time of 10 s. High power broadband proton decoupling (SPINAL-64) was used during the acquisition. Liquids ^{13}C spectra were collected using a simple one-pulse experiment and low-power proton decoupling (WALTZ-16). Typically, 256 scans were collected for solids and liquids ^{13}C spectra to obtain a good signal-to-noise ratio. ^{15}N chemical shifts were measured using a gs-HMBC experiment by detecting correlations to the nearby H_α. All spectra were processed with one zero-fill and no apodization. The line widths in the solid state, full width at half-maximum (fwhm), were determined by fitting the peaks to a Lorentzian line shape. The solid-state spectra were referenced externally to neat nitromethane for ^{15}N and neat tetramethylsilane (TMS) for ^{13}C. The ^{13}C shifts were rereferenced to 0.5% DSS as detailed by Morcombe and Zilm.[43] The solution-state spectra were externally referenced to 0.5% DSS in D_2O for ^{13}C and neat nitromethane for ^{15}N. A correction was applied to the ^{15}N measured chemical shifts to take into account the difference in magnetic susceptibility between D_2O and nitromethane (see Supporting Information). An exceptionally low drift rate of the magnetic field (less than 0.1 Hz/h for ^{13}C) obviated the need to use a deuterium field frequency lock and made possible straightforward external referencing. The principle components of the chemical shift tensor were determined using a Herzfeld-Berger analysis of the spinning side bands in the MAS spectra.[44,45] The software package *MatLab* was used to numerically integrate the peaks in the solid-state MAS spectra to determine the relative intensities of the various spinning sidebands.

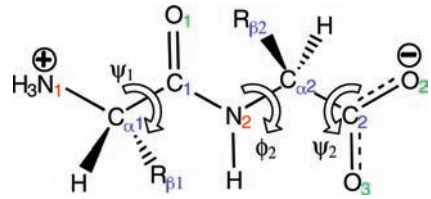

Figure 1. The labeling scheme used for the carbon and nitrogen sites among the dipeptides. The three backbone torsion angles (ψ_1, ϕ_2, ψ_2) used to characterize the structures are shown.

***Ab Initio* Calculations.** All *ab initio* calculations were performed using *Gaussian03* via the *WebMO* graphical user interface.[46] The B3LYP density functional was used for all calculations, and water was modeled as the solvent using the Polarizable Continuum Model (PCM). Geometry optimizations, single-point energy calculations, and NMR chemical shift calculations were conducted on each of the four dipeptides using a 6-311+G(d,p) basis set.

An initial geometry optimization was performed on each dipeptide in PCM water, starting from the geometry of the known crystal structure. Next, a series of all-parameter geometry optimizations were performed, to comprise a crude systematic search for lower-energy optimized structures. Input geometries for these optimizations sampled values of ψ_1 and ϕ_2 (see Figure 1 for the labeling scheme), at 90° increments, for a total of 16 input geometries (all other geometrical parameters were taken from the initial optimized structure). Additional optimizations were performed using starting geometries similar to those of low-energy converged structures, to arrive at a more complete picture of the low energy structures. Converged structures were grouped into distinct families according to energy and geometry, assuming each family represents a single minimum on the potential energy surface. For converged structures with energy within 4RT (about 10 kJ/mol) of the lowest energy, chemical shift calculations were carried out, using the GIAO method. Calculated chemical shifts were averaged using Boltzmann weighting at 298 K, including only the lowest-energy structure from

(39) Koch, J.; Germain, G. *Acta Crystallogr., Sect. B* **1970**, *26*, 410–417.
(40) Fletterick, R. J.; Hughes, R. *J. Phys. Chem.* **1971**, *75*, 918–922.
(41) Wang, A. H.-J.; Paul, I. C. *Cryst. Struct. Comm* **1979**, *8*, 269–273.
(42) Kameda, T.; Takeda, N.; S.; Ando; Ando, I.; Hashizume, D.; Ohashi, Y. *Biopolymers* **1998**, *45*, 333–339.
(43) Morcombe, C.; Zilm, K. *J. Magn. Reson.* **2003**, *162*, 479–486.
(44) Herzfeld, J.; Berger, A. *J. Chem. Phys.* **1980**, *73*, 6021–6030.
(45) Eichele, K.; Easylishen, R. E.; Herzfeld-Berger Analysis, 1.5; Dalhousie University & Universitat Tubingen, 2006.
(46) Frisch, M. J. T. et al.; Gaussian, Inc., Wallingford, CT, 2004.

Table 1. Geometrical and Chemical Shift Data for the Computed and Crystalline Low-Energy Conformers of the Four Dipeptides[a]

	ID	energy (kJ/mol)	ψ_1	ϕ_2	ψ_2	C_1═O	C_2═O	$C_{\alpha 1}$	$C_{\alpha 2}$	$C_{\beta 1}$	$C_{\beta 2}$	N_1	N_2
GG	A	0.0	158.6	−179.2	179.3	171.21	181.14	48.88	49.88	--	--	−382.17	−279.33
	B	0.7	−154.3	179.4	179.6	171.32	181.09	49.35	49.87	--	--	−381.48	−278.20
	C	4.4	159.5	91.7	2.6	172.46	183.29	49.13	49.82	--	--	−382.02	−278.93
	D	4.8	−160.2	−91.4	176.3	172.48	183.30	49.13	49.87	--	--	−381.93	−279.05
	E	5.0	−159.2	82.9	1.4	172.74	182.89	48.73	50.13	--	--	−382.74	−279.77
	F	5.1	159.4	−90.8	178.6	172.52	183.23	48.96	49.53	--	--	−382.41	−279.27
	G	5.6	178.8	−89.6	178.1	173.31	183.63	49.27	49.67	--	--	−381.67	−282.13
	H	5.7	−179.3	89.1	2.4	173.41	183.59	49.25	49.59	--	--	−381.68	−282.12
	avg.	--	--	--	--	171.72	181.79	49.08	49.85	--	--	−381.94	−279.20
	Crystal[42]		152.4	155.1	10.1	169.3	175.0	41.9	46.5	--	--	−353.4	−263.5
GA	A	0.0	−158.7	−154.1	165.3	170.53	185.85	49.00	59.21	--	23.53	−382.33	−266.49
	B	0.1	160.0	−153.8	165.0	170.66	185.87	48.87	59.79	--	24.02	−382.30	−266.82
	C	4.2	−159.0	−97.7	159.1	170.99	187.78	49.12	59.11	--	23.92	−381.95	−263.86
	D	4.6	159.7	−95.4	160.8	171.51	187.68	49.00	59.04	--	23.83	−382.37	−264.01
	avg.	--	--	--	--	170.69	186.13	48.95	59.43	--	23.79	−382.29	−266.25
	Crystal[41]		−163.5	−78.4	150.0	168.0	183.5	41.0	54.9	--	18.1	−352.9	−252.8
AG	A	0.0	149.1	178.8	179.8	175.96	181.22	60.39	49.91	21.79	--	−366.54	−278.90
	B	4.4	152.2	92.9	1.9	177.35	183.33	60.64	49.62	22.16	--	−366.63	−278.76
	C	6.8	146.9	−94.7	176.7	176.58	183.31	60.63	49.11	21.40	--	−366.53	−277.89
	avg.	--	--	--	--	176.18	181.62	60.43	49.83	21.82	--	−366.55	−278.83
	Crystal[39]		162.8	−72.7	155.3	172.5	180.9	51.3	46.6	21.8	--	−339.8	−269.7
AA	A	0.0	148.6	−154.8	164.5	175.41	185.89	60.93	59.77	21.69	23.83	−366.39	−266.04
	MD sims.[37]		145	−150	--	--	--	--	--	--	--	--	--
	Crystal[40]		165.4	−113.0	102.9	171.8	179.6	50.8	53.2	19.7	20.5	−340.6	−257.1

[a] The Boltzmann-weighted average chemical shifts for each compound are given as the averages. Energies of each conformer for that dipeptide, up to 7.5 kJ/mol (3RT), are reported relative to the lowest energy structure. Other details appear in the Supporting Information. Reference for each crystal structure is given in the "crystal" entry.

each distinct family of structures (described previously). The lowest energy structure was assigned a weighting factor of 1, and so forth. Carbon chemical shifts were referenced to DSS in PCM water, and nitrogen shifts were referenced to nitromethane in PCM nitromethane. A susceptibility correction was applied to the nitrogen shifts. The energy and geometry data for the lowest energy structures, within 3RT, are given in Table 1. Cartesian coordinates and other information are presented in the Supporting Information.

Results and Discussion

Our analysis is based on a detailed comparison of crystalline and liquid state chemical shifts and a comparison between the reported crystal structures and the calculated solvated structures. The nomenclature used to label various sites in the dipeptides is shown in Figure 1 and will be used throughout this paper. Our larger goal is to quantify the dependence of chemical shift on secondary structure, sequence identity, and solvation in this system of four dipeptides by focusing on ^{13}C and ^{15}N chemical shifts. In the crystalline state, the structures are established via X-ray diffraction, where the molecule serves as its own solvent. Here, we can establish an unequivocal connection between chemical shift and structure. When the dipeptides are solvated, many more conformational degrees of freedom are made accessible, accompanied by a change in the environment. In three of the four compounds, several low-energy conformations become thermally accessible. We hypothesize that, if the barriers to interconversion are low relative to kT and the interconversion rapid on the NMR time scale, the observed chemical shift will be a Boltzmann-averaged value over all the thermally accessible structures. The lowest energy calculated structures are shown alongside the crystal structures in Figure 2, and the remainder in the Supporting Information.

Crystal Structures. The crystal structures of the four dipeptides are very different. The α-polymorph of GlyGly has four equivalent molecules in its unit cell and a lattice structure that resembles an antiparallel β-sheet. This is important because GlyGly is the only dipeptide in our study in which the carbonyl oxygen forms a hydrogen bond in the lattice. In crystalline peptides, the carbonyl and carboxyl groups prefer to accept two hydrogen bonds per oxygen atom, while the amine and amide nitrogens donate only one hydrogen bond. Because of this, there are more hydrogen bond acceptor sites than donor sites per molecule, leading to a *proton deficiency* that is satisfied via three-center/bifurcated hydrogen bonds. As a stronger acceptor, the carboxyl oxygen generally forms hydrogen bonds with the amines while the carbonyl oxygen remains unbonded. The lattices in GlyAla, AlaGly, and AlaAla are all structured so that both the amine and amide hydrogens are oriented toward the carboxyl oxygens, while the internal carbonyl group resides in a hydrophobic cavity. These structural differences between GlyGly and the rest of the dipeptides are important to keep in mind while analyzing the NMR data. A more detailed discussion of the four crystal structures is presented in the Supporting Information.

Solid-State NMR and Hydrogen Bonding. Internal and Terminal Carbonyl Carbons. The electronic environment around these sp^2 carbons is highly anisotropic, as reflected in the chemical shift anisotropy (CSA). The principal components of the CSA tensors of the internal carbonyl carbons vary significantly with secondary structure and hydrogen bonding characteristics of the peptide backbone.[47,48] Of the three tensor components, δ_{22} is oriented along the C═O bond and is most sensitive to changes in the local hydrogen bonding environment. Studies of carbonyl tensors in several small peptides report that δ_{22} shifts downfield as the C═O···H−N distance decreases.[49−51]

(47) Oas, T.; Hartzell, C.; McMahon, T.; Drobny, G.; Dahlquist, F. *J. Am. Chem. Soc.* **1987**, *109*, 5956–5962.
(48) Wei, Y.; Lee, D.-K.; Ramamoorthy, A. *J. Am. Chem. Soc.* **2001**, *123*, 6118–6126.
(49) Ando, S.; Ando, I.; Shoji, A.; Ozaki, T. *J. Am. Chem. Soc.* **1988**, *110*, 3380–3386.
(50) Asakawa, N.; Kuroki, S.; Kurosu, H.; Ando, I.; Shoji, A.; Ozaki, T. *J. Am. Chem. Soc.* **1992**, *114*, 3261–3265.
(51) Takeda, N.; Ando, I. *Biopolymers* **1999**, *50*, 61–69.

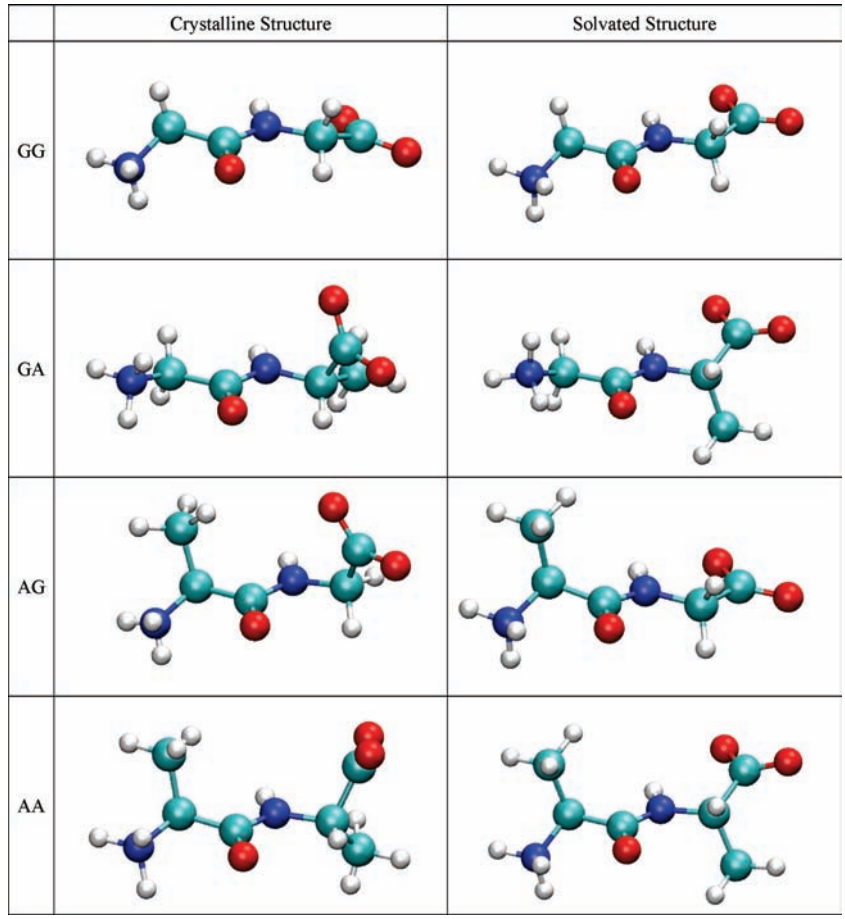

Figure 2. Crystalline conformations (left) and the lowest energy conformations (right) of the four dipeptides. Further details are given in the Supporting Information.

Table 2. $^{13}C{=}O$ Chemical Shifts (in ppm) and Derived Parameters in the Crystalline and Solvated States for the Internal ($C_1{=}O$) and Terminal Carbonyl ($C_2{=}O$) Carbons of the Four Dipeptides[a]

		fwhm (ppm)	δ_{11}	δ_{22}	δ_{33}	span (Ω) (ppm)	skew (κ)	δ_{iso} solid	δ_{iso} liquid	δ_{iso} (*ab initio* PCM)	$\Delta\delta_{iso}$ (liquid−solid)
GG	$C_1{=}O$	0.45	239.7	177.2	90.9	148.8	0.16	169.3	169.65	171.72	+0.4
	$C_2{=}O$	0.35	240.1	176.9	108.0	132.1	0.04	175.0	179.19	181.79	+4.2
GA	$C_1{=}O$	0.35	247.7	165.6	90.7	157.0	−0.05	168.0	168.87	170.69	+0.9
	$C_2{=}O$	0.20	240.8	197.0	112.8	128.0	0.32	183.5	182.72	186.13	−0.8
AG	$C_1{=}O$	0.38	246.8	178.9	91.8	155.0	0.12	172.5	173.34	176.18	+0.8
	$C_2{=}O$	0.19	242.5	186.4	113.9	128.6	0.13	180.9	179.05	181.62	−1.9
AA	$C_1{=}O$	0.35	251.4	171.4	92.5	158.9	−0.01	171.8	172.51	175.41	+0.7
	$C_2{=}O$	0.12	241.7	185.6	111.6	130.1	0.14	179.6	182.48	185.89	+2.9

[a] Chemical shift anisotropy parameters were derived at 4 kHz MAS. All the shifts are referenced to 0.5% DSS in D_2O. The line widths for the crystalline samples, reported as full width at half-maximum (fwhm), are from unapodized CP/MAS spectra, with each line fitted to a pure Lorenzian line shape.

When the hydrogen bond partner changes from a neutral amide (N−H) to a charged amine ($^+$N−H), this trend is reversed. Studies of terminal carboxyl carbon anisotropies are scarce because their CSAs are of limited utility in determining protein conformations. The orientation of the principal components around the carboxyl carbon reflects the local symmetry: δ_{11} and δ_{22} are in the O−C−O plane, and δ_{33} is perpendicular to it. Gu et al. have correlated the protonation state of the carboxyl group with changes in δ_{11} and δ_{22}.[52] They have also shown that, for protonated acids, an increase in hydrogen bonding strength leads to a decrease in shielding around the acidic carbon, whereas in deprotonated acids, stronger hydrogen bonding leads to more shielded values of δ_{22}.[53]

Establishing reliable correlations between hydrogen bond strength and CSA parameters is difficult here because each of the two carboxyl oxygens has multiple hydrogen bond partners, and the hydrogen-bonding geometries are not always linear. There are, however, some noteworthy trends (Table 2). Solid-state NMR data show that, across the four compounds, the carbonyl carbon line width is consistently larger than the carboxyl carbon line width suggesting that the carbonyl carbon has a shorter T_2. The span of the internal carbonyl carbon CSA

(52) Gu, Z.; McDermott, A. *J. Am. Chem. Soc.* **1993**, *115*, 4282–4285.

(53) Gu, Z.; Zambrano, R.; McDermott, A. *J. Am. Chem. Soc.* **1994**, *116*, 6368–6372.

Table 3. $^{13}C_\alpha$ Chemical Shifts and Anisotropy Parameters in the Crystalline and Solvated States for the N-Terminal Alpha Carbon ($C_{\alpha1}$), the C-Terminal Alpha Carbon ($C_{\alpha2}$), and the Side-Chain Methyl Carbons ($C_{\beta1}$ and $C_{\beta2}$)[a]

		fwhm (ppm)	δ_{11}	δ_{22}	δ_{33}	span (Ω) (ppm)	skew (κ)	δ_{iso} solid	δ_{iso} liquid	δ_{iso} (*ab initio* PCM)	$\Delta\delta_{iso}$ (liquid−solid)
GG	$C_{\alpha1}$	0.43	54.6	48.1	23.1	31.5	0.58	41.9	43.20	49.08	+1.3
	$C_{\alpha2}$	0.43	57.2	56.9	25.3	31.9	0.98	46.5	45.93	49.85	−0.6
GA	$C_{\alpha1}$	0.29	55.9	48.3	18.7	37.3	0.59	41.0	43.15	48.95	+2.2
	$C_{\alpha2}$	0.32	76.3	53.8	34.6	41.7	−0.08	54.9	53.88	59.43	−1.0
	$C_{\beta2}$	0.14	31.6	21.0	1.8	29.8	0.29	18.1	19.83	23.79	+1.7
AG	$C_{\alpha1}$	0.28	62.5	62.0	29.5	33.0	0.97	51.3	51.92	60.43	+0.6
	$C_{\alpha2}$	0.35	71.5	45.1	23.2	48.3	−0.09	46.6	46.01	49.83	−0.6
	$C_{\beta1}$	0.18	38.2	22.6	4.6	33.6	0.07	21.8	19.00	21.82	−2.8
AA	$C_{\alpha1}$	0.24	62.3	61.6	28.4	33.9	0.96	50.8	51.76	60.93	+1.0
	$C_{\alpha2}$	0.30	66.6	59.1	33.8	32.7	0.56	53.2	54.00	59.77	+0.8
	$C_{\beta1}$	0.15	32.8	21.3	5.1	27.7	0.17	19.7	19.06	21.69	−0.6
	$C_{\beta2}$	0.14	33.2	23.2	5.1	28.1	0.29	20.5	19.76	23.83	−0.7

[a] Chemical shift anisotropy parameters for the C_α and methyl carbons were derived at 1200 Hz MAS. All the shifts are referenced to 0.5% DSS in D_2O and given in ppm.

Table 4. ^{15}N Chemical Shifts and Anisotropy Parameters in the Crystalline and Solvated States for the Terminal Amine (N_1-H_3) and Amide (N_2-H) Nitrogens of the Four Dipeptides[a]

		fwhm (ppm)	δ_{11}	δ_{22}	δ_{33}	span (Ω) (ppm)	skew (κ)	δ_{iso} solid	δ_{iso} liquid	δ_{iso} (*ab initio* PCM)	$\Delta\delta_{iso}$ (liquid−solid)
GG	N_1-H_3	0.60	--	--	--	--	--	−353.4	−354.4	−381.94	−1.0
	N_2-H	0.34	−159.6	−296.8	−334.0	174.4	−0.58	−263.5	−265.7	−279.20	−2.2
GA	N_1-H_3	0.33	--	--	--	--	--	−352.9	−354.4	−382.29	−1.5
	N_2-H	0.20	−143.3	−289.4	−325.8	182.5	−0.60	−252.8	−251.3	−266.25	+1.5
AG	N_1-H_3	0.34	--	--	--	--	--	−339.8	−340.9	−366.55	−1.1
	N_2-H	0.31	−169.7	−313.2	−326.4	156.7	−0.83	−269.7	−265.5	−278.83	+4.2
AA	N_1-H_3	0.24	--	--	--	--	--	−340.6	−341.0	−366.39	−0.4
	N_2-H	0.28	−162.8	−294.7	−313.8	151.0	−0.75	−257.1	−252.0	−266.04	+5.1

[a] Chemical shift anisotropy parameters for the amide nitrogens were derived at 1 kHz MAS; parameters for the terminal amine nitrogens are not reported, as the anisotropy was insufficiently large to determine using the spinning side band approach. All the shifts are referenced to neat CH_3NO_2 and corrected for differences in magnetic susceptibility.

is ~25 ppm larger than that of the terminal carboxyl, indicating that the chemical shielding environment is more heterogeneous around the internal carbonyl compared to the terminal carboxyl. The average span for the internal carbonyl across the four dipeptides is 155.9 ± 4.4 ppm and for the carboxyl carbon is 129.7 ± 1.8 ppm, and the span does not correlate with the identity of the amino acid residue. With one exception, the skew parameter for all of the carbonyl and carboxyl carbon shift tensors is close to zero. This is consistent with the observation that δ_{22} is generally close to δ_{iso}, and indicates these carbon chemical shift tensors are nearly centrosymmetric. The exception is the terminal carboxyl carbon in GlyAla, which has an unusually large value for δ_{22} resulting in a skew of 0.32. Finally, while δ_{11} and δ_{22} change across the four dipeptides, δ_{33} stays relatively constant at 91.5 ± 0.8 ppm for the carbonyl and 111.6 ± 2.6 ppm for carboxyl carbons. Hartzell et al. showed that δ_{33} is perpendicular to the C−O plane, whereas δ_{11} and δ_{22} lie within the plane and thus are most affected by changes in lattice structure and hydrogen bonding patterns.[54]

C_α Carbons. The C_α carbon chemical shift is more sensitive to structure and residue identity than to the hydrogen-bonding environment because C_α usually does not participate in traditional hydrogen bonds. The span of the C_α anisotropy has been correlated with the peptide backbone conformation in protein systems.[55,56] This correlation is rooted in the different orientations of the principal tensor components in different structural forms. The orientation of the C_α tensor has been determined experimentally for valine and theoretically for a host of other amino acid residues.[57] It has also been shown that, in crystalline alanyl compounds, while the spans of the C_α carbons are independent of the α-helix/β-sheet structure, the orientation of the tensor changes dramatically between these two forms.[58,1] We used slow MAS to determine the chemical shift anisotropies around both the C_α and C_β carbons for each of the four dipeptides (Table 3). The data show that the C_α carbon CSA spans a range from 28 to 48 ppm in the four dipeptides. In our system, the skew of the N-terminal C_α tensor can be correlated with residue identity—glycyl $C_{\alpha1}$ has a skew value of about 0.58, whereas in alanyl $C_{\alpha1}$, it is close to 1.0. The C-terminal $C_{\alpha2}$ skew parameter shows no such trend, likely because the strongly charged carboxyl group overshadows smaller perturbations in electron density due to side chain identity.

Amide Nitrogen. Previous studies of the orientation of the amide chemical shift tensor in the molecular frame of L-alanyl-L-alanine have reported that δ_{11} and δ_{33} are within the peptide plane.[54,59] In some instances, δ_{33} is displaced out of the plane. The variability of the amide ^{15}N CSA principal values has been attributed to structural differences, lattice-dependent variations, and differences in hydrogen bonding.[60] The amide δ_{22} is especially sensitive to differences in polypeptide structure,[61] but unlike C_α, the orientation of the amide nitrogen CSA tensor tends not to change drastically in response to structure and

(54) Hartzell, C. J.; Whitfield, M.; Oas, T. G.; Drobny, G. *J. Am. Chem. Soc.* **1987**, *109*, 5966–5969.
(55) Tjandra, N.; Bax, A. *J. Am. Chem. Soc.* **1997**, *119*, 9576–9577.
(56) Hong, M. *J. Am. Chem. Soc.* **2000**, *122*, 3762–3770.
(57) Birn, J.; Poon, A.; Mao, Y.; Ramamoorthy, A. *J. Am. Chem. Soc.* **2004**, *126*, 8529–8534.
(58) Heller, J.; Laws, D. D.; Tomaselli, M.; King, S.; Wemmer, E.; Pines, A.; Oldfield, E. *J. Am. Chem. Soc.* **1997**, *119*, 7827–7831.
(59) Lee, D. K.; Wittebort, R. J.; Ramamoorthy, A. *J. Am. Chem. Soc.* **1998**, *120*, 8868–8874.
(60) Oas, T. G.; Hartzell, C.; Dahlquist, F.; Drobny, G. *J. Am. Chem. Soc.* **1987**, *109*, 5962–5966.
(61) Shoji, A.; Ozaki, T.; Fujito, T.; Deguchi, K.; Ando, S.; Ando, I. *J. Am. Chem. Soc.* **1990**, *112*, 4693–4697.

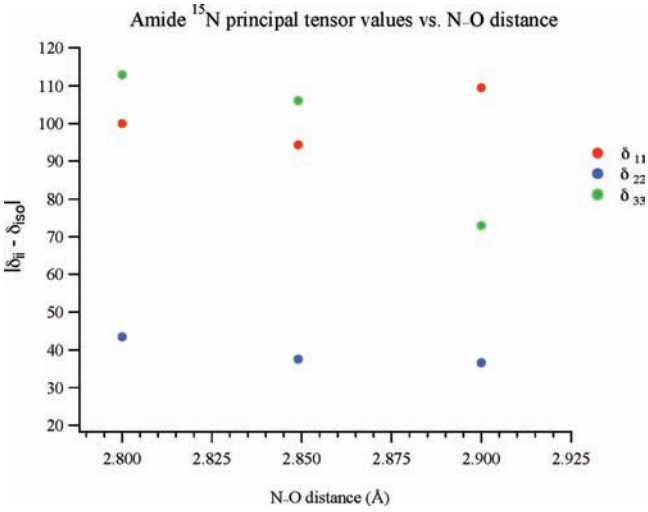

Figure 3. Principal values of the ^{15}N shift tensor for three of the dipeptides, as a function of the donor−acceptor hydrogen bond distance. The shortest distance (2.80 Å) is for AlaGly, the next (2.85 Å) for AlaAla, and the longest (2.90 Å) for GlyAla. The distances are taken from the reported X-ray crystal structures. Data for GlyGly are not shown in this plot, as its hydrogen bonding arrangement is very different from its counterparts in the other three dipeptides.

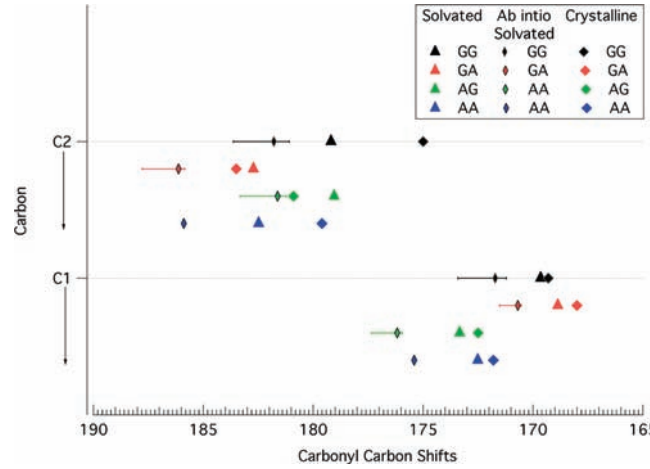

Figure 4. A graphical depiction of experimentally measured ^{13}C=O chemical shifts for the solvated dipeptides (triangles), crystalline peptides (squares), and the Boltzmann-weighted calculated shifts (diamonds). The horizontal bars on the computed shifts represent the range of shifts for that site across all the low-energy structures below 3RT. C1 is the internal carbonyl carbon, and C2 is the terminal carboxyl carbon.

hydrogen bonding.[62,63] CSA parameters for the amide nitrogen are presented in Table 4.

It has been shown that, in β-sheet-like geometries, the glycine amide nitrogen span is ~181 ppm, while in α-helical structures, it is ~161 ppm.[64] Extrapolation of this correlation to our system would suggest that GlyGly and possibly GlyAla exist in a β-sheet like structure in the crystalline state, while AlaGly and AlaAla exist in a more α-helix like geometry. Such correlations must, of course, be viewed cautiously because the model compounds used to establish these trends were not zwitterionic.

We have also correlated hydrogen bond strength with the magnitudes of the principal components of the amide nitrogen CSA tensor to learn which component is most sensitive to changes in hydrogen bonding. The donor−acceptor (N···O) distance reported in the crystal structure is a reasonable measure of hydrogen bond strength. Figure 3 plots $|\delta_{iso} - \delta_{ii}|$ for each of the three principle tensor values against the N···O bond length in angstroms. The data show that δ_{22} and δ_{33} both decrease as the hydrogen bond length increases, whereas δ_{11} shows no particular trend. This is a surprising result since tensor orientation data reports that δ_{11} is directly pointed toward the hydrogen bond. One would therefore expect δ_{11} to be most perturbed by changes in hydrogen bond strength. Understanding exactly how electron density affects shielding and chemical shift in these hydrogen-bonded systems would require extensive *ab initio* calculations in the solid state, which were beyond the scope of this project.

Trends in ^{13}C and ^{15}N NMR Chemical Shifts upon Solvation. Internal Carbonyl Carbon. The internal carbonyl carbon chemical shift moves downfield upon solvation for all four compounds (Figure 4). This change reflects a gain in hydrogen bonds, because upon solvation, proton deficiency is lifted and the internal carbonyl oxygen can accept two hydrogen bonds from water. It is interesting that, upon solvation, the carbonyl carbon in GlyGly is deshielded 0.4−0.5 ppm less than the other three dipeptides. This difference is explained by the crystal structure. Crystalline GlyGly has an antiparallel β-sheet like structure, which allows the carbonyl group to form one weak hydrogen bond with an amide hydrogen. In the other compounds, the carbonyl group has no hydrogen bond partners in the crystalline state. In GlyAla, for example, the carbonyl is tucked away in a hydrophobic pocket with no intermolecular hydrogen bond donors within a 4 Å radius around it. Upon solvation, all the intermolecular hydrogen bonds are likely replaced by hydrogen bonds with water. The internal carbonyl in GlyGly thus effectively gains one hydrogen bond upon solvation, whereas in the other dipeptides, it gains two hydrogen bonds as it goes from the crystalline to the solvated state. The data are consistent with a downfield shift of ~0.4 ppm per additional hydrogen bond, and the effect is approximately additive.

The obvious caveat here is that we cannot be certain whether the observed changes in chemical shift reflect changes in structure or changes in the hydrogen-bonding environment. *Ab initio* calculations of the NMR chemical shifts in the solvated conformations of these dipeptides reveal that the internal carbonyl carbon shift is much more sensitive to hydrogen bonding than to structural changes (see Supporting Information). This result, together with the fact that the magnitude of the observed change is similar for the four compounds, suggests that the observed changes in the chemical shift of the carbonyl carbon primarily reflect changes in its hydrogen-bonding environment. Previous studies of capped peptides have suggested preferential solvation at different carbonyl carbons.[65] Our results imply that the ^{13}C chemical shift of the internal carbonyl carbon can potentially be used to study peptide solvation because it is very sensitive to changes in the hydrogen-bonding environment.

Terminal Carboxyl Carbon. Unlike the internal carbonyl carbon, the terminal carboxyl carbon chemical shift does not exhibit a systematic trend upon solvation (Figure 4). For the

(62) Brender, J.; Ramamoorthy, A. *J. Am. Chem. Soc.* **2001**, *123*, 914–922.
(63) Poon, A.; Birn, J.; Ramamoorthy, A. *J. Am. Chem. Soc.* **2004**, *108*, 16577–16585.
(64) Chekmenev, E.; Zhang, Q.; Waddell, K.; Mashuta, M.; Wittebort, R. *J. Am. Chem. Soc.* **2004**, *126*, 379–384.
(65) Mehta, M. A.; Fry, E. A.; Eddy, M. T.; Dedeo, M. T.; Anagnost, A. E.; Long, J. R. *J. Phys. Chem. B* **2004**, *108*, 2777–2780.

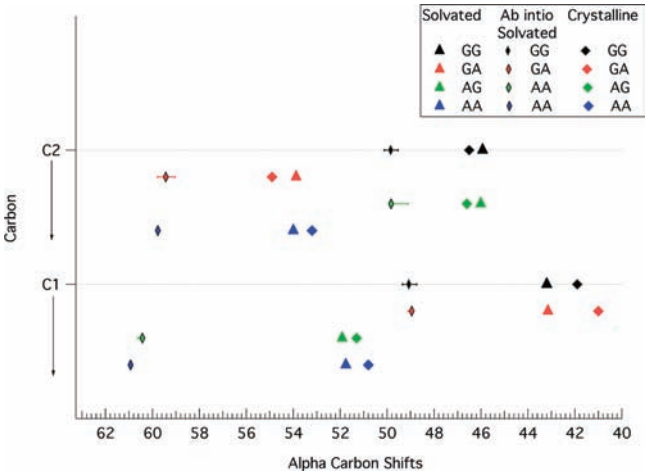

Figure 5. A graphical depiction of experimentally measured $^{13}C_\alpha$ chemical shifts for the solvated dipeptides (triangles), crystalline peptides (squares), and the Boltzmann-weighted calculated shifts (diamonds). The horizontal bars on the computed shifts represent the range of shifts for that site across all the low-energy structures below 3RT. C1 is $C_{\alpha1}$ and C2 is $C_{\alpha2}$.

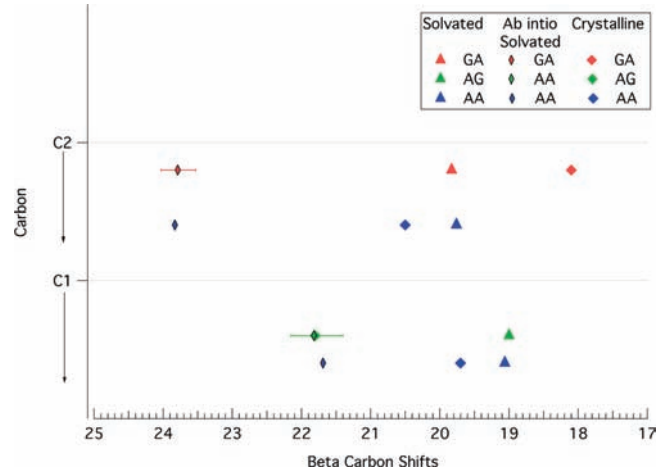

Figure 6. A graphical depiction of experimentally measured $^{13}C_\beta$ chemical shifts for the alanine residues of the solvated dipeptides (triangles), crystalline peptides (squares), and the Boltzmann-weighted calculated shifts (diamonds). The horizontal bars on the computed shifts represent the range of shifts for that site across all the low-energy structures below 3RT. C1 is $C_{\beta1}$ and C2 is $C_{\beta2}$, where applicable.

carboxyl group, solvation entails breaking three $C{=}O^-\cdots H{-}N^+$ and one $C{=}O^-\cdots H{-}N$ hydrogen bond, and forming four new $C{=}O^-\cdots H{-}O_w$ bonds. Since $^+N{-}H$ is a stronger hydrogen bond donor than water, the carboxyl carbon chemical shift is expected to move upfield in the solvated state. However, this is not the observed trend. *Ab initio* geometry optimizations indicate that solvation often involves a significant repositioning of the terminal carboxyl group. In general, dipeptides adopt a more extended structure in the solvated state compared to the crystalline state. Thus, unlike the internal carbonyl shift, the observed change in the terminal carboxyl carbon shift is a convolution of local structural changes and changes in the hydrogen bonding environment. To parse these effects apart, we calculated the ^{13}C NMR chemical shifts for both the reported crystalline conformation and the computationally calculated lowest energy solvated conformation of the dipeptides in both a vacuum and a PCM environment (see Supporting Information).

In the case of GlyGly and AlaAla, the computational ^{13}C NMR chemical shifts suggest that the terminal carboxyl group is structurally repositioned in the solvated conformation. The downfield shift affected by this structural change overpowers the shielding effect of replacing an $N^+{-}H$ donor with water. In the case of the other two dipeptides, AlaGly and GlyAla, computational data indicate that the structural change that occurs upon solvation moves the carboxyl carbon chemical shift upfield. Thus, even though changes in hydrogen bonding move the carboxyl shift upfield to some extent, in the case of AlaGly and GlyAla, the effect is primarily due to structural changes. The data indicate that, while the internal carbonyl carbon chemical shift is a sensitive reporter for changes in the hydrogen-bonding environment, structural changes are embedded primarily in the carboxyl carbon shift.

C_α Carbon. It is well-established that, for globular proteins in aqueous solutions, C_α chemical shifts are predominately sensitive to secondary structure. Our experimental observations of changes in chemical shifts upon solvation are consistent with this notion, now across multiple phases. Chemical shift trends for the C_α and C_β carbons upon solvation are shown graphically in Figures 5 and 6, respectively. The existence and importance of the putative $C_\alpha{-}H\cdots O{=}C$ hydrogen bond are controversial.[66–69] In this study, we have chosen to ignore them. When compared to the electrostatic forces in small zwitterions, these hydrogen bonds are very weak and unlikely to affect dipeptide conformations. Table 2 and Figure 5 show the observed ^{13}C chemical shifts for the C_α carbons in the crystalline and solvated dipeptides. In Figure 5, the upfield resonance is generally the N-terminal $C_{\alpha1}$ and the downfield resonance, the C-terminal $C_{\alpha2}$. This is likely due to the through-bond deshielding effect, which is larger at the C-terminus than at the N-terminus. The only exception is AlaGly in which the glycyl $C_{\alpha2}$ is upfield compared to $C_{\alpha1}$ which belongs to alanine. After residue identity, the C_α chemical shift is sensitive to the position relative to the C- and N-termini.

Figure 5 also shows that, while $C_{\alpha1}$ is always deshielded upon solvation, $C_{\alpha2}$ follows a less discernible trend. The magnitudes of the observed changes in $C_{\alpha1}$ and $C_{\alpha2}$ confirm the computational prediction that the $C_{\alpha1}$ chemical shift is more sensitive to structural changes than its C-terminal counterpart. Furthermore, the change in both the C_α chemical shifts in AlaGly is small. This is consistent with the computational prediction that AlaGly does not undergo a large structural change upon solvation.

Internal Amide. The amide nitrogen is an uncharged hydrogen bond donor. Upon solvation, the ^{15}N amide shift moves significantly downfield in three of the four dipeptides (Table 4, Figure 7). In the same three dipeptides, the amide forms a hydrogen bond with the terminal carboxyl group. In the fourth, GlyGly, it forms a hydrogen bond with the carbonyl oxygen. Correspondingly, the amide nitrogen shift moves upfield upon solvation in GlyGly, but moves downfield upon solvation in the other three dipeptides. The relative strengths of the different hydrogen bond acceptors, namely, $COO^- > H_2O > C{=}O$, helps explain this result. For GlyGly, solvation involves changing the

(66) Senes, A.; Ubarretxena-Belandia, I.; Engelman, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 9056–9061.
(67) Baures, P. W.; Beatty, A. M.; Dhanasekaran, M.; Helfrich, B. A. *J. Am. Chem. Soc.* **2002**, *124*, 11315–11323.
(68) Chamberlain, A. K.; Bowie, J. U. *J. Mol. Biol.* **2002**, *322*, 497–503.
(69) Yohannan, S.; Bowie, J. *J. Am. Chem. Soc.* **2004**, *126*, 2284–2285.

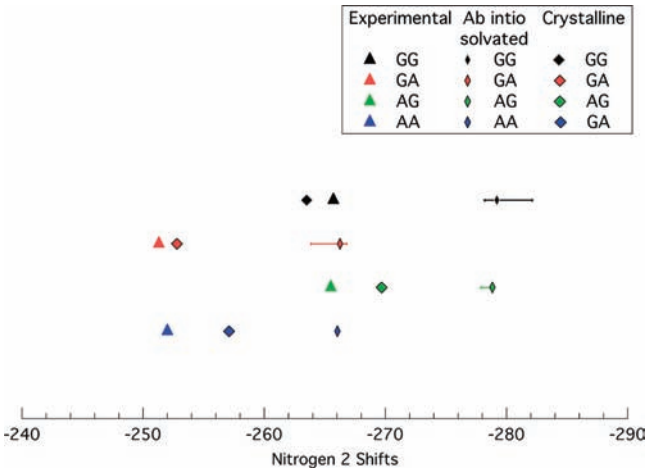

Figure 7. A graphical depiction of experimentally measured $^{15}N-H$ amide chemical shifts of the solvated dipeptides (triangles), crystalline peptides (squares), and the Boltzmann-weighted calculated shifts (diamonds). The horizontal bars on the computed shifts represent the range of shifts for that site across all the low-energy structures below 3RT.

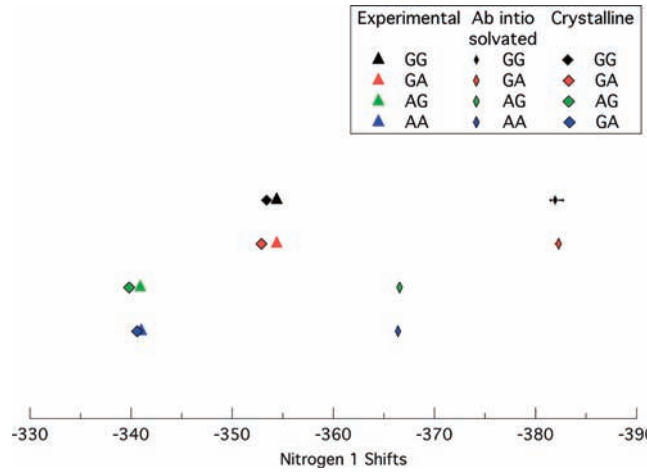

Figure 8. A graphical depiction of experimentally measured $^{15}NH_3^+$ amine chemical shifts of the solvated dipeptides (triangles), crystalline peptides (squares), and the Boltzmann-weighted calculated shifts (diamonds). The horizontal bars on the computed shifts represent the range of shifts for that site across all the low-energy structures below 3RT.

acceptor from the weak C═O to a stronger H_2O, thus, moving the chemical shift upfield. In the case of the other three dipeptides, solvation involves changing the acceptor from a strong COO^- to relatively weaker H_2O, moving the chemical shift downfield upon solvation. The data suggest that a strong hydrogen bond acceptor shields the amide nitrogen, while a weak hydrogen bond acceptor deshields the amide nitrogen. This inference is consistent with similar studies done on a set of large organic molecules by Li et al.[70] It is important to note that there are structural changes occurring upon solvation that affect the amide nitrogen chemical shift and further complicate the analysis. Our computational methods, however, have not been able to quantify the effect of these structural changes on the ^{15}N chemical shift.

Terminal Amine. In the crystalline dipeptides, the amine nitrogen is always hydrogen-bonded to the terminal carboxyl group in a head-to-tail arrangement.[71] We report ^{15}N chemical shifts for each of the four dipeptides in both the crystalline and the solvated states (Table 4). The data show that the amine ^{15}N chemical shift moves upfield upon solvation (Figure 8). The magnitude of the upfield shift is ~1 ppm for three of the four dipeptides. Furthermore, since the terminal amine nitrogen resides in a near-tetrahedral environment in the solid state, and is a free rotor in the solvated state, it is fair to assume that the changes in the amine chemical shift primarily reflect changes in its hydrogen-bonding environment. In all of the four dipeptides, solvation involves changing the hydrogen bond acceptor and replacing a carboxyl group with water. This is a well-documented effect.[72,73,70] Strong hydrogen bond acceptors, such as COO^-, tend to pull the proton away from the amine nitrogen. It is therefore conceivable that the ^+N-H bond length is marginally elongated when the hydrogen bond acceptor is charged (COO^-) versus when it is uncharged (H_2O). Solvation contracts this elongated N−H bond allowing the proton to shield the nitrogen more and move its chemical shift upfield.

(70) Li, Z. J.; Abramov, Y.; Bordner, J.; Leonard, J.; Medek, A.; Trask, A. V. *J. Am. Chem. Soc.* **2006**, *128*, 8199–8210.
(71) Gorbitz, C.; Etter, M. *Int. J. Pept. Protein Res.* **1992**, *39*, 93–110.
(72) Wei, Y.; McDermott, A. E.; Dios, A. C. d. *J. Am. Chem. Soc.* **1999**, *121*, 10389–10394.
(73) Lorente, P.; Limbach, H.-H. *Magn. Reson. Chem.* **2001**, *39*, S18–S29.

Computational Approaches to Dipeptide Solvation. Solvated Structures. The lowest energy structures from the *ab initio* PCM calculations are shown in Figure 2. In all four dipeptides, the "paddle" of the CO_2^- is aligned with the central amide dipole, based on simple electrostatic considerations. The same, however, does not hold true for the amino terminus where only GlyAla shows a fully extended geometry. Unlike the crystalline structure, the solvated state may execute hindered rotation about any of the sp^3 bonds, a feature often exhibited by small peptides. The observed NMR chemical shift is an average of various conformations that interconvert on a time scale much faster than the NMR data acquisition. However, such conformational dynamics are possible only if the solvated structure resides in a shallow energy minimum. In all four instances, we seek to understand the difference in structure between the crystalline ("C") and the lowest energy solvated ("L") conformers in terms of changes in structure and hydrogen bonding patterns between the crystalline and solvated states.

To separately measure structural effects from hydrogen bonding effects, we placed both the "C" and "L" conformations of each dipeptides in the same environment (vacuum) and calculated their chemical shifts. Conversely, we also placed the same structure (either "C" or "L") for each of the dipeptides, in different environments (vacuum and PCM) to explore how change in the solvation environment affects chemical shifts in the absence of structural changes. In the discussion below, we highlight some of the general trends. Detailed quantitative data for this comparison is provided in the Supporting Information Tables S.1−S.12.

GlyGly. The most significant difference between the crystal and solvated structures is the difference in the position of the terminal carboxyl group. In the crystalline conformation, the carboxy terminus puckers out of the plane of the molecule (ϕ_2 = 155°) forming a boat like structure, whereas in the lowest energy solution state conformation, this group relaxes back into the plane of the molecule (ϕ_2 = 180°). Crystal structure data show that this kink in the backbone is essential for strong intersheet hydrogen bonds in the crystal. In the solvated state, this constraint of intersheet hydrogen bonding is removed, and the carboxy terminus is able to relax back into the plane of the molecule maximizing the surface area and thus solvent exposure.

From the calculations, it is evident that chemical shift of the C_α carbons is more sensitive to structural changes (Table S.12 in Supporting Information), whereas the carbonyl carbon chemical shifts are more sensitive to changes in the hydrogen-bonding environment (Tables S.10 and S.11 in Supporting Information). The data also indicate that of the two C_α carbons, the ^{13}C chemical shift of $C_{\alpha1}$ (near the N-terminus) is more sensitive to structural rearrangement than that of $C_{\alpha2}$. Also, due to reduced steric hindrance, GlyGly has the shallowest solvated potential energy surface of the four dipeptides. Consequently, we converged on 8 solvated structures with energies below 3RT and far fewer for the other three dipeptides (Table 1).

GlyAla. In the crystalline state, each of the two terminal carboxyl oxygens forms a hydrogen bond with a different neighboring molecule, thus, linking the molecule to four other molecules in the lattice. At first glance, the "C" and "L" conformers of GlyAla appear quite different (Figure 2). However, careful analysis reveals that the two are separated by a 60° rotation about ϕ_2. NMR shifts for GlyAla, calculated for its crystalline conformation and for its lowest energy liquids conformations, in both vacuum and PCM also show that C_α carbons are sensitive to structure and the carbonyl carbons are sensitive to hydrogen bonding environment. Broadly speaking, the angle ψ_1 reflects changes around $C_{\alpha1}$ and ϕ_2 reflects changes around $C_{\alpha2}$. Our data indicate that a 6° difference in ψ_1 leads to a 13.7 ppm change in the chemical shift of $C_{\alpha1}$, while a 76° change in ϕ_2 leads to only a 4.7 ppm change in chemical shift for $C_{\alpha2}$. Yet again we see that the chemical shift of the $C_{\alpha1}$ is more sensitive to structural changes than $C_{\alpha2}$ (Table S.6 in Supporting Information).

AlaGly. Keeping with the general trend in dipeptide conformations thus far, the main difference between the "C" and "L" conformers of AlaGly is the position of the carboxyl group. In the crystal, it puckers out of the plane of the molecule (ϕ_2=−72.7°), and in the solvated state, it relaxes back (ϕ_2 = 178°). NMR shifts for AlaGly in its crystalline conformation and in its solvated conformation were calculated in both vacuum and PCM. Our calculations indicate that the $C_{\alpha1}$ and the terminal carboxyl (C_2') carbon chemical shifts report on structural changes that occur upon solvation, whereas the internal carbonyl (C_1') carbon shift is more sensitive to changes in hydrogen bonding environment (Tables S.7 and S.8 in Supporting Information). It is worth noting that experimentally we observe that the terminal carboxyl carbon C_2' is slightly more shielded in the "L" conformation than in the "C" conformation (Table 3). Our calculations verify that this upfield change in the C_2' chemical shift reflects a structural change rather than a change in the hydrogen-bonding environment.

AlaAla. Upon solvation, the ϕ_2 angle in AlaAla changes from −113° to −155°, reflecting a change from boat-like to a relaxed structure. The magnitude of this change, similar to that in GlyGly, is well above the predicted ±10° torsion angle dispersion range around the solvated AlaAla energy minimum suggesting that the changes are energetically significant.[37] The lack of a completely planar solvated conformation in AlaAla is due to steric hindrance from the two side-chain methyl groups. The position of the carboxyl oxygens in the crystalline state facilitates the intricate hydrogen bonding patterns observed in the lattice structure. Notably, the −113° value of ϕ_2 allows the head-to-tail arrangement of molecules to form a circular lattice structure with a 4-fold axis in the middle. This is in contrast to a ϕ_2 of −73° in AlaGly, which leads to a crystalline lattice made of linear columns. NMR chemical shift calculations, in both vacuum and PCM, again show that the C_α carbons are sensitive to structure and the carbonyl carbons are sensitive to hydrogen bonding environment (Tables S.1 and S.2 in Supporting Information, Figure 2). The internal carbonyl carbon shift is basically unchanged between the two conformations. The terminal carbonyl chemical shift reports on both structural and hydrogen bonding changes. Furthermore, unlike the other three dipeptides, the structural change between conformers "L" and "C" causes $C_{\alpha1}$ to become significantly more shielded. (Table S.3 in Supporting Information, Figure 2). AlaAla has been studied extensively and the literature is rich with information on its crystalline and solvated states, thus, providing a good comparison for the data presented here. The solvated structure determined from our *ab initio* calculations is very similar to that reported by Bour et al.[37] (Table 1). While MD simulations were beyond the scope of this project, it is satisfying to note that, at least in the case of AlaAla, the structures converged on by *ab initio* and MD simulations are consistent with one another.

We note four points from the computational effort. One, the solvated ("L") conformations for the dipeptides are generally more extended than the crystalline ("C") conformations. Across the four dipeptides, ψ_1 and ϕ_2 register most of the conformational change while the central ω angle remains unchanged upon solvation. Two, although both ψ_1 and ϕ_2 change upon solvation, ϕ_2 changes more than ψ_1. The former is directly related to the terminal carboxyl group indicating that upon solvation the terminal carboxyl group undergoes significant repositioning. Three, calculated NMR parameters show that the C_α carbon chemical shift is sensitive to structure, whereas the carbonyl and carboxyl chemical shifts generally report on the hydrogen bonding environment. Four, the NMR chemical shift of $C_{\alpha1}$ is more sensitive to structural changes than $C_{\alpha2}$. That is, changes in structure that alter the amino end of the dipeptide register as large changes in the $C_{\alpha1}$ chemical shift, but corresponding changes around the carboxyl end do not result in large changes in the $C_{\alpha2}$ chemical shift.

While neither a vacuum nor a PCM environment accurately describes the electronic environment in the crystalline or solvated states, the computational data are internally consistent, allowing us to identify viable trends within the computational data set. We surmise that an accurate description of the solvated environment may require an explicit solvent model. Given the scaling of the computation time with the number of electrons, *ab initio* explicit solvent calculations are computationally too costly at this time. Furthermore, explicit solvent modeling requires an algorithm that samples the conformation space—a challenging task, given the number of degrees of freedom. A rigorous computational effort of this sort has been recently undertaken on glycine,[18,20] although larger zwitterionic systems have yet to be explored. Molecular dynamics simulations provide an alternate means to incorporate explicit solvent effects; however, since the effects considered here are subtle, these approaches will fail to capture the essence due to the quality of the force fields available at the present time.

Even though we calculated ^{15}N chemical shifts in this study and report them here, the results are inconclusive, and hence, we choose not to discuss them. This unreliability of the computational ^{15}N shifts is a well-documented phenomenon

attributed to difficulties in predicting nitrogen shifts using DFT approaches.[74]

Conclusion

We have investigated structure and solvation in the four alanine- and glycine-containing dipeptides, by measuring the ^{13}C and ^{15}N NMR chemical shift and making detailed comparisons between crystal structures and the *ab initio* computed solvated structures. Several important trends emerge. One, the internal carbonyl carbon shift is a sensitive measure of the hydrogen-bonding environment and that site is always deshielded upon solvation. Two, the terminal carboxyl group undergoes significant structural repositioning in the solvated state; thus, although it is a site for strong hydrogen bonding, the changes in its chemical shift frequently reflect structural changes. Three, changes in the C_α chemical shift primarily reflect changes in the local conformation of the dipeptide. The N-terminal C_α chemical shift is more sensitive to structural changes than the C-terminal C_α. Four, trends in the amide ^{15}N shifts can be understood in terms of different the different types of hydrogen bond acceptors interacting with the N−H bond. Generally, in the case of uncharged nitrogen donors, a stronger acceptor gives rise to greater shielding at the amide nitrogen. And five, the terminal amine nitrogen is always shielded upon solvation. Unlike the amide donor, a charged nitrogen donor, like the amine, experiences a deshielding at the nitrogen.

We have shown that ^{13}C and ^{15}N NMR chemical shifts can serve as a sensitive probe for changes in structure and hydrogen bonding that occur upon solvation. With proper care and control of experimental parameters, the NMR chemical shifts in the solid and solvated state can become sensitive indicators of the electronic environment. However, a confluence of many factors leads to a dimensional collapse of information to one quantity, the chemical shift, that is difficult to interrogate. Thus, NMR data needs to be supplemented by crystallographic and computational data to understand the origins of the observed trends. Our study leaves some questions unanswered and raises some new ones. The accuracy and relevance of the *ab initio* PCM models, as applied to these dipeptide systems, are open to debate. The fact that the computed chemical shifts are not in complete agreement is likely due to matters of referencing as much as their computational accuracy, and yet *ab initio* calculations have allowed us to tease apart various contributing effects. We are reassured that the computational results reflect the experimentally observed changes in chemical shift upon solvation. We are confident that the high-level calculations reflect the low energy solvated structures, and while it may seem ironic that we are motivated by ideas of microsolvation, the bulk dielectric models are neither able to capture the site- nor event-specific hydrogen bonding events with the solvent. The experimental measurements at high magnetic field strength are unambiguous, though the ^{13}C and ^{15}N chemical shifts change upon solvation, and these changes must be due to changes in structure and differences in the solvent environment. The *ab initio* calculations serve as a useful guide, especially in instances where the peptide is able to interconvert between multiple conformations in solution.

The results presented from our study point to several new directions. To contrast the situation here, it would be interesting to investigate dipeptides with larger side chains, which, by their steric features, would likely adopt a single conformation in solution. Spectroscopic signatures of hydrogen bonds have captivated researchers for many years, and new results continue to shed light on this important area. Recent advances in ultra fast magic angle spinning NMR are making possible direct observation of ^{1}H in the solid state, which may, over time, provide a viable alternative to neutron diffraction. We are presently engaged in a study to detect weak $C_\alpha-H\cdots O$ hydrogen bonds in other peptide systems and to establish their NMR signature. For one model dipeptide, we have undertaken an extensive series of *ab initio* calculations to model explicit solvent molecules, and to investigate the accuracy of the PCM model. We are also attempting to measure CSA tensor orientations in polycrystalline samples using newly developed solid-state NMR methods. Work is underway in our laboratory to extend the study presented here to alanine- and glycine-containing tripeptides, as well as small, capped peptides.

Acknowledgment. We thank William Mohler (Oberlin College), Lon Buck, and Gary Drobny (University of Washington) for their assistance in the construction of an MAS controller; and Sydney Williams for her assistance in the preparation of some of the solvated samples. This work was supported by NSF-CAREER grant CHE-0449629, NSF-MRI grant CHE-0079470 toward the acquisition of a 600 MHz NMR spectrometer, NSF-MRI grant CHE-0420717 for the acquisition of a parallel computational cluster, and grant 42481-B4 from the donors of the Petroleum Research Fund of the American Chemical Society.

Supporting Information Available: Complete refs 17 and 46, details of magnetic susceptibility corrections, X-ray powder diffraction data, further *ab initio* results including energies and atomic coordinates, discussion of crystal structures, and chemical shift tensor analyses. This material is available free of charge via the Internet at http://pubs.acs.org.

JA902917S

(74) Fadda, E.; Casida, M. E.; Salahub, D. R. *J. Phys. Chem. A* **2003**, *107*, 9924–9930.