Conformational Analysis of Alanine Dipeptide from Dipolar Couplings in a Water-Based Liquid Crystal

Christoph F. Weise and James C. Weisshaar*

Department of Chemistry, University of Wisconsin-Madison, 1101 University Avenue, Madison, Wisconsin 53706-1396

Received: August 27, 2002; In Final Form: January 28, 2003

The proton NMR spectra of unlabeled alanine dipeptide (Ac-L-Ala-NHMe) at 300 MHz and of alanine dipeptide with a single ^{13}C label at 500 MHz are obtained in the lyotropic liquid-crystalline solvent cesium pentadecafluorooctanoate in water (CsPFO/H_2O). Simulations of the spectra yield 9 and 13 dipolar couplings D_{ij}, respectively, many with absolute sign determined. We fit the set of dipolar couplings by systematically varying the flexible dihedral angles ϕ and ψ while freezing local geometric details from the electronic structure calculations of Suhai and co-workers (Han, W. G.; Jalkanen, K. J.; Elstner, M.; Suhai, S. *J. Phys. Chem. B* **1998**, *102*, 2587). The orientation tensor is optimized at each combination of dihedral angles. Remarkably, a single conformer P_{II} ($\phi \approx -85°$, $\psi \approx +160°$) fits both sets of couplings within experimental error. The orientation tensor can be understood in terms of a simple rocking motion that dips the central methyl group into the fluorocarbon core of the CsPFO bicelle while alternately exposing both hydrogen-bonding pockets of P_{II} to interfacial or bulk water. The search for a minority conformer such as α_R (right-handed alpha helix, often favored by theory) using the larger data set was inconclusive. The data support localization of the peptide within the P_{II} well rather than the broad sampling of ϕ in the range from $-60°$ to $-180°$ (a "β/P_{II} minimum") found by certain models. We suggest that the P_{II} geometry is stable primarily because it maximizes the opportunity for peptide−water cooperative hydrogen bonding, whereas the α_R geometry is stable primarily because of its large dipole moment. Our result corroborates recent work on short polypeptides suggesting that they preferentially sample configurations that fluctuate about P_{II}-like structures, in contrast to the usual random-coil models.

I. Introduction

The use of empirical atomic-level force fields and semiempirical quantum mechanical methods in large-scale molecular dynamics (MD) simulations of protein structure and dynamics is a burgeoning enterprise.[1] Improved force fields would lead to better NMR structures of whole proteins, better drug design, and more reliable molecular-level insight into biochemical mechanisms. Toward this end, it remains important to test the accuracy of each method on smaller model systems by all available means. Model peptides such as Ac-L-Ala-NHMe (herein alanine dipeptide, Figure 1) provide the basic elements of secondary structure (one flexible backbone dihedral angle of each type, ϕ and ψ) in the simplest possible form. Accordingly, alanine dipeptide has become a standard test case for all models of peptide solvation.[2−4] A recent study[5] compared five different force fields with results from a semiempirical quantum mechanical method (QM/MM using the self-consistent tight-binding method SCCTB on the peptide). The closely related alanine tripeptide was studied with six different force fields.[6] The results are encouraging in the sense that the models agree broadly on which regions of (ϕ, ψ) space are sampled at 300 K; the dispersion in calculated relative free energies of local minima across models is only 1−2 kcal/mol. On the other hand, the energy accuracy is not yet sufficient to predict conformer populations accurately, and the models vary widely in terms of the geometries of local minima, the breadths of these minima, and the heights of barriers connecting them.

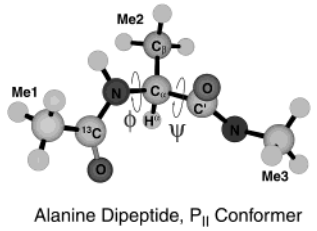

Figure 1. Alanine dipeptide in its P_{II} conformation showing ϕ, ψ, and the proton and ^{13}C labels.

On the experimental side, the conformational flexibility of small peptides presents a severe challenge to traditional solution-phase structural tools. Earlier experimental work on alanine dipeptide includes infrared absorption spectra,[7] Raman spectra,[8] and the circularly polarized variants of these,[8] as well as traditional NMR studies of scalar couplings and nuclear Overhauser effect (NOE) intensities.[9] The usual three-bond J couplings and NOE intensities from NMR spectra of normal solutions rule out certain conformers but do not unambiguously determine the dihedral angles ϕ and ψ, particularly when the measurement averages over two or more conformers.[10] For interpretation of vibrational spectra of peptides, detailed electronic structure calculations have shown the importance of including several explicit water molecules.[3] Two-dimensional nonlinear vibrational spectroscopies now under development promise more incisive structural information.[11]

In this work, we explore liquid-crystal NMR (LX-NMR) spectroscopy as a new experimental test of dipeptide conforma-

* Corresponding author. E-mail: weisshaar@chem.wisc.edu.

10.1021/jp026854i CCC: $25.00
Published on Web 03/18/2003

tion. We measure direct magnetic dipolar couplings for alanine dipeptide partially oriented in a water-based (lyotropic) liquid-crystalline medium CsPFO/water.[12] LX-NMR spectroscopy was invented in the 1960s,[13,14] when its focus was on rigid solute molecules and on the study of phase transitions in the liquid crystals themselves. Subsequently, flexible solutes were analyzed in conjunction with modeling of conformer ordering in nematic phases.[15,16] LX-NMR spectroscopy has enjoyed a recent revival in the structural refinement of whole proteins that adopt a single dominant conformer.[17,18] A clear strength of LX-NMR spectroscopy as a solution-phase structural tool is the direct, well-understood relationship between the measured dipolar coupling constants D_{ij} and the inverse cube of the internuclear distance r_{ij}, where i and j label two nuclei. Because motion about ϕ and ψ alters internuclear distances, dipolar couplings are quite sensitive to conformation. Our plan is to fit a set of D_{ij} as a superposition of contributions from various discrete conformers whose geometries are taken from electronic structure calculations of the dipeptide with explicit solvent water molecules.[3] Each conformer in the fit requires five adjustable parameters that describe the unknown angular distribution of the partially oriented molecular frame in space.

Remarkably, we showed earlier[19] that nine $^1H-^1H$ couplings in unlabeled alanine dipeptide could be fit within experimental error by a single conformer called polyproline II (P_{II}, $\phi \approx -85°$, $\psi \approx +160°$) and by no other single conformer. P_{II} is one of the two conformers most often preferred by theory; the other is the right-handed alpha helix (α_R, $\phi \approx -80°$, $\psi \approx -50°$). In this paper, we use a single ^{13}C label to provide four additional couplings, in principle allowing us to test two-conformer models against the data and potentially to estimate the relative free energies of P_{II} and α_R. However, P_{II} alone fits all 13 couplings within experimental error, which makes the search for statistically significant contributions from possible minority conformers inconclusive. The data support localization of the peptide within the P_{II} well rather than the broad sampling of ϕ in the range from $-60°$ to $-180°$ (a "β/P_{II} minimum") found by certain models. We argue against the possibility that the liquid-crystal medium perturbs alanine dipeptide conformational preferences or that P_{II} might orient much more strongly than α_R. Surprisingly, a simple physical picture based on a recent comparative study of methyl-substituted formamides[20] explains the observed orientation tensor of P_{II} at the CsPFO bicelle/water interface.

Given that P_{II} is the dominant conformer in the solution phase, alanine dipeptide provides an especially simple example of a large geometry change from gas phase to solvated peptide. Suhai et al.'s electronic structure study of alanine dipeptide with four explicit water molecules in an Onsager cavity embedded in a dielectric continuum provides intriguing clues to the nature of the solute−solvent interactions driving the geometry change.[3] Neither P_{II} nor α_R is an energy minimum in the gas phase. We suggest that the P_{II} geometry is stable primarily because it maximizes peptide−water cooperative hydrogen bonding, whereas the α_R geometry is stable primarily because of its large dipole moment. Our conclusion is in qualitative accord with early[21,22] and recently revived[23,24] suggestions based on circular dichroism and NMR data that *many* "disordered" short-chain polypeptides and perhaps *most* "random-coil" states of whole proteins in fact consist of P_{II}-like segments linked by sharp bends.

II. Experimental Techniques

At 42 wt % in water, cesium pentadecafluorooctanoate (CsPFO) forms a lyotropic liquid crystal comprising oblate, disk-shaped *bicelles* about 22 Å thick and 80 Å in diameter (aggregation number about 190) that occupy roughly 24% of the volume of the solution.[25] When heated to the isotropic phase (above 37 °C) and then cooled in the strong magnetic field of an NMR spectrometer, the bicelles form a *macroscopically oriented phase* with its director−−the symmetry axis of the mean field−−aligned parallel to the magnetic field. In LX-NMR spectroscopy, the liquid-crystalline solvent medium subjects the solute to an *orientationally anisotropic mean field.*[12] The orientational mean field in CsPFO/D_2O is stronger than that in phospholipid-based systems recently used in studies of larger biological molecules,[18,26] yet not so strong as to obviate first-order analysis of spectra. Small water-soluble peptides sample primarily an aqueous environment but acquire substantial orientation from transient encounters with the oriented CsPFO bicelles. The resulting NMR spectrum exhibits moderately strong intramolecular, direct magnetic dipolar couplings between nuclear spins.

CsPFO was prepared as follows: Equimolar amounts of cesium carbonate and pentadecafluorooctanoic acid were dissolved in 50:50 hexane/butanol through gentle heating to 40 °C and mild stirring. The product was recrystallized twice from 50:50 hexane/butanol and was dried in a vacuum oven at 60 °C to remove residual solvent, with purity verified by NMR spectroscopy. The peptide solute, surfactant, and water were weighed into an NMR tube to make approximately 1 g of sample with a solute concentration of ∼50 mM. The samples were heated to the isotropic phase (∼40 °C) and mixed in the NMR tube to obtain a homogeneous solution. The transition temperature from nematic to isotropic phase in the CsPFO occurred at the same temperature reported by Boden[12] to within 0.5 °C, as judged by the sharp onset of dipolar couplings as the temperature was slowly lowered. Proton spectra acquired at this higher temperature are liquidlike, with fairly narrow lines (1.5 Hz fwhm) and the usual J couplings.

Our quantitative studies and modeling efforts have focused primarily on dipolar-coupled NMR spectra of two samples. The first spectrum, briefly described earlier,[19] is of unlabeled Ac-L-Ala-NHMe (Bachem) dissolved at 0.3 wt % in 42 wt % CsPFO/D_2O, studied at 300 MHz and 25 °C with a Bruker AVANCE spectrometer in Chapel Hill, NC. Because the H^N protons readily exchange, the species under study is $CH_3CON(D)C(H)(CH_3)CON(D)CH_3$. The second spectrum uses one ^{13}C label to provide additional dipolar couplings. It is 0.3 wt % in 41.5 wt % CsPFO/D_2O, studied with a 500-MHz Varian INNOVA spectrometer at 26 °C in Madison, WI. The species under study is then $CH_3{}^{13}CON(D)C(H)(CH_3)CON(D)CH_3$. To prepare the ^{13}C-labeled compound, 118 mg of unlabeled alanine (Bachem) was treated with 340 μL of double-^{13}C-labeled acetic anhydride (Isotech) in 5 mL of methanol. Unlabeled reagents were purchased from Aldrich (Milwaukee, WI). After 24 h, the solvent was evaporated, and the residue washed with cold chloroform (87% yield). The product was treated with 265 mg of 1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride (EDCI), 250 mg of 1-hydroxybenzotriazole (HOBt), and 130 μL of 4-methylmorpholine in a solution of 15 mL of DMF and 10 mL of CH_2Cl_2. After the reaction had been allowed to proceed overnight, the solvent was evaporated under reduced pressure, and the product was separated by flash chromatography with powdered silica as the stationary phase and a mobile phase of 5% v/v methanol in chloroform. The solvent was evaporated from the fractions containing the dipeptide (identified by TLC), and the residue was separated with a second column using a mobile phase of 10% v/v methanol in chloroform. The purity

and identity of labeled alanine dipeptide were verified by ^{1}H NMR spectroscopy.

One-dimensional proton spectra were acquired with a pulse−observe sequence in the ^{1}H channel. Spectra at 300 MHz were acquired under conditions of ^{2}H decoupling. Conditions for the acquisition of proton spectra at 500 Hz were as follows: The frequency lock was turned off. A GARP decoupling sequence was used in the ^{2}H channel with the decoupler power selected as a compromise between optimal decoupling and minimal heating of the sample. The water resonance in the ^{1}H spectra of alanine dipeptide in CsPFO/acetate buffer was saturated with a presaturation pulse. The large number of couplings in the spectrum was detrimental to S/N, particularly in the methyne region. To acquire sufficient S/N, 64 transients were accumulated, at a minor cost to resolution due to slight fluctuations in the magnetic field during the acquisition. The number of points collected was 63488 over a sweep width of 8000 Hz, for a spectral resolution of 0.126 Hz. In all cases, the temperature was regulated to ±0.1 °C.

It is important to define *conformer* clearly to avoid confusion. In this paper, a conformer samples a distinct local free energy minimum in (ϕ, ψ) space. A conformer adopts a range of geometric configurations while executing fast torsional and vibrational motions within its free energy well. Throughout, we refer to the acetyl methyl protons as H^{Me1}, the C^β methyl protons as H^{Me2}, the *N*-methylamide methyl protons as H^{Me3}, and the C^α proton as H^α (Figure 1). In the 500-MHz spectrum of the labeled compound, ^{13}C refers to the single label at the carbonyl within the acetyl group.

III. Results

A. NMR Spectra. The 300-MHz proton NMR spectrum of alanine dipeptide in CsPFO/D_2O was shown previously (ref 19, Figure 2). The isotropic spectrum at 47 °C shows well-resolved features for H^α and for each of the three methyl groups, with a single significant scalar coupling $J(H^\alpha,H^{Me2}) = +7.3$ Hz. Additional *J* couplings extending over five or six bonds would measure 0.1 Hz or less and are neglected. The two H^N protons exchange with D_2O solvent and are not observed in the proton spectrum. The LX-NMR dipolar-coupled spectrum of the partially oriented alanine dipeptide at 25 °C was obtained with deuterium decoupling to remove dipolar coupling between H^α and the two D^N deuterons. The apparent fwhm of the narrowest lines increases from 3.6 Hz in the isotropic spectrum to 7 Hz in the LX-NMR spectrum. Despite the net orientation, the substantial motional narrowing of the spectrum indicates that the peptide is *not permanently bound to the bicelle surface* (which would produce a solidlike powder pattern), but rather samples many environments and orientations on the NMR time scale of microseconds to milliseconds.

The 500-MHz proton NMR spectrum of the dipeptide with the ^{13}C-labeled carbonyl group in the acetyl moiety is shown in Figure 2. The isotropic spectrum (top trace) now shows three resolved *J* couplings detailed below. The line width is 1.5 Hz fwhm. The dipolar-coupled spectrum at 26 °C (middle trace) shows narrowest features of 3 Hz fwhm. Deuterium decoupling somewhat sharpens the spectrum in the H^α region by removing the effects of small couplings to the two ND groups. The 500-MHz spectrum is more nearly in the first-order limit than the 300-MHz spectrum. Cursory comparison shows that additional dipolar couplings due to the presence of the ^{13}C label are resolved.

Assignment of the main features of the spectrum is straightforward. All chemical shifts move downfield (toward larger

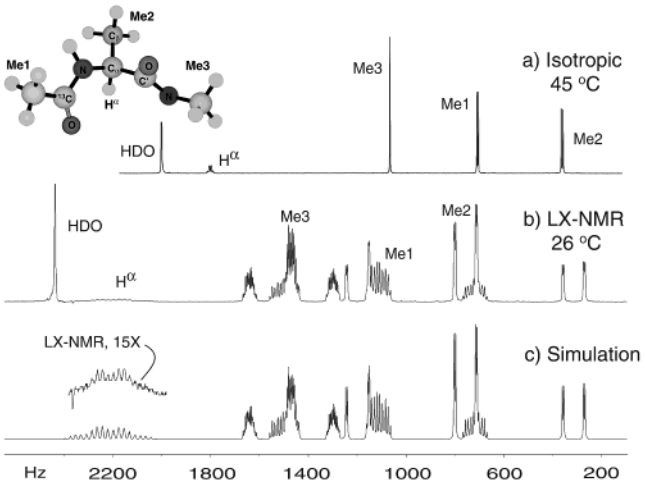

Figure 2. 500-MHz proton NMR spectra of alanine dipeptide with single ^{13}C label obtained in 41.5 wt % CsPFO in D_2O. Frequency scale with 0 Hz = 0 ppm. (a) Isotropic spectrum at 45 °C. (b) Oriented LX-NMR spectrum at 26 °C. (c) Example of best-fit simulated spectrum using XSIM and sign combination C1. See text.

frequencies in hertz) in the oriented sample as a result of the change in magnetic susceptibility of the bulk. In addition, the net orientation of the sample averages each chemical shift tensor differently in the LX-NMR spectrum than in the isotropic spectrum, accounting for the *differential* change in chemical shifts among different sets of equivalent protons. Dipolar couplings produce 1:2:1 intramethyl triplets that are typically widely split because of the small internuclear distance. The three sets of triplets are easily seen, and we assign them to lie in the same order as in the isotropic spectrum. The H^α resonance is split multiple times by both dipolar and scalar couplings. Without deuterium decoupling, H^α appears as a broad, featureless envelope. The irregularity of the pattern of some of the finer splittings is indicative of deviations from the first-order coupling regime (dipolar couplings comparable to differences in chemical shifts). A detailed analysis of the spectrum is presented below.

B. Dipolar Coupling Constants. The oriented spectra at 300 and 500 MHz form the basis of our conformational analysis of alanine dipeptide. Here, we describe extraction of dipolar coupling constants from the spectra. The Hamiltonian governing nuclear spin coupling in an anisotropic medium is well-known.[13,15] The scalar and dipolar coupling constants J_{ij} and D_{ij} appear beside the same spin operators[14,27]

$$\hat{H} = -\sum_i (1-\sigma_i)\nu_0 \hat{I}_{zi} + \sum_{i<j}(J_{ij}+2D_{ij})\hat{I}_{zi}\hat{I}_{zj} + \frac{1}{2}\sum_{i<j}(J_{ij}-D_{ij})(\hat{I}_i^+\hat{I}_j^- + \hat{I}_i^-\hat{I}_j^+) \quad (1)$$

The first term is the effective chemical shift in the anisotropic medium. The dipolar and scalar coupling constants appear asymmetrically in the second term (diagonal, first-order coupling) and the third term (off-diagonal, second-order coupling). In favorable circumstances, this allows unambiguous determination of the sign of D_{ij} when the sign of the corresponding J_{ij} is known by other means.

The oriented spectra were fit with the program XSIM (UNIX) or SpinWorks (MS Windows version) available from Dr. Kirk Marat (University of Manitoba).[28] The program takes inputs of estimates for the chemical shifts, *J* couplings, and dipolar couplings of the magnetically active nuclei and calculates the

NMR transition frequencies and intensities for the partially ordered spin system. Once manual adjustment of the spectral parameters has led to a sufficient resemblance between the experimental and calculated spectra, some 1000 calculated NMR transitions in the simulated spectrum are assigned by the user to corresponding resolved features in the experimental spectrum; this is a many-to-one mapping. The couplings and shifts can then be optimized through an automated least-squares routine to minimize the deviation between calculated and experimental frequencies.

1. 300-MHz Spectrum. The lower trace in Figure 2 of ref 19 is our best simulation of the 300-MHz experimental spectrum using the program XSIM with a uniform line-width parameter of 6 Hz fwhm. Input to the model Hamiltonian includes four adjustable chemical shifts δ, the single scalar coupling $J(H^{\alpha},H^{Me2})$ fixed at +7.3 Hz by the isotropic solution at 40 °C, and a set of nine adjustable proton−proton dipolar coupling constants D_{ij}. The value of $J(H^{\alpha},H^{Me2})$ differs negligibly from those derived from ^{1}H spectra of alanine dipeptide in normal D_2O at 20−40 °C. The sign is taken from the literature.[9]

The fitting procedure uses no intensity information. In first-order spectra, the intensity patterns would be independent of the relative signs of the D_{ij}, but the 300-MHz spectrum has sufficiently strong second-order coupling that the nuances of the intensity patterns in the region 750−1500 Hz are quite sensitive to the relative signs of the D_{ij}. Exhaustive exploration of *all* sign combinations shows that *exactly two* fit the spectral intensities qualitatively better than all others; they differ in the sign of two 9 Hz couplings. In the two best fits, both of these couplings are positive or both are negative. The best-fit D_{ij} values are collected in Table 1. The estimated accuracy varies from ±0.2 to 0.6 Hz. The error estimates are a semiquantitative measure of the range over which each parameter can vary (with all others reoptimized) before the quality of the fit definitely degrades either in frequency match or in intensity match. The small remaining systematic intensity deviations between experiment and the best fit probably occur because XSIM is fitting many lines of slightly different calculated frequency to the *center* frequency of each resolved feature and differential line-broadening mechanisms are neglected.

2. 500-MHz Spectrum. A similar fitting procedure was carried out on the more extensively coupled 500-MHz spectrum of ^{13}C-labeled alanine dipeptide (Figure 2b), with the best-fit spectrum shown in Figure 2c and the parameters given in Table 1. The magnitudes of three scalar couplings were obtained from ^{1}H spectra in the CsPFO solution at 40 °C (isotropic phase), while the signs were taken from the literature.[29] The scalar couplings differ negligibly from those derived from ^{1}H spectra of alanine dipeptide in normal D_2O at 20−40 °C. The new dipolar coupling $D(^{13}C,H^{Me2})$ gives rise to isolated doublets in the H^{Me2} region and is easily determined. Dipolar couplings between ^{13}C and the H^{Me1} and H^{Me3} groups were determined by matching the elaborate features in the corresponding methyl regions shown in greater detail in Figure 3. In the H^{α} region, the signal is attenuated by the water presaturation pulse, and the resolution is compromised by the sheer number of couplings to the centrally located H^{α}. The parameters $D(^{13}C,H^{\alpha})$ and $\sigma(H^{\alpha})$ were optimized manually and fixed during automated optimizations. The reported uncertainty in these parameters is accordingly large. In addition, because of their small size, the magnitudes of $D(H^{Me2},H^{Me3})$ and $D(H^{Me1},H^{Me2})$ were also determined manually and held constant during the automated optimizations.

Extraction of sign information from the 500-MHz spectrum is nontrivial. The 300-MHz spectrum of the unlabeled compound is particularly useful in determining sign information in that the reduced dispersion of the resonances leads to stronger second-order perturbations of the multiplet patterns. The media used to align the labeled and unlabeled alanine dipeptide species are sufficiently similar to permit transfer of sign information for the larger dipolar couplings between the spectra. For couplings of 5 Hz or larger, the ratio of corresponding coupling magnitudes D(500 MHz)/D(300 MHz) lies in the range 1.08−1.23 (Table 1). Consistently larger couplings at 500 MHz suggest moderately enhanced orientation of the bicelles themselves in the higher magnetic field. We transfer the *signs* of the seven largest ^{1}H−^{1}H couplings from the 300-MHz spectrum of unlabeled alanine dipeptide to the 500-MHz spectrum of the ^{13}C labeled alanine dipeptide intact. For the four ^{13}C−^{1}H couplings, all possible sign combinations were explored. Figure 3 shows an example of the subtle discrepancies between experiment and simulation that can constrain some sign choices. To fit the spectrum, the couplings $D(H^{\alpha},H^{Me1})$ and $D(H^{\alpha},H^{Me3})$ must both be positive or both be negative. $D(H^{Me1},^{13}C)$ and $D(H^{Me3},^{13}C)$ must both be positive or both be negative. The signs of $D(H^{\alpha},^{13}C)$ and of $D(H^{Me2},^{13}C)$ are completely undetermined. The 500-MHz spectrum is thus fit very well by 16 combinations of coupling signs among those couplings with

TABLE 1: Best-Fit Parameters from Dipolar-Coupled NMR Spectra of Alanine Dipeptide at 300 and 500 MHz

parameter[a]	unlabeled alanine dipeptide, 300 MHz	^{13}C-labeled alanine dipeptide, 500 MHz	ratio[b] $\frac{D(500\text{ MHz})}{D(300\text{ MHz})}$	r_{ij} (Å)[c]	S_{ij}[d]
$\delta(H^{Me1})$	719.3 ± 0.6	1078			
$\delta(H^{Me2})$	506.8 ± 0.6	721.7			
$\delta(H^{Me3})$	931.2 ± 0.5	1433.7			
$\delta(H^{\alpha})$	1374.8 ± 0.8	2169			
$J(H^{\alpha},H^{Me2})$	+7.3	+7.3			
$J(H^{Me1},^{13}C)$	−	−6.1			
$J(H^{\alpha},^{13}C)$	−	+2.4			
$D(H^{Me1})$	−111.2 ± 0.3	−131.7 ± 0.3	1.18	1.78	−0.0124
$D(H^{Me1},^{13}C)$	−	+25.3 ± 0.3, −19.2 ± 0.3[e]	−	2.15	−0.0083
$D(H^{\alpha},H^{Me1})$	±9.5 ± 0.4[f]	±11.7 ± 0.3[f]	1.23	4.47	−0.0087
$D(H^{Me1},H^{Me2})$	$\lvert D\rvert$ < 0.6	$\lvert D\rvert$ < 0.7	−	5.39	<0.00039
$D(H^{Me1},H^{Me3})$	+4.6 ± 0.2	+5.0 ± 0.3	1.09	7.36	−0.0166
$D(H^{Me2})$	+127.8 ± 0.3	+147.6 ± 0.4	1.15	1.77	+0.0137
$D(H^{Me2},^{13}C)$	−	∓2.9 ± 0.3[g]	−	4.06	+0.0064
$D(H^{\alpha},H^{Me2})$	−41.1 ± 0.4	−48.1 ± 0.3	1.17	2.59	+0.0070
$D(H^{Me2},H^{Me3})$	$\lvert D\rvert$ < 0.6	$\lvert D\rvert$ < 0.3	−	5.34	<0.0002
$D(H^{Me3})$	−50.2 ± 0.3	−57.0 ± 0.3	1.14	1.77	−0.0053
$D(H^{Me3},^{13}C)$	−	±2.6 ± 0.3[e]	−	5.73	−0.0162
$D(H^{\alpha},H^{Me3})$	±9.3 ± 0.3[f]	±10.0 ± 0.3[f]	1.08	4.45	−0.0073
$D(H^{\alpha},^{13}C)$	−	+9.5 ± 1.5, −11.9 ± 1.5[g]	−	2.51	−0.0050

[a] Proton labels as in Figure 1: H^{Me1}, acetyl methyl; H^{Me2}, C^{α} methyl; H^{Me3}, *N*-methylamide methyl; H^{α}, C^{α} methine. In the 500-MHz spectrum, the acetyl carbonyl carbon is the ^{13}C. Hamiltonian parameters (eq 1) are chemical shifts δ, scalar coupling constants J, and dipolar coupling constants D. Two different sign combinations fit the 300-MHz spectrum equally well. Sixteen different sign combinations fit the 500-MHz spectrum equally well, as detailed in Supplementary Table S1. [b] Ratio of proton−proton dipolar coupling at 500 MHz to that at 300 MHz. [c] Average internuclear distance according to best-fit P_{II} geometry: $\phi = -90°$, $\psi = +140°$, $C_7^{ax\prime}$ local geometry from ref 3. [d] $S_{ij} = \langle P_2(\cos\theta_{ij})\rangle$ as in eq 3 for the best one-conformer fit of the 500-MHz data (P_{II} geometry: $\phi = -90°$, $\psi = +140°$, $C_7^{ax\prime}$ local geometry) to the C1 combination of dipolar coupling signs (all of the *upper* signs). For intramethyl couplings, S_{ij} refers to the orientation of the methyl symmetry axis. [e] This pair of couplings must be taken as either +/+ or −/− (both upper or both lower signs). [f] This pair of couplings must be taken as either +/+ or −/− (both upper or both lower signs). [g] The sign of these dipolar couplings is completely flexible; either choice fits the spectrum well with *any* sign combination for other couplings.

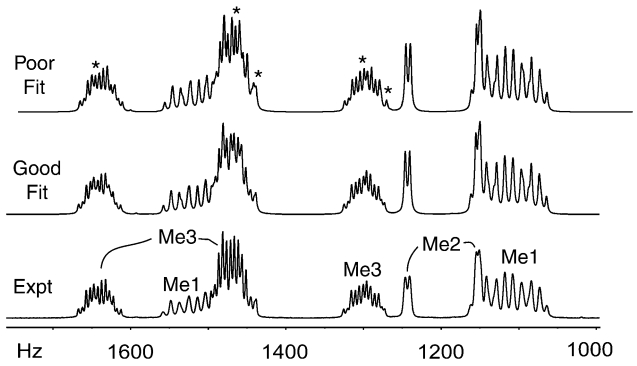

Figure 3. (lower trace) Expanded view of methyl region of the LX-NMR spectrum of Figure 2. (middle trace) Best-fit simulation using the C1 sign combination (upper signs in Table 1). The nuances of the intensity patterns are well reproduced. Some line frequencies in the Me3 region are slightly in error. XSIM is constrained to use a single line width; it is slightly too narrow at the right and too broad at the left. (upper trace) Best fit to the sign combination that is the same as C15 except $D(H^{Me3},{}^{13}C)$ is changed in sign from −2.6 to +2.6 Hz. The intensity pattern is significantly less well reproduced in the Me3 triplet as marked by asterisks.

magnitudes larger than 1 Hz. The sign ambiguities in $D(H^{Me1},{}^{13}C)$ and $D(H^{\alpha},{}^{13}C)$ lead to magnitude ambiguities as well because there are two ways to obtain the same magnitude $|J_{ij} + 2D_{ij}|$ in the first-order term of eq 1. A summary of the 16 sign combinations that fit the spectra is contained in Table 1; they are written out and labeled as C1−C16 in Supplementary Table S1.

C. Fitting Coupling Constants to Geometries Based on Theory. The dipolar coupling between nuclei *i* and *j* depends on the internuclear distance r_{ij} and on the angular distribution of the vector $\mathbf{r}_{ij}$ in space[15]

$$D_{ij} = -K_{ij}\langle r_{ij}^{-3} P_2(\cos\theta_{ij})\rangle \tag{2a}$$

$$\approx -K_{ij}\sum_{n=1}^{N} f_n \langle r_{ij}^{-3} P_2(\cos\theta_{ij})\rangle_n \tag{2b}$$

$$\approx \frac{2}{3}\sum_{n=1}^{N} f_n \mathrm{Tr}(\boldsymbol{S}_n \mathbf{G}_{ij}^{\,n}) \tag{2c}$$

Here, K_{ij} contains fundamental constants; θ_{ij} is the angle between $\mathbf{r}_{ij}$ and the *z* axis (magnetic field axis); and P_2 is the second Legendre polynomial, $P_2(x) = (3x^2 - 1)/2$. The brackets indicate an average over all motion of the molecule, including internal rotation of methyl groups, molecular vibration, translation relative to the bicelle/water interface, and overall molecular reorientation in space. Equation 2a is quite general. Equation 2b writes the dipolar coupling as a discrete sum over *N* distinct conformers, each representing a free energy minimum in (ϕ, ψ) space of the dipeptide with f_n being the fractional population of conformer *n*.[16] This *rotational isomeric states approximation*[30] assumes that barriers to isomerization are large enough compared with k_BT that the dipeptide spends much more time in local free energy minima than in the act of isomerization. Equation 2c further assumes that, *within* each free energy well, the internal vibrational−torsional motion separates from tumbling in space, i.e., vibrational averaging of the internal geometry is fast on the time scale of reorientation.[16,27] The average implied by the brackets in eq 2b then can be factored. $\mathbf{S}_n$ is the symmetric, traceless 3 × 3 orientation tensor describing the angular distribution in space of a set of molecule-fixed Cartesian axes for conformer *n*. $\mathbf{G}_{ij}^{\,n}$ is the 3 × 3 "geometry matrix" for the *ij* coupling within conformer *n*, consisting of r_{ij}^{-3} times products of projection cosines of the *ij* internuclear vector onto the same molecule-fixed axes. The peptide of interest here is nonplanar so that each $\mathbf{S}_n$ has five independent elements.

Finally, in the simple case of a single dominant conformer whose reorientation separates from vibration, we can write

$$D_{ij} \approx -K_{ij} S_{ij} \langle r_{ij}^{-3}\rangle \tag{3}$$

where the orientational order parameter is $S_{ij} = \langle P_2(\cos\theta_{ij})\rangle$. For the intramethyl $^1H-{}^1H$ couplings, we quote the order parameter for the methyl rotor symmetry axis, which is obtained from the motionally averaged order parameter for the $^1H-{}^1H$ pairs on multiplication by −2.[31]

The number of experimental dipolar couplings is nine from the 300-MHz spectrum and 13 from the 500-MHz spectrum, far too few to determine detailed geometries a priori. To a first approximation, alanine dipeptide consists of two rigid amide planes moving on dihedral *hinges* (ϕ, ψ) held at tetrahedral angles to each other (Figure 1). Clearly, dipolar couplings between pairs of nuclei whose separation changes with ϕ or ψ will be particularly sensitive to the conformation. Our general strategy is to take $N = 1$ or 2 candidate geometries from theory and fit the experimental coupling constants to eq 2c in a least-squares sense, with the 5*N* independent elements of the products $f_n\mathbf{S}_n$ as fitting parameters. The molecule-fixed axis system used to calculate the $\mathbf{G}_{ij}^{\,n}$ matrices is attached to the central $-C^{\alpha}(H^{\alpha})(Me2)-$ moiety. We explore the use of different local geometric details (all coordinates other than ϕ and ψ) from among the energy minima found in the recent density functional electronic structure calculations of Suhai and co-workers.[3] They carried out complete geometry optimizations for alanine dipeptide plus four explicit water molecules in a spherical cavity embedded in a dielectric medium (Onsager cavity). Five local energy minima (marked by red X's in Figure 4) were found and named according to common structural motifs in proteins: P_{II} (polyproline II or polyglycine II), β_2' (beta sheet), α_R' (right-handed helix), α_L' (left-handed helix), and $C_7^{ax\prime}$ (seven-member "ring"). These geometries differ substantially in ϕ and ψ but vary only mildly in the other local geometric details. The geometries optimized with and without the Onsager cavity are very similar and behave identically in the fitting attempts.

1. 300-MHz Spectrum. With nine dipolar couplings from the 300-MHz spectrum, we could attempt only one-conformer fits.[19] For each choice of frozen local geometry, we varied ϕ and ψ on a 10° grid of 1296 conformer geometries and solved the least-squares problem at each (ϕ, ψ), i.e., the entire **S** matrix was optimized at each (ϕ, ψ). Methyl rotors were assumed to sample all internal rotation dihedral angles uniformly without intermethyl correlations. The resulting rms error contour plot using frozen local geometry elements from Suhai et al.'s $C_7^{ax\prime}$ minimum[3] (shown in Supplementary Figure 1 of ref 19) shows that a *single conformer* with $\phi = -90°$ and $\psi = +140°$ fits all nine data extremely well when the +9.5, +9.3 Hz sign combination (Table 1) is chosen for $D(H^{\alpha},H^{Me1})$ and $D(H^{\alpha},H^{Me3})$. The minimum rms error is only 0.14 Hz, compared with estimated rms experimental uncertainty of 0.3 Hz. No single error in a dipolar coupling exceeds the error estimate from the spectral fits. The range of angles over which the fit remains very good (rms error < 0.4 Hz) is quite narrow ($\Delta\phi = \pm 8°$ and $\Delta\psi = \pm 16°$). Two other local minima in the rms error plot have rms errors below 1 Hz, but both of these include specific errors that far exceed the experimental uncertainty. Significantly, single conformers near the α_R geometry fit the

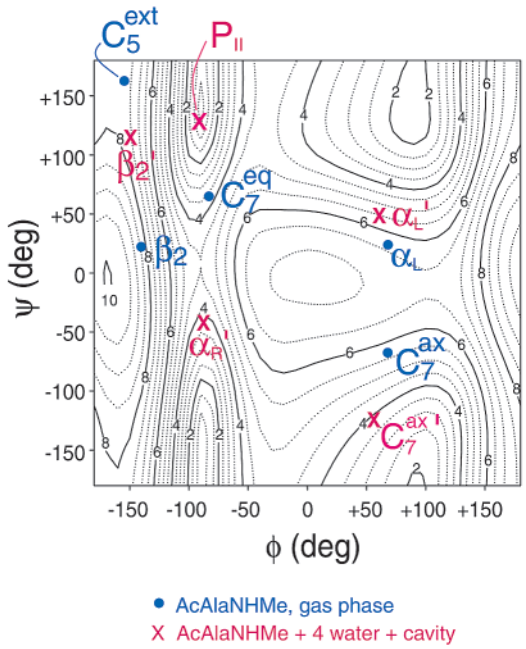

Figure 4. Contour plot of root-mean-squared best-fit error in hertz (*not* an energy contour plot) for single-conformer fits to 13 dipolar couplings from the 500-MHz spectrum using the sign combination C1 and local geometric elements from Suhai et al.'s $C_7^{ax\prime}$ minimum (ref 3). Superimposed are markers of Suhai et al.'s optimized geometries for gas-phase alanine dipeptide (blue dots) and for the four-water cluster in an Onsager cavity (red X's). P_{II} and α_R are not gas-phase energy minima.

data very poorly, with rms error in the 3−5 Hz range. When the alternative −9.4, −9.2 Hz sign combination is chosen, the rms error of the P_{II} conformer grows to 4.5 Hz, and no (ϕ, ψ) combination fits the data at all well.

These results depend somewhat on which of Suhai et al.'s computed minima provides the local geometric details. The minima vary in the local structure of the hydrogen bonds to water, which, in turn, affects details of the geometry about both amide bonds. For the +9.5, +9.3 Hz sign choice, the rms error at the P_{II} minimum among five choices of local geometry varied from 0.14 to 0.7 Hz, with four local geometries giving quantitatively good fits with rms errors below 0.45 Hz. In all cases, the best fit is a minimum in the region of P_{II} having ϕ of −80° or −90° and ψ in the range of +160 ± 20°. In all cases, a single conformer near the geometry of α_R fits the data very poorly (rms error larger than 4 Hz). Thus, we tentatively concluded from the 300-MHz results that P_{II} is the predominant conformer in our CsPFO/water solutions. Effects of vibrational averaging are discussed in section D below.

2. 500-MHz Spectrum. The four additional dipolar couplings from the ^{13}C-labeled 500-MHz spectrum in principle permit a search for a small contribution from one additional conformer, because each conformer in the fit uses up five adjustable parameters. The most likely candidate according to many theoretical treatments is α_R. First, we tested whether P_{II} alone can fit all 13 couplings. Of the 16 sign combinations consistent with the spectrum (Table 1), 14 permit no single conformer fit with rms error below 3 Hz. The other two combinations (C1 and C5 in Supplementary Table S1) differ only in the sign of $D(H^{Me2},{}^{13}C)$. The choice +2.9 Hz yields best-fit rms errors of 1.6−2.0 Hz, depending on local geometry. The choice −2.9 Hz (combination C1, which corresponds to taking all of the upper signs in Table 1) gives rms errors in the range 0.2−0.7 Hz, very similar to the 300-MHz results. The estimated rms uncertainty in the dipolar couplings themselves is 0.5 Hz. The lowest error is 0.22 Hz, again occurring at $\phi = -90°$, $\psi = +140°$ with local geometry taken from Suhai et al.'s $C_7^{ax\prime}$ minimum. The error plots for sign combination C1 using various local geometries all show a narrow trough of reasonable fits oriented vertically on the ϕ, ψ plot (example in Figure 4). In all cases, the best-fit value of ϕ is either −80° or −90° and the best-fit value of ψ lies in the broader range from +140° to +180°.

Guided by theory, we looked for *two-conformer* fits to the 16 sign combinations consistent with the spectra. In these searches, we used local geometric elements from Suhai et al.'s $C_7^{ax\prime}$ minimum throughout. In each case, one of the conformers was frozen to be either P_{II} ($\phi = -90°$, $\psi = +140°$) or α_R ($\phi = -60°$, $\psi = -60°$). The complete range of (ϕ, ψ) possibilities for the second conformer was sampled on a 10° grid as before. For the C1 sign combination that was fit to 0.22 Hz rms by P_{II} alone, the "P_{II} + other" fits can only do better. Indeed, a wide range of second conformers yield fits to 0.13 Hz rms or better, but the statistical significance of the improvement is dubious. In all regions distinct from the narrow trough of good one-conformer fits to P_{II}, the matrix $f_1\mathbf{S}_1$ for the P_{II} part was 3−30 times larger than $f_2\mathbf{S}_2$ for the second conformer. We judge the "size" of each $f_i\mathbf{S}_i$ by the root-mean-square of the matrix elements. Perhaps significantly, admixture of α_R to P_{II} does not improve the fit.

There remains the possibility that two conformers combined with two independent **S** matrices can fit some of the 15 other sign combinations as well as P_{II} alone fits C1. This indeed occurs. For example, in the P_{II} + other searches, the sign combinations C3, C5, and C7 can all be fit by two conformers to within 0.1 Hz rms. Moreover, two of these sign combinations, C5 and C7, are well fit by comparable amplitudes of P_{II} and α_R. For the other 12 combinations, the rms error is 4 Hz or larger. We also explored "α_R + other" fits to all 16 possible sign combinations. For C1, C3, C5, and C7 we found fits with rms error of 0.5 Hz or less. For all other sign combinations, there were no good fits (lowest rms error of 6 Hz or more). Among the regions of good fits, α_R and the other conformer had comparable amplitudes as estimated by the rms value of the elements of $f_i\mathbf{S}_i$. The other conformer did not match any region of Ramachandran space preferred by theory.

Thus, *many* admixtures of two conformers with similar amplitudes can fit several of the sign combinations. The P_{II} conformer is special in its ability to fit one sign combination by itself, but there seems to be nothing special about α_R as the choice of the fixed conformer in two-conformer searches. Evidently, the four new ^{13}C couplings do not provide sufficient independent information to narrow the search usefully for two-conformer fits. The likely reason is that several pairs of the 13 internuclear vectors lie approximately parallel to each other, sometimes depending on ϕ and ψ. This reduces the effective number of linearly independent equations. As this number approaches 11 or even 10, which it seems to do for a variety of choices of ϕ and ψ for the second conformer, eq 2c approaches 10 equations in 10 unknowns. Very good two-conformer fits thus arise frequently. The 13 data without better independent sign information can neither identify a small contribution from a minority conformer nor distinguish between the one-conformer fit to P_{II} and a variety of two-conformer fits.

It remains highly significant that, for one and only one of the 16 sign combinations of dipolar couplings consistent with the spectrum, all 13 data are fit very well by the single conformer P_{II} using only five adjustable **S**-matrix elements. This

certainly need not occur mathematically. It could be fortuitous, but this seems unlikely. If two conformers of roughly comparable amplitude created the couplings but a single **S** matrix and geometry provided a good fit, that geometry need not match one of the two conformers preferred by theory. P_{II} alone fits the data and is consistently one of two prime theoretical candidates for the most stable conformer. Furthermore, the corresponding **S** matrix has an appealing physical interpretation (section IV.B). Finally, P_{II} is consistent with the values of ϕ indicated by certain scalar coupling $J(H^N,H^\alpha)$. Although not definitive, the dipolar coupling data provide powerful evidence in support of P_{II} as the dominant conformer of alanine dipeptide in the CsPFO/water medium.

D. Vibrational Effects. In using eq 2c, we have assumed that vibrational averaging within each conformer can be mimicked well by a single effective geometry. That is, we have replaced $\langle r_{ij}^{-3}\rangle$ by $(r_{ij}^0)^{-3}$, the inverse cube of the internuclear distance in the best-fit geometry derived from one of Suhai et al.'s local geometries by altering only ϕ and ψ. Alanine dipeptide is quite flexible, and MD simulations indicate that the range of thermal excursions about ϕ and ψ is about $\pm 20°$ rms.[5,6] Methyl rotor motion and high-frequency vibrations must be considered as well. The effects of high-frequency vibrations within a conformer well on dipolar couplings have been treated as a series expansion of correction terms to the measured couplings, including both harmonic and anharmonic effects.[31,32] The theory requires vibrational frequencies, anharmonicity constants, and mean-squared displacements along each normal coordinate of the molecule *in solution*. In a recent detailed study of formamide,[20,33] we attempted such corrections using parameters from electronic structure calculations on the *gas-phase* molecule. The results are extremely sensitive to the assumed magnitude of displacements along two low-frequency, out-of-plane modes. The gas-phase calculations evidently grossly overestimate their mean-squared displacements. Thus, we did not attempt to apply the standard theory to alanine dipeptide using gas-phase normal modes for ϕ, ψ torsion.

Why might a single geometry mimic the average over configurations sampled in solution by alanine dipeptide? Part of the answer lies in the relatively large distances between the coupled nuclei under study. For high-frequency vibrations, harmonic corrections fall off as r_{ij}^{-2} and as the inverse of the vibrational frequency.[31] Somewhat fortuitously, large-amplitude torsional motion about ϕ and ψ modulates only the larger proton−proton distances (Table 1). The much smaller intramethyl distances and $r(H^\alpha-H^{Me2})$ are modulated only by higher-frequency vibrations on the order of 1200 cm^{-1} and above.

To estimate the magnitude of vibrational effects, we use a hybrid approach. For a single conformer, we can write the experimental dipolar coupling as

$$D^{ij}_{expt} = {}^2/_3\mathrm{Tr}(\mathbf{S}\mathbf{G}^{ij}) = {}^2/_3\mathrm{Tr}[\mathbf{S}(\mathbf{G}^{ij}_0 + \Delta\mathbf{G}^{ij}_{int})] \qquad (4)$$

where $\mathbf{G}^{ij}_0$ is built from the equilibrium geometry and $\Delta\mathbf{G}^{ij}_{int}$ corrects for all internal motion. We partition the correction $\Delta\mathbf{G}^{ij}_{int}$ into separable contributions from large-amplitude ϕ and ψ motions and higher-frequency vibrations to obtain

$$\mathbf{G}^{ij} = (\mathbf{G}^{ij}_0)_{Me} + (\Delta\mathbf{G}^{ij}_{\phi\psi})_{Me} + \Delta\mathbf{G}^{ij}_{vib} \qquad (5)$$

Here, the subscript Me indicates an average over internal rotation of all three methyl rotors, described below. The standard harmonic correction for high-frequency vibrations is carried out with the methyl rotors fixed at their equilibrium positions. In principle, one should average the $\mathbf{G}^{ij}$ matrices over all three types of motion as indicated and reoptimize the **S** matrix for best fit of the experimental data. Instead, we investigated each effect separately. We used the local geometric elements from Suhai et al.'s $C_7^{ax\prime}$ minimum in each case.

To mimic effects of large-amplitude motion about a given (ϕ_0, ψ_0), we calculated $\langle\mathbf{G}^{ij}(\phi_0,\psi_0)\rangle_{\Delta\phi,\Delta\psi}$ as an *unweighted average* over ϕ and ψ in the square region $\Delta\phi$, $\Delta\psi = \pm 20°$ about the central values (ϕ_0, ψ_0). The averaged **G** matrices were inserted into eq 2c, and the **S** matrix was reoptimized on a 10° grid over the entire Ramachandran space to generate a best-fit rms error plot with ϕ, ψ motion included. The result is quantitatively remarkably similar to the single-geometry plot of Figure 4. The error is minimized at the same P_{II} geometry with an rms error of 0.26 Hz for the ϕ, ψ average compared with 0.22 Hz for the single geometries. In other words, averaging the **G** matrices over $\pm 20°$ in both ϕ and ψ affects the quality of the single-conformer fit less than varying the choice of local geometric elements among Suhai et al.'s minima (section III.C above). If the range of motion is taken to be $\pm 10°$ or $\pm 30°$ and the **S** matrix is reoptimized, the rms fitting error at the P_{II} minimum becomes 0.23 or 0.34 Hz, respectively. The adjustments to the **S** matrix are minor when the reference molecule-fixed coordinate frame is placed on the central alanine moiety, as we have done thus far. In that choice of frame, torsion about ϕ and ψ changes the orientation of the methyl rotor axes only moderately. If instead the **S** matrix is referenced to axes fixed in one or the other of the amide planes, then torsion about ϕ or ψ causes large-scale reorientation of one or more methyl-group axes. Not surprisingly, the reoptimized **S** matrix is then quite different from that of the single-geometry fit. However, the data are still fit to 0.25 Hz rms for motion on the $\pm 20°$ grid. Evidently, a single geometry *centrally located* in the P_{II} conformer bowl can fit dipolar coupling data generated by a broad ϕ, ψ average about the P_{II} minimum very well.

Next, we tried two different models of methyl-group torsion. In electronic structure calculations on isolated alanine dipeptide, the 3-fold torsional potentials for Me1 and Me3 have barrier heights of only 0.3 kcal/mol, less than $k_BT = 0.6$ kcal/mol at 35 °C. The tetrahedral geometry about C^α creates a 3.5 kcal/mol 3-fold barrier for Me2, with the preferred orientation placing the rotor staggered relative to the other three bonds to C^α. We tested two models of methyl motion: isotropic, uncorrelated sampling for all three rotors (which was used in all of the single- and double-conformer fits described above) and Boltzmann-weighted, uncorrelated sampling using the gas-phase potential for all three rotors. For each motional model, we average the **G** matrix for each *ij* pair over the assumed methyl motions in 4° angular steps and reoptimized the **S** matrix as in the treatment of ϕ, ψ motion. The isotropic model gives an rms error of 0.22 Hz, as already seen. The Boltzmann average gives an rms error of 0.4 Hz. The **S** matrix changes only subtly. The effect of two different treatments of methyl motion is comparable to the effect of different choices of local geometry.

Finally, we investigated the effects of ordinary high-frequency vibrational motions on the dipolar couplings. The largest effects will arise for the shortest internuclear distances. For proton−proton couplings, these are the intramethyl distances and $r(H^a-H^{Me2})$. To simplify the problem, we carried out gas-phase ab initio vibrational mode calculations [B3LYP/6-31+G(d,p)] on the model fragments alanine, *cis*-*N*-methylformamide, and acetamide. We transferred Cartesian displacements and frequencies from these small molecules onto the dipeptide and computed harmonic corrections to the measured dipolar couplings using

perturbation theory.[34] Fits were performed over the same 10° grid of geometries. The lowest rms error (0.56 Hz) fell once again in the P_{II} region, with the optimal values of ϕ and ψ shifted slightly to −80° and +160°, respectively. The vibrational correction to each coupling can then be estimated as $^2/_3\mathrm{Tr}(\mathbf{S}\Delta\mathbf{G}^{ij}_{vib})$, where **S** is the reoptimized tensor. For the best-fit P_{II} conformer, the largest correction is 3.4 Hz (6%) for $D(H^{Me3})$. All other corrections are 1−2% or less. The **S** matrix elements from the corrected fit at the best P_{II} geometry differs by less than 10% from the matrix elements obtained without corrections.

Our illustrative calculations suggest why fitting the data to a single, frozen geometry can succeed. Considered individually, each of the three contributions from internal motion ($\phi-\psi$ torsion, methyl internal rotation, and high-frequency vibrations) is readily compensated using a single geometry with moderate adjustment of the **S** matrix. For the range of internuclear distances studied, small changes in an **S** matrix anchored to the central C^α region can compensate the effects of the average over thermal internal motion.

E. Orientation of P_{II}. In Figure 5a, we show the single-conformer best-fit geometry (P_{II}: $\phi = -90°$, $\psi = +140°$, $C_7^{ax\prime}$ local geometry) in the coordinate system that diagonalizes **S**. The eigenvalues $S_{11} = -0.0182$, $S_{22} = +0.0167$, and $S_{33} = +0.0015$ yield an asymmetry parameter $\eta = 0.84$.[31] The accuracy of the fitting parameters for an assumed geometry is closely related to the accuracy of the input dipolar couplings, which is on the order of 1% for the stronger couplings. We estimate that the **S** matrix elements are accurate to 1−2%; they vary by less than that for different local input geometries. The orientation of the principal axes is shown. P_{II} is rather flat, and the smallest eigenvalue corresponds to the axis perpendicular to the plane of the molecule. The orientation of the principal axes of **S** bears a strong resemblance to the orientation of the principal axes of the inertial tensor, but we see no physical reason that this should be so. We provide Cartesian coordinates relative to the chosen molecule-fixed axes in Supplementary Table S2. The best-fit **S** matrix expressed in these same coordinates is given in Supplementary Table S3.

The one-conformer fit to P_{II} allows us to compute orientational order parameters S_{ij} for each internuclear vector using r_{ij}^{-3} from the input geometry (eq 3, Table 1). For intramethyl couplings, we give the order parameter of the methyl rotor symmetry axis. Several of the order parameters are shown in Figure 5a. Large positive values mean that the vector orients preferentially along the magnetic field axis, perpendicular to the bicelle/water interface. Large negative values mean that the vector orients parallel to the bicelle/water interface. Small values mean little orientational order or preferred orientation about the magic angle $\theta = 54.7°$. The fact that S(Me1−Me3) is almost as large as the S_{11} eigenvalue of the **S** matrix means that the Me1−Me3 axis lies nearly parallel to the first principal axis of the orientation tensor, as seen clearly in the end view of Figure 5b. Similarly, the Me2 symmetry axis lies roughly parallel to the second principal axis. We discuss a physical model consistent with these observations in section IV.B below.

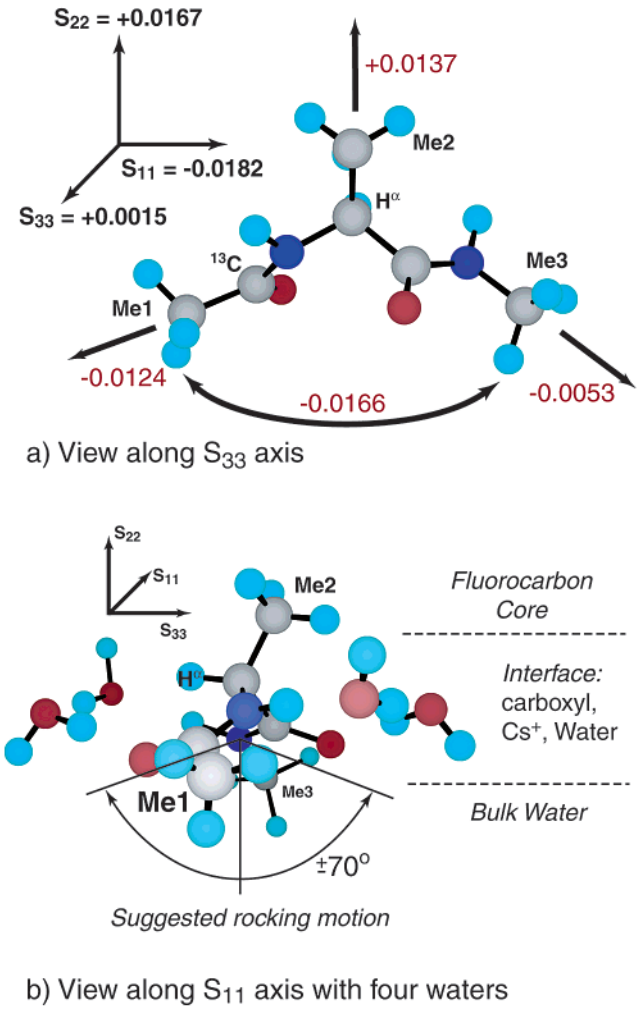

Figure 5. (a) Alanine dipeptide in P_{II} conformation placed in the Cartesian reference frame that diagonalizes the **S** matrix at the P_{II} minimum. Principal values of **S** are given. The S_{33} axis lies perpendicular to the plane of the page. Red numbers give selected values of the orientational order parameter S_{ij} for the three methyl rotor symmetry axes and for the Me1−Me2 average proton−proton axis. See Table 1 for complete list of S_{ij}. (b) View along the S_{11} axis including four explicit water molecules at the minimum-energy positions calculated for P_{II} in ref 3. The order parameters are consistent with a simple rocking motion as shown.

IV. Discussion

A. Comparison with Other Experimental Results. The dipolar coupling data in CsPFO/water strongly indicate that the P_{II} structure is the dominant conformer. A legitimate question is whether the liquid-crystalline medium substantially perturbs the structure of alanine dipeptide compared with an isotropic aqueous solution. After all, only configurations sampled by the dipeptide during excursions to the bicelle/water interface contribute to the measured dipolar couplings. To address this question, we investigated whether the P_{II} structure found in CsPFO/water is consistent with other experimental results in normal water.

Alanine dipeptide has been the subject of numerous experimental studies over the years. In early interpretations of CD and traditional NMR data, the P_{II} structure was considered likely to be important in aqueous solution, but C_7^{eq}, C_5^{ext}, and α_R were frequently mentioned as well.[8,9,35] The NMR scalar couplings $J(H^N,H^\alpha) = 6.0$ Hz and $J(C',H^\alpha) = 2.4$ Hz [here called $J(H^\alpha,{}^{13}C)$] can be correlated with the dihedral angle ϕ by fitting protein structural data to Karplus curves.[36,37] According to the most recent fits to experiment, the value of $J(H^N,H^\alpha)$ in aqueous solution points to $\phi = -74°$, whereas the value of $J(C',H^\alpha)$ points to $\phi = -96°$.[37] Both results lie within 11° (but on opposite sides) of our best-fit value of −85° derived from dipolar couplings in the liquid-crystal medium. Given the dispersion in the protein data that went into the fits, especially for $J(C',H^\alpha)$, this is fairly reasonable agreement. The difficulty is that none of the known scalar couplings is sensitive to ψ and the derived

values of ϕ do not incisively distinguish between P_{II} and α_R conformations. If the P_{II} conformer indeed dominates in aqueous solution as we expect, measurement of all available scalar couplings in thoroughly labeled alanine dipeptide could determine ψ as well as ϕ and definitively settle the question.

Linear vibrational spectra of alanine dipeptide have been obtained by IR absorption and Raman scattering, including differential circularly polarized variants. Interpretation of these spectra was the goal of Suhai et al.'s electronic structure study of the P_{II} conformer with four explicit water molecules in an Onsager cavity.[3] His calculated vibrational spectra for P_{II} indeed resemble the experimental spectra much more closely than calculations based on the gas-phase molecule. It remains difficult to make a unique assignment to P_{II}. In particular, some calculated bands of α_R match experiment also, but some strong bands do not.

Overall, the experimental information from scalar couplings and from vibrational spectra in H_2O and D_2O appears consistent with our conclusion that the dominant conformer is P_{II}. There is no evidence that the liquid-crystal medium perturbs peptide structure.

In closely related work, two new experimental studies have addressed the same conformational question in the molecule alanine *tripeptide*, which also has only two flexible backbone angles (one ϕ, one ψ) plus titratable carboxyl and amino groups at either end. Hamm and co-workers[38] obtained two-dimensional vibrational spectra of the tripeptide in D_2O at pD = 1 (both N and C termini protonated). Analysis using a local oscillator model with parameters transferred from electronic structure theory yields the angle between the two localized carbonyl transition dipole moment vectors and the coupling strength between them. The results are consistent with (ϕ, ψ) in the vicinity of ($-60°$, $+140°$), quite similar to our conclusions for alanine dipeptide. Schweitzer-Stenner and co-workers[39,40] obtained polarized visible Raman and FTIR spectra of trialanine in all three charge states by varying pD. A global theoretical fit to the data using a similar model finds that the preferred ϕ varies from $-66°$ to $-95°$ as pD increases from 1 to 6 to 12, while the preferred ψ varies only from $+140°$ to $+150°$ over the same range. These results are very similar to our conclusions for the neutral alanine dipeptide in CsPFO/water.

B. Orientational Preferences. The LX-NMR method provides information about the geometry of the dipeptide *while it samples the oriented bicelle/water interfacial region*. Perhaps P_{II} orients more strongly than α_R at the bicelle interface, so that the method is intrinsically less sensitive to α_R. By analogy to the lipid bilayer/water interface,[41] we expect the interface between the packed fluorocarbon chains of the bicelle core and bulk exterior water to be rich in water itself, carboxyl groups, and Cs^+ counterions.[42] To learn more about the orientation mechanism, we recently studied the entire family of mono-, di-, and trimethyl-substituted formamides in the same CsPFO/water system.[33,43] Remarkably, the strength of orientation as measured by the largest order parameter S_{ij} varied a factor of 25 from the weakly orienting formamide to the strongly orienting *trans-N*-methylformamide and *N,N*-dimethylformamide. In these simple amides, the presence of an *N*-methyl group trans to the amide carbonyl group apparently causes strong orientation within the interface and even preferential partitioning to the interface. For the strongly orienting amides, the S_{ij} values are consistent with a single, dominant orientation within the interface that "dips" the *trans*-methyl group into the fluorocarbon core of the bicelle while immersing the carbonyl group in water, carboxyl ions, and Cs^+ ions, a strong hydrogen-bonding environment.

P_{II} orients about 8 times more weakly than the two strongly orienting amides; it is comparable to the "moderate orienters" *N*-methylacetamide and *cis-N*-methylformamide. Accordingly, all relevant conformers of alanine dipeptide have NH, not $N-CH_3$, in the position trans to the amide carbonyls. The formamide study suggests that orientations of P_{II} that can simultaneously direct one or both carbonyl groups *outward* from the bicelle and one or more methyl groups *inward* toward the bicelle core might be preferred.

Indeed, the **S** matrix from the single-conformer fit to the P_{II} geometry permits a simple and appealing, *albeit speculative*, physical interpretation that can guide future work. As shown in Figure 5a, the fit yields a large *positive* order parameter for the Me2 axis ($+0.014$, preferential orientation parallel to the magnetic field, perpendicular to the bicelle surface) and an even larger negative order parameter (-0.017, preferential orientation parallel to the bicelle surface) for the Me1−Me3 axis. In a *perfectly* oriented sample, all S_{ij} values must lie in the range $+1$ to -0.5. However, as described in detail elsewhere, we estimate that in 42 wt % CsPFO complete orientation of the solute dipeptide *while it visits the interface* would give rise to order parameters in the approximate range -0.026 to $+0.052$. The estimate uses bicelle dimensions of an 80-Å diameter and a 22-Å thickness with 24% of the volume occupied by bicelle (core plus interface), taken from X-ray scattering measurements. We assume that the dipeptide partitions equally between bulk and interface, in proportion to volume (interfacial thickness assumed to be 4 Å, which gives bulk about 5 times larger volume than interface), so that the fraction of peptide at the interface is 0.12. We account for the imperfect orientation of the bicelles themselves ($S_{\text{bicelle}} \approx 0.9$)[44] and make a "cylindrical bicelle" approximation in which 64.5% of the bicelle surface is oriented perpendicular to the magnetic field and 35.5% parallel (the "belt" of the disk). In this context, $S_{ij} = -0.017$ for the Me1−Me3 axis is a large order parameter, roughly 65% as large as possible assuming no preferential partitioning. This motivates a simple interpretation. Evidently, the motional average within the bicelle/water interface leaves the Me1−Me3 axis strongly oriented parallel to the interfacial plane. The order parameter of the Me2 axis is about 25% of the maximum possible ($+0.014$ out of $+0.052$).

The large negative order parameter for the Me1−Me3 axis motivates a simple model of the orientational motion of alanine dipeptide at the interface. We suggest that the dipeptide rocks the S_{22} and S_{33} axes (Figure 5a) by some $\pm 70°$ by rotating about the S_{11} axis (an axis nearly parallel to the Me1−Me2 axis). This motion maintains the orientation of the Me2 axis essentially inward toward the fluorocarbon core (Figure 5b). Both the principal values of the **S** matrix and the signs and magnitudes of all of the S_{ij} components derived for P_{II} (Table 1) are consistent with this idea. The range of motion about the rocking axis must be large enough that the third principal axis (nearly perpendicular to both the Me2 and the Me1−Me3 axes) can sample values of θ on either side of the magic angle (54.7°) and average S_{33} to nearly zero. To illustrate that this is physically possible, we averaged $P_2(\cos\theta)$ over pure rotation about the S_{11} axis with uniform sampling of the angle of rotation over various ranges. For example, rotation in the range $0-68°$ moves the S_{22} axis over the range $\theta = 0-68°$ and the S_{33} axis over the range $\theta = 90-158°$ while holding the S_{11} axis at $\theta = 90°$. This gives motionally averaged eigenvalues of $S_{11} = -0.500$, $S_{22} = +0.466$, and $S_{33} = +0.034$ (on the scale $+1.0$ to -0.5). The ratios of these eigenvalues match experiment. Scaled to the feasible range $+0.052$ to -0.026, these become $S_{11} = -0.026$,

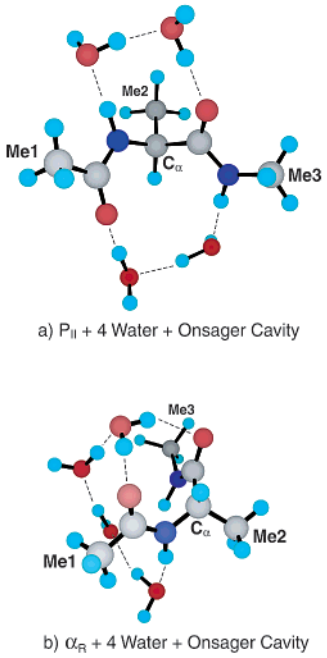

Figure 6. Geometries of energy minima in Suhai et al.'s B3LYP calculations of alanine dipeptide plus four explicit water molecules in an Onsager cavity (ref 3). (a) P_{II} minimum. (b) α_R minimum. Dashed lines indicate hydrogen-bonding contacts.

$S_{22} = +0.024$, and $S_{33} = +0.002$. These model numbers are about 1.5 times the experimental values, indicating that the real motion is not quite so simple.

We emphasize that a wide variety of orientational distributions are undoubtedly consistent with the observations. However, the suggested rocking motion is physically appealing and provides a sensible generalization of the ideas developed in the comparative study of substituted formamides in CsPFO/water.[33,43] According to the cluster electronic structure calculations of ref 3 (Figures 5b and 6a), the P_{II} conformer provides two "pockets" that expose the two CO···HN units simultaneously to solvent, allowing a network of hydrogen bonds to water without steric interference from H^{α} or any of the methyl groups. The two pockets lie on opposite faces of the P_{II} conformer, which is rather flat. We take the locations of the four water molecules in the cluster calculations as indicative of particularly hydrophilic regions of the dipeptide. In the suggested motion, the two hydrogen-bonding pockets are alternately exposed primarily to the interface or to the bulk, both of which provide good hydrogen-bonding environments. If this speculation is correct, the molecule prefers to dip Me2 into the fluorocarbon core rather than Me1 or Me3. Either of the latter two choices would permit exposure of both hydrogen-bonding pockets to water. However, both Me1 and Me3 lie one bond closer to a hydrogen-bonding NH or carbonyl group than Me2, perhaps rendering the two ends of the dipeptide "less hydrophobic" than the center. It would be very interesting to substitute ethyl for methyl at each position and measure the effect on orientation.

The competing conformer α_R places the two carbonyl axes parallel to each other on the same side of the dipeptide, which is the reason it has the largest dipole moment. The optimized geometry for α_R plus four waters in the Onsager cavity shows a single water bridging the two carbonyls and hydrogen-bonded in a chain to the other three waters (Figure 6b).[3] In the same spirit, a *single* orientation of Suhai et al.'s α_R could simultaneously aim both carbonyls and three of four waters outward from the bicelle surface while again dipping Me2 into the fluorocarbon core. Such an orientation favorably aligns the dipole moment with the interfacial electric field as well. On the basis of such ideas, α_R might well orient *more* strongly than P_{II}, which would make the LX-NMR experiment in CsPFO/water more sensitive to α_R than to P_{II}. However, these are speculative arguments, not incisive ones.

C. Comparison with Theoretical Results. The main goal of this work is to provide quantitative experimental benchmarks that test theoretical models of peptide solvation in water. The alanine dipeptide has become a standard test case for all solvation models.[2,45−49] A variety of model potentials have been used in large-scale MD simulations. Smith and co-workers review the results in their study using CHARMM22 for the dipeptide and TIP3P for the water.[47] Recent QM/MM simulations by Hu and Hermans used semiempirical quantum mechanical forces for the dipeptide and either SPC or TIP3P models for 362 surrounding water molecules in 6-ns simulations.[5] In a very different approach aimed at assigning vibrational spectra, Suhai and co-workers[3] carried out electronic structure calculations (density functional theory, B3LYP) on a cluster of four explicit water molecules about alanine dipeptide embedded in an Onsager cavity. We used their optimized geometries in fitting the dipolar coupling data. Nemukhin and co-workers[50] used a hybrid QM/MM method to study the same four-water cluster in isolation.

The most recent theoretical work consistently finds two regions of ϕ, ψ space to be most stable. These are the β/P_{II} region and the α_R region. By P_{II}, we mean $\phi \approx -75°$ and $\psi \approx +140°$; by β, we mean $\phi \approx -140°$ and $\psi \approx +150°$; and by α_R, we mean $\phi \approx -80°$ and $\psi \approx -50°$. In the most recent MD simulations with model force fields,[5,47] some variant of a β/P_{II} local minimum and the α_R minimum typically lie within 1−2 kcal/mol of each other. Perhaps coincidentally, the *energy* differences from electronic structure treatments of the dipeptide/water cluster in an Onsager cavity[3] are quite similar to the *free energy* differences from MD simulations. Among the MD simulations, the locations and breadths of the β and P_{II} minima; the heights of the barriers separating β from P_{II} and P_{II} from α_R; and the free energy differences among β, P_{II}, and α_R vary substantially with the choice of model potential. AMBER favors a localized α_R minimum; CEDAR, GROMOS, and OPLS favor both β and P_{II} geometries; and CHARMM22 finds localized α_R and P_{II} structures essentially equal in free energy. The recent large-scale QM/MM simulations at 300 K predict a *delocalized* structure for alanine dipeptide.[5] Trajectories spend 50% of the time in a broad β/P_{II} well with ϕ ranging from −60° to −180°, 30% of the time in a broad α_R well, and another 15% of the time in the transition region between the two.

We believe that the dipolar coupling data cannot be explained by a broad sampling of ϕ encompassing both β and P_{II} geometries. The data are extremely well fit by a *single geometry* P_{II}. We showed in section III.D that a single *centrally located* geometry can mimic dipolar couplings that would arise from uniformly sampled ±20° or even ±30° torsion in ϕ and ψ about the P_{II} minimum. Whereas the data are rather insensitive to the breadth of excursion about $\phi = -90°$, $\psi = +140°$, they are quite sensitive to the choice of the central geometry (Figure 4). If the molecule spent comparable fractions of time in both P_{II} and β, as simulations based on AMBER, GROMOS, and OPLS models suggest, then the best-fit geometry would lie in between, with $\phi \approx -120°$. Our preference for P_{II} rather than a broad β/P_{II} region is more consistent with the CHARMM22 results[47] (for which both the P_{II} and α_R wells are deep and well isolated) than with the other treatments, including the semiempirical QM/MM results.

The scalar coupling $J(H^N,H^\alpha) = +6.0$ Hz corroborates this conclusion. Protein data show that as ϕ varies from $-180°$ to $-120°$ to $-60°$, $J(H^N,H^\alpha)$ increases sinusoidally from 4 to 10 Hz and back down to 4 Hz.[36,37] $J(H^N,H^\alpha)$ reaches the experimental value 6 Hz twice, at $\phi \approx -165°$ and again at $\phi \approx -75°$. Frequent sampling of ϕ in the β-like range from $-120°$ to $-140°$ would drive the average value of $J(H^N,H^\alpha)$ above 6 Hz and is thus ruled out. In particular, a broad distribution of ϕ ranging from $-180°$ to $-60°$ and peaked near $-120°$ would give a $J(H^N,H^\alpha)$ value of about 8 Hz, in disagreement with experiment.

Intriguingly, the most recent nonlinear vibrational spectroscopic study of the closely related molecule alanine tripeptide at low pD infers the presence of a minority species in addition to P_{II} from the observed inhomogeneous line broadening of the amide I band.[51] Detailed MD and spectral simulations using the GROMOS96 force field in conjunction with quantum mechanical vibrational frequency calculations at the sampled geometries indicate that the data are consistent with the minority species being α_R at the level of about 20% population. For the fully protonated form of trialanine, the usual model potentials[6] differ in their predicted distributions of ϕ and ψ in ways quite reminiscent of alanine dipeptide. Both the vibrational results and scalar couplings from NMR spectroscopy are inconsistent with a contribution from β. Analogous nonlinear vibrational studies of alanine dipeptide might help in the search for a minority contribution from α_R.

D. Water Bridges. With P_{II} rather firmly established as the dominant conformer of alanine dipeptide in water, theoretical models that find a P_{II} minimum gain credence and provide new insight into the detailed nature of solvation of the dipeptide. As discussed by Suhai et al.,[3] *water greatly modifies the preferred structure of the peptide compared with the gas phase*. In making that assertion, we trust electronic structure theory to predict the geometries of gas-phase local minima correctly; there are no experimental data. A variety of calculations[3,46,52,53] agree that the gas-phase minima in the upper left quadrant of Ramachandran space comprise the structures C_7^{eq} ($\phi = -80°$, $\psi = +70°$), C_5^{ext} ($-160°$, $+165°$), and β_2 ($-135°$, $+25°$), shown as blue dots in Figure 4. Absent water, neither P_{II} nor α_R is an energy minimum in electronic structure theory or a free energy minimum in the CHARMM22 and other model potentials. According to the B3LYP calculations of Suhai et al.,[3] already in the *gas-phase cluster*, the presence of four explicit water molecules about alanine dipeptide destroys C_7^{eq} and C_5^{ext} as local minima, shifts β_2 to β_2' ($-150°$, $+115°$), and creates new minima at P_{II} ($-90°$, $+130°$), at α_R' ($-80°$, $-45°$) and at α_L' ($+60°$, $+50°$). Among the four-water-cluster minima, P_{II} is most stable, followed by β_2' (+1.9 kcal/mol), α_R' (+2.5 kcal/mol), and α_L' (+2.8 kcal/mol). Subsequent reoptimization of the four-water clusters within an Onsager cavity changes the *geometries* very little but renders P_{II} and α_R' equal in energy, followed by the left-handed helix α_L' (+1.2 kcal/mol) and β_2' (+1.9 kcal/mol). The polarizable medium preferentially stabilizes the two helical structures by about 2 kcal/mol because they have the largest electric dipole moments. Quite similar changes in geometry and stability of local minima occur using the CHARMM22/TIP3P force fields to compare gas-phase and solvated alanine dipeptide.[47]

What causes the changes in conformation on solvation? Although the energy minima are static, the detailed geometries from the water-cluster calculations[3] provide clues to the kinds of solvent configurations that are probably important in the real system of nonrigid peptide and mobile water molecules. Suhai et al.'s P_{II} minimum is shown in Figure 6a and in side view in Figure 5b. Faced with the limited range of choices available with four explicit water molecules, P_{II} makes two *double-water bridges*. Other conformers such as $C_7^{ax'}$ optimize with one *single-water* bridge and one *triple-water* bridge.[3] We can thus begin to see in detail how the gas-phase intrachain potential (which favors the C_5^{ext} and C_7^{eq} structures) combines with the effects of explicit hydrogen bonding to determine the overall free energy minimum. P_{II} seems better configured to make strong water bridges than the gas-phase conformers. As already mentioned, the P_{II} geometry provides two pockets in which water molecules can hydrogen bond with the two CO···HN units simultaneously, without steric interference from H^α or any of the methyl groups. It allows nearly linear hydrogen-bond angles throughout both double-water bridges, even though the two bridges differ substantially in overall length. The distances spanned by the two CO···HN units to which water dimers bind and the near coplanarity within each CO···HN unit might be important. In contrast, the C_7^{eq} conformer places the relevant CO···HN moieties at skewed angles awkward for water bridges. The C_5^{ext} conformer places each CO···HN unit nearly coplanar but interposes H^α between the outside CO···HN pair, perhaps disrupting hydrogen-bonding possibilities.

The solvated P_{II} structure of Figure 6a suggests that *cooperative hydrogen bonding* could be important in accurate prediction of peptide structure and energetics.[54] By cooperative, we mean that the total hydrogen-bond strength in a carbonyl−water−water−NH chain is substantially larger than the sum of analogous individual, isolated hydrogen-bond strengths. Charge transfer between components leads to progressively enhanced attractive interactions between adjacent groups as the chain increases in length. Eight δ^+, δ^- partial charges alternate around the "super chain" comprising ···amide−water−water−amide′−water−water··· in Figure 6a. In addition to cooperativity within each double-water bridge, the two bridges can interact by stabilizing the N^+/O^- resonance structure of *both* amide groups in a kind of interchain cooperativity.

The P_{II} structure fully utilizes the hydrogen-bonding capacity of the two CO···HN units, maximizes peptide−water cooperativity, and leaves all four waters able to participate in further hydrogen bonds with the next solvation "shell". In contrast, the α_R' structure (Figure 6b) places the two carbonyls parallel to each other on the same side of the dipeptide. The structure is more compact than P_{II}. The four waters form a linear chain connecting the two carbonyls to a single NH group on the opposite side of the molecule. The water at one end of the chain bridges both carbonyls. The α_R' structure also indicates substantial hydrogen-bonding cooperativity, but it resides predominantly in the water−water bonds of the chain. These waters could enjoy a similar interaction in bulk water. Thus, we suggest that the geometry change from gas phase to P_{II} is driven primarily by optimization of hydrogen-bonding capacity, whereas the geometry change from gas phase to α_R is driven primarily by the creation of a large dipole moment that orients surrounding water molecules to enhance stability. As discussed below, this idea has implications for the structure of small polypeptide chains in water.[24]

It is intriguingly difficult to demonstrate that explicit cooperativity of hydrogen bonding is necessary for accurate modeling. Such cooperativity is automatically present in ab initio calculations and has a substantial energetic effect in gas-phase water clusters. In model force fields, one way to obtain cooperativity is by inclusion of non-pairwise-additive terms involving polarizability in the model force field.[55] Yet, MD simulation using force fields that *are* pairwise-additive *also* find P_{II} and

α_R as the lowest free energy structures.[5,47] Energetic cooperativity might still be significant in these simulations, because it can arise in Coulombic, pairwise additive potentials by advantageous placement of an array of positive and negative charges.

Detailed theoretical analysis aimed at fully understanding the stability of the P_{II} geometry in aqueous solution is worthwhile. We have drawn rather general conclusions from the particular geometries that optimize the total electronic energy in Suhai et al.'s cluster calculations. Because electronic structure calculations and MD simulations (which include entropic effects) are in remarkable agreement on alanine dipeptide, it would be informative to extract from the simulations the joint probability distribution of single-, double-, and triple-water bridges at different sites for different peptide conformers and to learn how the distribution varies with model potential. Additional electronic structure calculations including more water molecules are highly desirable as well.

V. Conclusions

Quantitative fits of dipolar coupling constants for alanine dipeptide in CsPFO/D_2O with a single geometry are completely consistent with P_{II} being the dominant conformer. The data are fit within experimental error by P_{II} alone; in a statistical sense, additional conformers could not improve the fit. Our dipolar coupling data conflict with theoretical models that find broad, shallow wells encompassing both the P_{II} and β conformers. A limited set of dipolar couplings with imperfect knowledge of relative signs cannot mathematically rule out the possibility of two significant conformers. The critical experimental sign information for dipolar couplings could presumably be obtained by more sophisticated multidimensional NMR methods.

Perhaps the best argument against the presence of a substantial population of α_R that orients much more weakly than P_{II} in CsPFO/D_2O comes from recent vibrational spectroscopy on the closely related molecule alanine tripeptide *in aqueous solution*.[38−40] Although interpretation of both infrared and Raman spectra depends on theoretical input, the results again point to P_{II} as the dominant conformation.

In the aggregate, these results appear to corroborate long-standing assertions that short polypeptides, and, indeed, denatured proteins, predominantly consist of segments that persistently adopt the P_{II} conformation in aqueous solution. This idea was suggested in 1968 by Tiffany and Krimm on the basis of circular dichroism (CD) spectra.[21] It stands in sharp contrast to the usual random-coil model of polypeptides. More recent NMR and CD studies of small polypeptides in aqueous solution reach the same conclusion.[22,23] A simple theoretical model using a purely repulsive, soft-core potential between polyalanine atoms finds that P_{II}-like conformations dominate the low-lying features of the potential energy landscape.[24] We suggest that the strength of cooperative hydrogen bonding between P_{II}-like conformations and water drives the propensity for formation of extended P_{II} helical segments.

The conformational problem in small peptides remains a daunting one. There is still no clear prescription for solving two- or three-conformer problems incisively. At present, the best approach combines dipolar coupling data, *J* couplings, and NOE data from solution-phase NMR and vibrational spectroscopies and computational results. Interpretation of the data has relied substantially on theory, making it difficult to devise independent tests of different model potentials. For small systems, theory can probably identify the plausible conformers, but their relative free energies are not sufficiently accurate as yet. Experiments that help to define which conformers are important, to determine their relative free energies, and to characterize the breadth of free energy minima and the heights of the barriers connecting them remain very important.

Clear experimental determination of the relative importance of two or three free energy minima for a variety of small peptides would provide the most stringent test of the accuracy of theory and MD simulation. For this purpose, we believe that an extensive body of dipolar coupling data with clear sign information, augmented by all of the *J* couplings that modern NMR spectroscopy can measure, would provide the best model-independent test. Our work on alanine dipeptide only begins to suggest the true power of LX-NMR spectroscopy in this arena.

Acknowledgment. We especially thank Prof. Ed Samulski and Dr. Chi-Duen Poon for helping to conceive and launch this project while J.C.W. was on sabbatical at Chapel Hill. Dr. Charlie Fry of the UW-Madison instrument center helped us obtain the NMR spectrum at 500 MHz. The Gellman group assisted in synthesis of the labeled compound. We thank the NSF (CHE-0077517) for generous support of this research.

Supporting Information Available: Tables listing sign combinations of dipolar couplings that fit 500-MHz spectrum well, geometry of P_{II} conformer in the chosen molecule-fixed axis system, and elements of the best-fit **S** matrix for the P_{II} conformer expressed in the same axis system. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Karplus, M. *Acc. Chem. Res.* **2002**, *35*, 321.
(2) Pettitt, B. M.; Karplus, M. *J. Phys. Chem.* **1988**, *92*, 3994.
(3) Han, W. G.; Jalkanen, K. J.; Elstner, M.; Suhai, S. *J. Phys. Chem. B* **1998**, *102*, 2587.
(4) Edinger, S. R.; Cortis, C.; Shenkin, P. S.; Friesner, R. A. *J. Phys. Chem. B* **1997**, *101*, 1190.
(5) Hu, H.; Hermans, J. *Proteins*, manuscript submitted.
(6) Mu, Y.; Kosov, D. S.; Stock, G. *J. Chem. Phys.*, in press.
(7) Grenie, Y.; Avignon, M.; Garrigou-Lagrance, C. *J. Mol. Struct.* **1975**, *24*, 293.
(8) Deng, Z.; Polavarapu, P. L.; Fiord, S. J.; Hecht, L.; Barron, L. D.; Ewig, C. S.; Jalkanen, K. *J. Phys. Chem.* **1996**, *100*, 2025.
(9) Madison, V.; Kopple, K. D. *J. Am. Chem. Soc.* **1980**, *102*, 4855.
(10) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G. I.; Skelton, N. J. *Protein NMR Spectroscopy: Principles and Practice*; Academic Press: New York, 1996.
(11) Hamm, P.; Lim, M.; DeGrado, W. F.; Hochstrasser, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2036.
(12) Boden, N.; Corne, S. A.; Jolley, K. W. *J. Phys. Chem.* **1987**, *91*, 4092.
(13) Saupe, A.; Englert, G. *Phys. Rev. Lett.* **1963**, *11*, 462.
(14) Emsley, J. N.; Lindon, J. C. *NMR Spectroscopy Using Liquid Crystal Solvents*; Pergamon Press: Oxford, U.K., 1975.
(15) Gochin, M.; Pines, A.; Rosen, M. E.; Rucker, S. P.; Schmidt, C. *Mol. Phys.* **1990**, *69*, 671.
(16) Photinos, D. J.; Poliks, B. J.; Samulski, E. T.; Terzis, A. F.; Toriumi, H. *Mol. Phys.* **1991**, *72*, 333.
(17) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111.
(18) Prestegard, J. H. *Nat. Struct. Biol.* **1998**, *5*, 517.
(19) Poon, C.-D.; Samulski, E. T.; Weise, C. F.; Weisshaar, J. C. *J. Am. Chem. Soc.* **2000**, *122*, 5642.
(20) Weise, C.; Weisshaar, J. C. *J. Phys. Chem. B*, manuscript submitted.
(21) Tiffany, M. L.; Krimm, S. *Biopolymers* **1968**, *6*, 1379.
(22) Woody, R. W. *Adv. Biophys. Chem.* **1992**, *2*, 37.
(23) Shi, Z.; Woody, R. W.; Kallenbach, N. R. *Adv Protein Chem* **2002**, *62*, 163.
(24) Pappu, R. V.; Rose, G. D. *Protein Sci*. **2002**, *11*, 2437.
(25) Holmes, M. C.; Reynolds, D. J.; Boden, N. *J. Phys. Chem.* **1987**, *91*, 5257.
(26) Ramirez, B. E.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 9106.
(27) *Nuclear Magnetic Resonance of Liquid Crystals*; Emsley, J. W., Ed.; D. Reidel: Dordrecht, The Netherlands, 1985.
(28) Marat, K. *XSIM*; University of Manitoba: Winnipeg, Manitoba, Canada, 1998.
(29) Hansen, E. O. *Prog. NMR Spectrosc.* **1981**, *14*, 175.

(30) Flory, P. J. *Macromolecules* **1974**, *7*, 381.
(31) Diehl, P. Structure of Rigid Molecules Dissolved in Liquid Crystalline Solvents. In *Encyclopedia of Nuclear Magnetic Resonance*; Grant, D. M., Harris, R. K., Eds.; John Wiley and Sons: New York, 1996; Vol. 5; p 4591.
(32) Vaara, J.; Kaski, J.; Jokisaari, J.; Diehl, P. *J. Phys. Chem.* **1997**, *101*, 5069.
(33) Weise, C. Structure and Orientation of the Alanine Dipeptide and Small Amides in a Water-based Liquid Crystal. Ph.D. Thesis, University of Wisconsin-Madison, Madison, WI, 2003.
(34) Sykora, S.; Vogt, J.; Bosiger, H.; Diehl, P. *J. Magn. Reson.* **1979**, *36*, 53.
(35) Jalkanen, K. J.; Suhai, S. *Chem. Phys.* **1996**, *208*, 81.
(36) Hu, J. S.; Bax, A. *J. Am. Chem. Soc.* **1997**, *119*, 6360.
(37) Case, D. A.; Scheurer, C.; Bruschweiler, R. *J. Am. Chem. Soc.* **2000**, *122*, 10390.
(38) Woutersen, S.; Hamm, P. *J. Phys. Chem. B* **2000**, *104*, 11316.
(39) Schweitzer-Stenner, R.; Eker, F. *Biophys. J.* **2002**, *82*, 874.
(40) Schweitzer-Stenner, R. *Biophys. J.* **2002**, *83*, 523.
(41) White, S. H.; Wiener, M. C. The liquid-crystallographic structure of fluid lipid bilayer membranes. In *Membrane Structure and Dynamics*; Merz, K. M., Roux, B., Eds.; Birkhäuser: Boston, 1996; pp 127−144.
(42) Hiemenz, P. C.; Rajagopalan, R. *Principles of Colloid and Surface Chemistry*, 3rd ed.; Marcel Dekker: New York, 1997.
(43) Weise, C. F.; Weisshaar, J. C. *J. Phys. Chem.*, manuscript submitted.
(44) Johannesson, H.; Furo, I.; Halle, B. *Phys. Rev. E* **1996**, *53*, 4904.
(45) Anderson, A. G.; Hermans, J. *Proteins: Struct., Funct., Genet.* **1988**, *3*, 262.
(46) Tobias, D. J.; Brooks, C. L. *J. Phys. Chem.* **1992**, *96*, 3864.
(47) Smith, P. E. *J. Chem. Phys.* **1999**, *111*, 5568.
(48) Schmidt, A. B.; Fine, R. M. *Mol. Simul.* **1994**, *13*, 347.
(49) Smart, J. L.; Marrone, T. J.; McCammon, J. A. *J. Comput. Chem.* **1997**, *18*, 1750.
(50) Nemukhin, A. V.; Grigorenko, B. L.; Bochenkova, A. V.; Topol, I. A.; Burt, S. K. *J. Mol. Struct. (THEOCHEM)* **2002**, *581*, 167.
(51) Woutersen, S.; Mu, Y.; Kosov, D. S.; Stock, G.; Hamm, P. *J. Chem. Phys.* in press.
(52) Headgordon, T.; Headgordon, M.; Frisch, M. J.; Brooks, C. L.; Pople, J. A. *J. Am. Chem. Soc.* **1991**, *113*, 5989.
(53) Gould, I. R.; Cornell, W. D.; Hillier, I. H. *J. Am. Chem. Soc.* **1994**, *116*, 9250.
(54) Weinhold, F. *J. Mol. Struct. (THEOCHEM)* **1997**, *398*, 181.
(55) Banks, J. L.; Kaminski, G. A.; Zhou, R. H.; Mainz, D. T.; Berne, B. J.; Friesner, R. A. *J. Chem. Phys.* **1999**, *110*, 741.