Determination of Relative N−H^N, N−C′, C^α−C′, and C^α−H^α Effective Bond Lengths in a Protein by NMR in a Dilute Liquid Crystalline Phase

Marcel Ottiger and Ad Bax*

Contribution from the Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520

Received July 28, 1998. Revised Manuscript Received October 2, 1998

Abstract: Weak alignment of solute macromolecules with the magnetic field can be achieved in a dilute, aqueous liquid crystalline phase of planar phospholipid micelles, consisting of mixtures of dimyristoylphosphatidylcholine (DMPC) and dihexanoylphosphatidylcholine (DHPC). Alignment of proteins in such a medium is sufficiently weak to retain the simplicity of the isotropic solution NMR spectrum but strong enough to permit accurate measurement of residual one-bond dipolar couplings. Highly accurate one-bond N−H^N, C^α−H^α, C^α−C′, and C′−N and two-bond C′−H^N dipolar couplings were measured in ^{13}C/^{15}N-enriched ubiquitin. Together with knowledge of the protein's three-dimensional structure, the dipolar couplings permit calculation of the relative, vibrationally corrected average bond lengths for these interactions. Assuming a C′−N bond length of 1.329 Å (Engh, R. A.; Huber, R. *Acta Crystallogr.* **1992,** *A47,* 392−400), the relative C^α−C′ distance of 1.526 Å is found to be in excellent agreement with results from Engh and Huber (1.525 Å). Using a C′−N bond length of 1.329 Å as a reference, N−H^N (1.041 ± 0.006 Å) and C^α−H^α (1.117 ± 0.007 Å) are considerably longer than equilibrium or average internuclear distances derived from ab initio calculations, electron diffraction, neutron diffraction, or microwave spectroscopy. The increase in effective N−H^N and C^α−H^α bond lengths is attributed to a decrease in the corresponding dipolar couplings resulting from fast librations, which must be of considerably larger amplitude than the C^α−C′ and C′−N angular fluctuations. Accurate knowledge of the relative effective N−H^N, C^α−H^α, C^α−C′, C′−N, and two-bond C′−H^N effective internuclear distances is essential for determining the magnitude of the molecular alignment tensor, for using the dipolar couplings in macromolecular structure determination, and for extracting angular information from recently described cross correlation experiments.

Introduction

Liquid crystal NMR is a well-established technique for obtaining highly accurate structures of small rigid molecules.[1] In favorable cases, detailed structural information can even be obtained if the solute undergoes conformational averaging.[2] The extremely sensitive relation between dipolar coupling and internuclear distance and orientation relative to the magnetic field, together with the large number of internuclear couplings that can be determined in such small, strongly oriented molecules, indeed provides an enormous amount of valuable structural information.[1−3] The high degree of order often results in observable dipolar couplings between spins at opposite ends of the molecule. Unfortunately, the resulting multitude of couplings leads to intractable ^{1}H NMR spectra in molecules containing more than a dozen hydrogen atoms. One successful approach to scale down the degree of order, and thereby decrease the magnitude of the dipolar couplings and the spectral complexity, uses spinning of the liquid crystalline NMR sample at an angle relative to the magnetic field.[4] Although this approach is intrinsically very powerful, no application to biological macromolecules has been reported to date.

For proteins, with inherently very complex NMR spectra, only minute degrees of molecular alignment can be tolerated before the NMR spectra become intractable. Such weak alignment can be obtained either from the protein's own magnetic susceptibility anisotropy[5−9] or from using a very dilute aqueous lyotropic liquid crystalline medium[10,11] consisting of planar

(1) Saupe, A.; Englert, G. *Phys. Rev. Lett.* **1963,** *11,* 462−465. Emsley, J. W.; Lindon, J. C. *NMR Spectroscopy using liquid crystal solvents*; Pergamon Press: New York, 1975. Lesot, P.; Merlet, D.; Courtieu, J.; Emsley, J. W.; Rantala, T. T.; Jokisaari, J. *J. Phys. Chem. A* **1997,** *101,* 5719−5724.

(2) Emsley, J. W.; Luckhurst, G. R.; Stockley, C. P. *Proc. R. Soc. London, A* **1982,** *381,* 117−138. Samulski, E. T.; Toriumi, H. *J. Chem. Phys.* **1983,** *79,* 5194−5199. Photinos, D. J.; Samulski, E. T.; Toriumi, H. *J. Chem. Phys.* **1991,** *94,* 2758−2772. Polson, J. M.; Burnell, E. E. *J. Chem. Phys.* **1995,** *103,* 6891−6902.

(3) Diehl, P. In *Accurate Molecular Structures, Their Determination and Importance*; Domenicano, A., Hargittai, I., Eds.; Oxford University Press: Oxford, 1992; pp 299−321. Emsley, J. W. In *Encyclopedia of Nuclear Magnetic Resonance*; Grant, D. M., Harris, R. K., Eds.; Wiley: Chichester, 1996; pp 2788−2799.

(4) Fung, B. M. In *Encyclopedia of Nuclear Magnetic Resonance*; Grant, D. M., Harris, R. K., Eds.; Wiley: Chichester, 1996; pp 2745−2751. Caldarelli, S.; Hong, M.; Emsley, L.; Pines, A. *J. Phys. Chem.* **1996,** *100,* 18696−18701.

(5) Bastiaan, E. W.; MacLean, C.; van Zijl, P. C. M.; Bothner-By, A. A. *Annu. Rep. NMR Spectrosc.* **1987,** *9,* 35−77. Bothner-By, A. A. In *Encyclopedia of Nuclear Magnetic Resonance*; Grant, D. M., Harris, R. K., Eds.; Wiley: Chichester, 1996; pp 2932−2938.

(6) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1995,** *92,* 9279−9283.

(7) King, H. C.; Wang, K. Y.; Goljer, I.; Bolton, P. H. *J. Magn. Reson. Ser. B* **1995,** *109,* 323−325.

(8) Tjandra, N.; Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1996,** *118,* 6264−6272.

(9) Tjandra, N.; Omichinski, J. G.; Gronenborn, A. M.; Clore, G. M.; Bax, A. *Nat. Struct. Biol.* **1997,** *4,* 732−738.

(10) Bax, A.; Tjandra, N. *J. Biomol. NMR* **1997,** *10,* 289−292.

phospholipid micelles, often referred to as bicelles.[12] Mixtures of dimyristoylphosphatidylcholine (DMPC) or ditridecanoylphosphatidylcholine (DTPC) and dihexanoylphosphatidylcholine (DHPC) in water form a particularly stable nematic liquid crystalline phase over a large range of conditions.[12−14] The thickness of the bicelles corresponds to that of the DMPC or DTPC bilayer (∼40 Å), and the diameter is a function of the DMPC:DHPC or DTPC:DHPC ratio, and of the bicelle concentration and temperature,[13,14] with typical values in the 400−800 Å range.

The bicelles themselves, and molecules anchored in these phospholipid bilayers, are both highly ordered.[15] However, proteins in the aqueous phase, which separates the bicelles by many hundreds of angstroms, are on average only very weakly aligned. Changing the interbicelle distance by altering their concentration can be used to "tune" the degree of protein alignment[10] such that one-bond dipolar couplings in the protein are sufficiently large to permit their accurate measurement, but small enough to avoid increased complexity of the NMR spectrum.[11]

One-bond dipolar couplings have recently been included as constraints in the NMR structure determination of several proteins.[9,16,17] Provided that the effective internuclear distances are accurately known, the dipolar couplings translate directly into angular constraints for the corresponding bond vectors relative to the molecular alignment tensor. These constraints are therefore fundamentally different from the strictly local NOE and *J* coupling parameters, which constrain atom positions only relative to those in their immediate vicinity. Inclusion of such dipolar constraints has been shown to improve considerably the accuracy of NMR structures, as indicated by a reduction of ϕ/ψ pairs outside of the most-favored region of the Ramachandran map[9] and improved agreement between predicted and measured changes in ^{15}N and $^{13}C'$ shifts as a function of protein alignment.[17,18]

Dipolar couplings between two atoms A and B are related to the inverse cube of the internuclear distance, $\langle r_{AB}^{-3}\rangle$, where the brackets indicate vibrational averaging. When different types of one-bond dipolar couplings are measured for a given protein, accurate knowledge of their relative $\langle r_{AB}^{-3}\rangle$ values is needed to incorporate such couplings simultaneously in the structure calculation protocol. For a protein of unknown structure, the magnitude of the alignment tensor is derived from the distributions of dipolar couplings observed for the various types of interactions, normalized relative to the dipolar couplings observed for, for example, backbone amide $N-H^N$ interactions.[19] Besides scaling for the magnetogyric ratios of the nuclei involved, this normalization also requires multiplication by $\langle r_{AB}^3\rangle/\langle r_{NH}^3\rangle$ values.

Which $N-H^N$ bond length to use in ^{15}N relaxation studies is the subject of ongoing debate, with neutron diffraction yielding larger values than ab initio calculations.[20] $N-H^N$ bond lengths obtained from neutron diffraction range from 1.02 to 1.04 Å.[20,21] Solid-state NMR on model peptides yields distances of 1.06−1.07 Å for the $N-H^N$ bond,[22,23] and slightly shorter distances (1.04−1.05 Å) for $N-D^N$ bonds.[24] These distances are derived from the corresponding dipolar couplings observed in solid-state NMR spectra. However, as the dipolar couplings are attenuated by fast angular fluctuations (librations) of the $N-H^N$ bond vector, the apparent internuclear distances derived in this manner are expected to be larger than the equilibrium distances.

Typical ab initio values for the equilibrium C−H bond length, r_e, of sp^3-hybridized carbons are 1.085 ± 0.005 Å.[20] Electron diffraction and microwave spectroscopy yield bond lengths which are approximately 0.01 Å longer,[25] as expected on the basis of the anharmonicity of the bond stretching.[25,26] Similarly, values of 1.09−1.10 Å are obtained from high-resolution neutron diffraction structures.[20]

Here we present optimized measurements of one-bond and two-bond dipolar couplings in the protein ubiquitin, dissolved in dilute liquid crystalline media. We demonstrate that even the small dipolar couplings between ^{13}C and ^{15}N can be measured with high accuracy and that the dipolar couplings correlate well with the X-ray structure of this protein. Knowledge of the crystal structure permits us to calculate relative effective lengths for the $N-H^N$, $C^\alpha-H^\alpha$, $C^\alpha-C'$, and $C'-N$ bonds and the $C'-H^N$ distance. These parameters are essential for determining the relative ratios of the magnitudes of the different dipolar coupling tensors and also provide information on the amplitude of the fluctuations of $N-H^N$ and $C^\alpha-H^\alpha$ bond vectors relative to those of $N-C'$ and $C^\alpha-C'$. Cross-correlation between dipole−dipole and either dipole−dipole or chemical shift anisotropy relaxation mechanisms contains quantitative information on the relative orientation of the two tensors, which provides a method to derive dihedral angle information.[27] Accurate knowledge of the one-bond $^1H-^{15}N$ and $^1H-^{13}C$ dipolar couplings is also needed for this approach.

Experimental Section

Two samples of uniformly $^{13}C/^{15}N$-enriched ubiquitin (VLI Research, Southeastern, PA) in liquid crystalline medium were prepared as described previously.[14] For one sample, the liquid crystal consisted of 50 mg/mL (5% w/v) DMPC/DHPC, with a molar ratio of $q = 3.0$, in 10 mM phosphate buffer, pH 6.6, 93% H_2O, 7% D_2O. The second sample was identical to the first, except that a small amount of cetyltrimethylammonium bromide (CTAB) was added to yield a molar ratio of DMPC:DHPC:CTAB = 30:10:1. Addition of CTAB can stabilize dilute liquid crystalline samples,[28] but more importantly for

(11) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111−1114.

(12) Sanders, C. R.; Schwonek, J. P. *Biochemistry* **1992**, *31*, 8898−8905. Sanders, C. R.; Landis, G. C. *Biochemistry* **1995**, *34*, 4030−4040. Ram, P.; Prestegard, J. H. *Biochim. Biophys. Acta* **1988**, *940*, 289−294. Sanders, C. R.; Prestegard, J. H. *Biophys. J.* **1990**, *58*, 447−460.

(13) Vold, R. R.; Prosser, P. S. *J. Magn. Reson.* **1996**, *B113*, 267−271.

(14) Ottiger, M.; Bax, A. *J. Biomol. NMR* **1998**, *12*, 361−372.

(15) Salvatore, B. A.; Ghose, R.; Prestegard, J. H. *J. Am. Chem. Soc.* **1996**, *118*, 4001−4008. Sanders, C. R.; Hare, B. J.; Howard, K. P.; Prestegard, J. H. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, *26*, 421−444. Metz, G.; Howard, K. P.; van Liemt, W. B. S.; Prestegard, J. H.; Lugtenburg, J.; Smith, S. O. *J. Am. Chem. Soc.* **1995**, *117*, 564−565.

(16) Bewley, C. A.; Gustafson, K. R.; Boyd, M. R.; Covell, D. G.; Bax, A.; Clore, G. M.; Gronenborn, A. M. *Nat. Struct. Biol.* **1998**, *5*, 571−578.

(17) Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 6836−6837.

(18) Ottiger, M.; Tjandra, N.; Bax, A. *J. Am. Chem. Soc.* **1997**, *119*, 9825−9830.

(19) Clore, G. M.; Gronenborn, A. M.; Bax, A. *J. Magn. Reson.* **1998**, *133*, 216−221.

(20) Jeffrey, G. In *Accurate Molecular Structures, Their Determination and Importance*; Domenicano, A., Hargittai, I., Eds.; Oxford University Press: Oxford, 1992; pp 270−298.

(21) Lehmann, M. S.; Koetzle, T. F.; Hamilton, W. C. *J. Am. Chem. Soc.* **1972**, *94*, 2657−2660. Koetzle, T. F.; Golic, L.; Lehmann, M. S.; Verbist, J. J.; Hamilton, W. C. *J. Chem. Phys.* **1974**, *60*, 4690−4696.

(22) Roberts, J. E.; Harbison, G. S.; Munowitz, M. G.; Herzfeld, J.; Griffin, R. G. *J. Am. Chem. Soc.* **1987**, *109*, 4163−4169.

(23) Wu, C. H.; Ramamoorthy, A.; Gierasch, L. M.; Opella, S. J. *J. Am. Chem. Soc.* **1995**, *117*, 6148−6149.

(24) Hiyama, Y.; Niu, C.-H.; Silverton, J. V.; Bavoso, A.; Torchia, D. A. *J. Am. Chem. Soc.* **1988**, *110*, 2378−2383.

(25) Demaison, J.; Wlodarczak, G. *Struct. Chem.* **1994**, *5*, 57−66.

(26) Henry, E. R.; Szabo, A. *J. Chem. Phys.* **1985**, *82*, 4753−4761.

(27) Reif, B.; Hennig, M.; Griesinger, C. *Science* **1997**, *276*, 1230−1233. Yang, D.; Konrat, R.; Kay, L. E. *J. Am. Chem. Soc.* **1997**, *119*, 11938−11940. Feng, X.; Eden. M.; Brinkmann, A.; Luthman, H.; Eriksson, L.; Graslund, A.; Antzutkin, O. N.; Levitt, M. H. *J. Am. Chem. Soc.* **1997**, *119*, 12006−12007. Brutscher, B.; Skrynnikov, N. R.; Bremi, T.; Brüschweiler, R.; Ernst, R. R. *J. Magn. Reson.* **1998**, *130*, 346−351.

the present application it can also significantly change the orientation and magnitude of the protein alignment tensor.[29] Final protein concentrations were 0.7 mM in 240 μL volumes, using thin-wall Shigemi microcells (Shigemi Inc., Allison Park, PA).

All NMR spectra were recorded on a Bruker DMX600 spectrometer operating at a ^{1}H resonance frequency of 600 MHz equipped with a triple-resonance, three-axis pulsed field gradient probehead. Data sets in the aligned state were recorded at 35 °C, and isotropic spectra were recorded at 25 °C. Special precautions were taken to ensure that the temperatures and the degree of protein alignment remained unchanged during each entire set of measurements (see below). Spectra were processed using the NMRPipe software package,[30] and peak positions were determined by contour averaging using the program PIPP,[31] as described previously.[32]

Dipolar couplings were derived from the difference in splittings in the aligned and the isotropic states. ${}^1D_{N-HN}$ dipolar couplings were extracted from 2D IPAP [^{15}N,^{1}H]-HSQC spectra[33] which were recorded as data matrices of 2 × 256* × 1024* complex points (n* denotes n complex points), with acquisition times of 128 ms (t_1, ^{15}N) and 57.0 ms (t_2, ^{1}H). For measurement of ${}^1D_{C^\alpha H^\alpha}$, 3D constant-time (HA)CA-(CO)NH spectra[34] without ^{1}H-decoupling during ^{13}C$^\alpha$ evolution were recorded as data matrices of 64* × 42* × 1024* complex points, with acquisition times of 25.3 ms (t_1, ^{13}C$^\alpha$), 29.4 ms (t_2, ^{15}N), and 57.0 ms (t_3, ^{1}H). ${}^1D_{C^\alpha-C'}$ values were obtained from 2D H(N)CO experiments[35] which were recorded without decoupling of ^{13}C$^\alpha$ during carbonyl evolution and as data matrices of 300* × 1024* complex points, with acquisition times of 153 ms (t_1, ^{13}C$'$) and 57.0 ms (t_2, ^{1}H). Both ${}^1D_{C'-N}$ (measured along F_1) and ${}^2D_{C'-HN}$ (along F_2) were obtained from 2D [^{15}N,^{1}H]-HSQC spectra, recorded without ^{13}C$'$ decoupling during ^{15}N evolution and as data matrixes of 512* × 1024* complex points, with acquisition times of 282 ms (t_1, ^{15}N) and 57.0 ms (t_2, ^{1}H). To increase ^{15}N resolution, composite pulse decoupling of in-phase ^{15}N−{^{1}H^N} magnetization was used.[36] The actual pulse scheme (Figure 1) uses mixed-constant-time evolution[37] in the t_1 dimension and therefore yields an additional 5−20% narrower ^{15}N line widths compared to the earlier experiment,[36] and 40−50% better resolution than a conventional [^{15}N,^{1}H]-HSQC spectrum.[36,38] To reduce t_1-noise-like artifacts and signal overlap resulting from the side chain NH$_2$ signals, an additional, phase-cycled ^{1}H filtering pulse is added immediately prior to t_1 ^{1}H decoupling.

Acquired data were apodized with a squared sine-bell in the directly detected dimension and a sine-bell in the indirectly detected dimension-(s), both shifted by 72° and truncated at 176°. Data were extensively zero-filled prior to Fourier transformation to yield high digital resolution.

Even though the magnitude of the alignment is not very sensitive to small temperature changes (ca. 0.3% per °C at 35 °C),[14] special care was taken to ensure constant sample conditions during acquisition of the experiments in the aligned state: All four experiments were recorded in one series without interruption, and the average applied radio frequency power per scan within each experiment and between the four experiments was kept constant by insertion of dummy decoupling periods prior to the relaxation delay.[39] Recycling periods
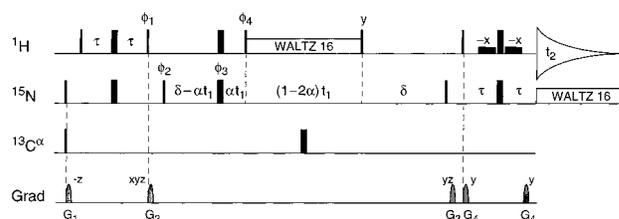
(28) Losonczi, J. A.; Prestegard, J. H. *J. Biomol. NMR*, in press.

(29) Ramirez, B. E.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 9106−9107.

(30) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277−293.

(31) Garrett, D. S.; Powers, R.; Gronenborn, A. M.; Clore, G. M. *J. Magn. Reson.* **1991**, *95*, 214−220.

(32) Wang, A. C.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 2483−2494.

(33) Ottiger, M.; Delaglio, F.; Bax, A. *J. Magn. Reson.* **1998**, *131*, 373−378.

(34) Tjandra, N.; Bax, A. *J. Am. Chem. Soc.* **1997**, *119*, 9576−9577.

(35) Grzesiek, S.; Bax, A. *J. Magn. Reson.* **1992**, *96*, 432−440. Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G., III; Skelton, N. J. *Protein NMR Spectroscopy. Principles and Practice*; Academic Press: San Diego, 1995; pp 499−500.

(36) Bax, A.; Ikura, M.; Kay, L. E.; Torchia, D. A.; Tschudin, R. *J. Magn. Reson.* **1990**, *86*, 304−318.

(37) Grzesiek, S.; Bax, A. *J. Biomol. NMR* **1993**, *3*, 185−204. Logan, T. M.; Olejniczak, E. T.; Xu, R. X.; Fesik, S. W. *J. Biomol. NMR* **1993**, *3*, 225−231.

(38) Bodenhausen, G.; Ruben, D. J. *Chem. Phys. Lett.* **1980**, *69*, 185−189.

Figure 1. Pulse scheme of the semi-CT HSQC experiment for measurement of ${}^1D_{C'N}$ and ${}^2D_{C'HN}$. Narrow and wide pulses correspond to 90 and 180° flip angles, respectively. Unless indicated, pulses are applied with phase x. The ^{13}C$^\alpha$ 180° pulse applied during the t_1 evolution period to decouple ^{13}C$^\alpha$ from ^{15}N has a duration of $\sqrt{3}/(2\Delta\delta)$ s (where $\Delta\delta$ is the frequency difference between the centers of the ^{13}C$^\alpha$ and ^{13}C$'$ regions). For optimal ^{13}C$^\alpha$ decoupling, this pulse is applied at time $t_1/2$ prior to the last 90° ^{15}N pulse. In practice, it may be simpler to program this pulse to be located at the midpoint of the ^{1}H composite pulse decoupling period, particularly when using gradient-enhanced coherence pathway selection by inserting an encoding gradient in the second δ delay and changing the back-INEPT accordingly.[61] The low power 90°$_{-x}$ pulses surrounding the final ^{1}H 180° pulse are part of the WATERGATE scheme,[62] needed for solvent suppression if no coherence pathway selection is used. Delay durations: $\tau = 2.25$ ms; $\delta = 5.3$ ms. The value of α equals $\delta/t_{1,\text{max}}$. Phase cycling: $\phi_1 = y, -y$; $\phi_2 = 2(x), 2(-x)$; $\phi_3 = 4(x), 4(y), 4(-x), 4(-y)$; $\phi_4 = 8(y), 8(-y)$; Receiver $= x, 2(-x), x, -x, 2(x), -x$. All gradients are sine-bell shaped with 25 G/cm at their center. Gradient durations: $G_{1,2,3,4} = 5, 1.1, 2.0, 0.5$ ms. Quadrature detection in the t_1 dimension is obtained by altering ϕ_2 in the usual States-TPPI manner.

were kept relatively long (about 1.8 s). Each series of experiments in the liquid crystalline state was started with a 4 h dummy experiment to allow complete equilibration of the sample in the magnet, and several short ^{1}H-coupled [^{15}N,^{1}H]-HSQC experiments were recorded within each series of experiments. Dipolar couplings measured from these HSQC spectra were used to check that the degree of protein alignment had remained constant during the whole series of measurements.

Fits of the dipolar coupling data to the X-ray structure of ubiquitin[40] to obtain orientation and magnitude of the rhombic alignment tensor were made with in-house written software, using a Powell minimization procedure as described previously.[8] Proton positions were added to the crystal structure with the program MOLMOL.[41]

The following nomenclature convention is used: the two coupled nuclei are listed in sequential order, such that the first atom is always closer to the N-terminus; for interresidue couplings the residue assignment is based on the NH group (e.g., $D_{C'N}$ is the residual dipolar coupling between C$'_{i-1}$ and N$_i$, where i is the residue number).

Results

Theoretical Background. In a suitably chosen reference frame, the dipolar coupling between two nuclei, A and B, in a solute macromolecule of fixed shape can be written as

$$D^{AB}(\theta,\phi) = D_a^{AB}\{(3\cos^2\theta - 1) + (3/2)R(\sin^2\theta\cos 2\phi)\} \quad (1)$$

where R is the rhombicity defined by D_r^{AB}/D_a^{AB}; D_a^{AB} and D_r^{AB} (in units of hertz) are the axial and rhombic components of the traceless second rank diagonal tensor D given by $(1/3)[D_{zz}^{AB} - (D_{xx}^{AB} + D_{yy}^{AB})/2]$ and $(1/3)[D_{xx}^{AB} - D_{yy}^{AB}]$, respectively, with $|D_{zz}^{AB}| > |D_{yy}^{AB}| \geq |D_{xx}^{AB}|$; θ is the angle between the A−B interatomic vector and the z axis of the tensor; and ϕ is the angle which describes the position of the projection of the A−B interatomic vector on the $x-y$ plane, relative to the x axis.

(39) Wang, A. C.; Bax, A. *J. Biomol. NMR* **1993**, *3*, 715−720.

(40) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531−544.

(41) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 52−55.

D_a^{AB} subsumes various constants, including the gyromagnetic ratios of the two nuclei γ_A and γ_B, the inverse cube of the distance between the two nuclei, $\langle r_{AB}^{-3} \rangle$, where the brackets indicate vibrational averaging, the generalized order parameter **S** for fast angular fluctuations of the internuclear vector[42] which provides a first-order correction for the effect of rapid internal motion on D^{AB},[8,43] and the unitless axial component, A_a, of the molecular alignment tensor **A**:

$$D_a^{AB} = -(\mu_o h / 16\pi^3)\mathbf{S}\gamma_A\gamma_B\langle r_{AB}^{-3}\rangle A_a \qquad (2)$$

A has the same meaning as the diagonalized Saupe ordering matrix;[1] the symbol **A** is used because in protein NMR, **S** is the symbol for the Lipari−Szabo generalized order parameter describing the effect of internal angular fluctuations.[42] Describing the effect of internal motions by a simple scaling constant **S** assumes that the alignment tensor is not influenced by the internal motions. For a globular protein this assumption is reasonable because (a) the internal motions occur on a much faster time scale than the overall rotational diffusion and (b) the internal motions have a negligible effect on the overall shape of the protein.

Measurement of Dipolar Couplings. Experiments were carried out for ubiquitin dissolved in two similar liquid crystalline phases which yield different alignment tensor orientations and magnitudes. For the bicelles consisting of a pure DMPC/DHPC mixture, the alignment tensor is nearly axially symmetric and its orientation approximately coincides with those of the inertia and rotational diffusion tensors.[11] When a small amount of positive charge is deposited on the bicelle surface by adding a small mole fraction of CTAB, weak electrostatic interaction between the bicelle and the protein results in an increase in protein ordering for ubiquitin and an alignment tensor orientation which is no longer dominated by the shape of the protein.[29] Dipolar couplings measured in this second liquid crystalline sample are in equally good agreement with the X-ray structure of the protein as those measured in the uncharged bicelle sample, indicating that the structure of the protein is not significantly affected by the very weak electrostatic interaction with the membrane. As the alignment tensor orientation and rhombicity differ substantially for the two samples, the dipolar couplings measured for ubiquitin are also quite different.[29] Availability of the two data sets therefore improves the accuracy of the relative effective bond lengths.

$^{15}N-^1H^N$ dipolar couplings were measured using the IPAP [^{15}N,^1H]-HSQC experiment, which yields a $^{15}N-\{^1H^N\}$ doublet in the F_1 dimension in which the upfield and downfield doublet components can be separated into different spectra in order to minimize spectral overlap.[33] Figure 2A shows the superimposed spectra containing the downfield and upfield Ile[13] doublet components and indicates an increase in splitting from 93.3 to 100.2 Hz when going from the isotropic to the liquid crystalline state, corresponding to a dipolar contribution of 6.9 Hz.

$^{13}C^\alpha-^1H^\alpha$ dipolar couplings were previously measured from a ^1H-coupled [$^{13}C^\alpha$,$^1H^\alpha$]-CT-HSQC experiment.[11] Because even the decoupled CT-HSQC experiment typically already shows considerable overlap, it is advantageous to record such couplings from a 3D experiment. The (HA)CA(CO)NH triple resonance experiment, previously used for measuring relaxation interference between the $^{13}C^\alpha$ CSA and the $^{13}C^\alpha-^1H^\alpha$ dipolar coupling,[34] is ideally suited for this purpose and yields well-resolved $^{13}C^\alpha-\{^1H^\alpha\}$ doublets. As an example, Figure 2B compares the
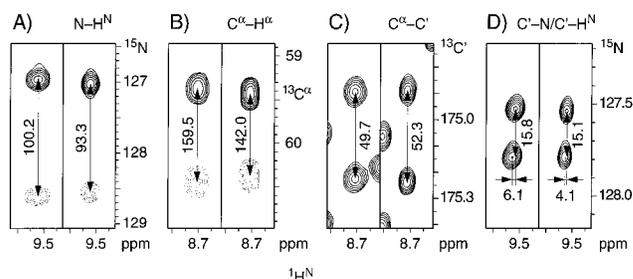
(42) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559.
(43) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. *Nat. Struct. Biol.* **1997**, *4*, 292−297.

Figure 2. Small regions of the aligned (left strip) and isotropic (right strip) of (A) the 2D IPAP [^{15}N,^1H]-HSQC, (B) the 3D constant-time (HA)CA(CO)NH, (C) the 2D H(N)CO, and (D) the 2D [^{15}N,^1H]-HSQC spectra, for measurement of $^1D_{NH}$, $^1D_{C\alpha H\alpha}$, $^1D_{C\alpha C'}$, and $^1D_{C'N}/^2D_{C'HN}$, respectively. All doublets shown correspond to Ile[13] of ubiquitin, and the $J + D$ splittings are indicated in Hz. In (A), the sum and the difference of the in-phase and antiphase spectra are overlaid and represented by thick and thin contours, respectively. In (B), positive and negative contours are drawn with thick and thin lines, respectively. All spectra were recorded with the uncharged liquid crystal sample at 35 °C (aligned) and 25 °C (isotropic). The strips each have a width of 0.2 ppm in the ^1H dimension.

$^{13}C^\alpha-\{^1H^\alpha\}$ doublet of Ile[13] in the isotropic and aligned state. The increase in splitting upon alignment corresponds to a 17.5 Hz dipolar coupling.

Previously, $^{13}C^\alpha-^{13}C'$ dipolar couplings were also obtained from a [$^{13}C^\alpha$,$^1H^\alpha$]-CT-HSQC experiment, in that case without $^{13}C'$ decoupling in the t_1 dimension.[11] As the dispersion of such a 2D spectrum is relatively poor, and the resolution in the t_1 dimension is limited by the duration of the constant-time $^{13}C^\alpha$ evolution period ($\sim$28 ms), a better way to measure these couplings uses the HNCO experiment.[35] By removing the $^{13}C^\alpha$ decoupling pulse, normally applied at the midpoint of the $^{13}C'$ evolution period, $^{13}C'-\{^{13}C^\alpha\}$ doublets are obtained. The resolution in this dimension is now determined by the favorable transverse relaxation properties of the $^{13}C'$ resonance, permitting more accurate measurement of this coupling. Note that the $^{13}C'$ line width rapidly increases with magnetic field strength and that such measurements therefore are best carried out at 500 MHz ^1H frequency or lower. In our case, where all experiments needed to be recorded under identical conditions, experiments were carried out at 600 MHz, however. For ubiquitin, resonance overlap in the 600 MHz 2D H(N)CO version of the experiment is relatively limited, and this 2D version was used to rapidly measure the $^{13}C^\alpha-^{13}C'$ couplings with high accuracy. Figure 2C compares the $^{13}C'-\{^{13}C^\alpha\}$ doublets of Ile[13] in the isotropic and aligned states. A decrease from 52.3 to 49.7 Hz upon alignment corresponds to a dipolar contribution of −2.6 Hz.

Owing to the low gyromagnetic ratio of ^{15}N, $^{13}C'-^{15}N$ dipolar interactions are quite small (eq 2). However, at 600 MHz ^1H frequency, ^{15}N also has the longest transverse relaxation time and therefore offers the best resolution of all NMR observable nuclei in a protein. ^{15}N resolution is increased by 40−50% on average by conducting the semiconstant time experiment in the "in-phase" mode of Figure 1, over what is obtained in the regular HSQC experiment.[36,38] For proteins much larger than ubiquitin, and particularly for perdeuterated proteins, it becomes advantageous to use the effect of relaxation interference between the ^{15}N CSA and the $^{15}N-^1H$ dipolar interaction, so that accurate $D_{C'N}$ couplings can be measured in proteins of 30 kDa and possibly larger.[44,45] As $^{13}C'$ decoupling is omitted in the pulse scheme of Figure 1, the $^{15}N-^1H$ correlation shows an E.COSY pattern[46] (Figure 2D), where the vertical splitting is the $^{15}N-^{13}C'$ coupling and the horizontal displacement in the ^1H dimension corresponds to the two-bond $^{13}C'-^1H^N$ coupling.[47]

Table 1. D_a Values for Five Internuclear Interactions in Ubiquitin, Measured in Two Liquid Crystalline Phases

	N$-$H^N	C$^\alpha$$-H^\alpha$	C$^\alpha$$-C'$	C$'$$-$N	C$'$$-$H^N
Uncharged LC					
number of D couplings					
measured	68	66[a]	64	67	67
used for fit	58	58	55	59	58
D_a values from fit (Hz)					
X-RAY1[b,e]	9.42	-18.87	-1.87	1.14	-2.99
X-RAY2[c,e]	9.47	-18.98	-1.88	1.15	-2.94
X-RAY3[d,e]	9.71	-19.11	-1.87	1.16	-2.94
error analysis					
σ_{fit}[f] (%)	1.9	1.9	1.6	2.2	3.3
RMSD[g] (%)	1.5	1.7	1.4	2.3	2.7
Charged LC					
number of D couplings					
measured	69	66[a]	59	69	67
used for fit	58	58	50	59	57
D_a values from fit (Hz)					
X-RAY1[b,e]	15.85	-32.08	-3.14	1.92	-5.21
X-RAY2[c,e]	15.79	-32.21	-3.17	1.93	-5.17
X-RAY3[d,e]	16.03	-32.45	-3.20	1.94	-5.22
error analysis					
σ_{fit}[f] (%)	2.0	2.5	1.9	1.7	3.0
RMSD[g] (%)	2.2	2.1	1.4	1.8	2.6

[a] Excluding Gly C$^\alpha$H$_2$. [b] X-ray coordinates of Vijay-Kumar et al. (PDB code 1UBQ).[40] Alignment tensor orientation is $\theta = 35.5 \pm 0.4°$, $\phi = 43.9 \pm 0.8°$, $\psi = 51.9 \pm 1.8°$ with $R = 0.163 \pm 0.009$ for the uncharged liquid crystal; $\theta = 34.0 \pm 0.4°$, $\phi = 32.0 \pm 0.7°$, $\psi = 20.9 \pm 0.8°$ with $R = 0.480 \pm 0.010$ for the charged liquid crystal. Error estimates were determined as described in the text. [c] Coordinates obtained after refinement of the structure factors of 1UBQ using Engh and Huber parameters.[53] [d] X-ray coordinates of Alexeev et al. (PDB code 1UBI).[54] [e] Protons were added with the program MOLMOL[41] using default parameter settings. [f] Standard error obtained by linear regression of predicted versus measured data, as in Figure 3. [g] Numbers represent one-half of the pairwise RMS difference of the D_a values obtained from 25 independent pairs of half the number of dipolar couplings for each of the five classes (see text for further explanations).

$^1D_{C'N}$ and $^2D_{C'H^N}$ couplings for Ile^{13} are relatively small (0.7 and -2 Hz, respectively), but duplicate experiments indicate they can be measured with high reproducibility (see below).

For the neutral liquid crystalline sample, a total of 333 dipolar couplings could be measured accurately (Supporting Information), without significant interference from resonance overlap (Table 1). For the sample containing charged bicelles, the degree of alignment was approximately 1.5 times larger. As a result, ^{1}H multiplet widths are increased due to larger homonuclear dipolar couplings, and a slightly smaller number (330) of dipolar couplings could be measured reliably (Supporting Information).

Fitting of Dipolar Couplings to Ubiquitin Structure. Data for the C-terminal residues 71$-$76, which have very low generalized order parameters,[48] were excluded from all analyses. The orientation and magnitudes of the D^{AB} tensors (eq 1) are then obtained from best fitting of eq 1 to the observed dipolar couplings. Variables in this search are the three Euler angles describing the orientation of the diagonalized alignment tensor (diagonalized Saupe matrix) relative to the frame of the X-ray coordinates, the rhombicity R, and for each type of dipolar coupling the magnitude of D_a^{AB}. In the first round of this fitting, values for two residues (Gln^{33} and Lys^{48}) and several dipolar interactions in other residues were identified which yielded particularly poor fits between measured and "best-fit" dipolar couplings (open circles in Figure 3). These interactions all correspond to loop regions with high temperature factors and substantial conformational variability between X-ray structures of different crystal forms.[40,49] These couplings were therefore omitted in a second round of fitting, and the results of this second round are graphically represented in Figure 3. For the neutral liquid crystal data (Figure 3), the final fit was performed with a total of 288 residual dipolar couplings (Table 1). The orientation of the alignment tensor obtained in this manner is described by the Euler rotation angles $\theta = 35.5°$, $\phi = 43.9°$, and $\psi = 51.9°$ with respect to the frame of the Brookhaven Protein Data Bank coordinates of ubiquitin.[40] The D_a values are listed in Table 1 and the rhombicity R is 0.163. Fits of comparable quality were obtained for 282 dipolar couplings measured in the sample containing positively charged bicelles (Table 1). The orientation of the alignment tensor for this second sample is quite different ($\theta = 34.0°$, $\phi = 32.0°$, and $\psi = 20.9°$) and it is also much more rhombic ($R = 0.48$) (Table 1).

Error Analysis. Before discussing the relation between the various D_a values and the corresponding bond lengths we first discuss the sources of error and uncertainty in these parameters. The standard errors of the linear regressions of predicted versus measured data (Figure 3, Table 1) indicate uncertainties in the D_a values ranging from 1.6% for $D_a^{C\alpha C'}$ to 3.3% for $D_a^{C'HN}$ in the neutral liquid crystal sample and 1.7% for $D_a^{C'N}$ to 3.0% for $D_a^{C'HN}$ in the charged liquid crystal sample. As the errors may not be randomly distributed we performed a second test as well: For each of the five classes the data were split into two nonoverlapping, equalized groups of residues. Individual fits of these groups were performed as described above. This procedure was repeated 25 times, each time generating two independent D_a values. Using only half of the available dipolar couplings is expected to increase the intrinsic uncertainty in a derived D_a value by $\sqrt{2}$. The pairwise rms difference between the 25 pairs of D_a values is $\sqrt{2}$ larger than the D_a uncertainty in a half data set, and therefore equals twice the uncertainty in the D_a of the full set. The uncertainties obtained in this manner for the D_a values are listed in Table 1 and they agree well with the results from the linear regression analysis.

On average, the random error in the D_a values is approximately 2% (Table 1). In the calculation of relative effective bond lengths (see below) this error propagates to 2% times $\sqrt{2}$ for the ratio of two D_a values. With D_a^{AB} proportionate to $\langle r_{AB}^{-3}\rangle$, this translates into a 0.95% error in the relative internuclear distance, $r_{AB}/r_{C'N}$ or $r_{AB}/r_{C\alpha C'}$. After averaging the r_{AB} values obtained using C$'$$-$N and C$^\alpha$$-C'$ as reference distances, and using the data from both the neutral and charged liquid crystal samples, the error is further reduced to about 0.6%.

The goodness of the correlations in Figure 3, and thereby the accuracy of the derived D_a values, are primarily limited by small differences between the average orientation of the internuclear vectors in the crystal structure and in solution, and by the fact that the assumption of a uniform **S** value is not

(44) Wang, Y.-X.; Marquardt, J. L.; Wingfield, P.; Stahl, S. J.; Lee-Huang, S.; Torchia, D. A.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 7385$-$7386.

(45) Pervushin, K.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12366$-$12371.

(46) Griesinger, C.; Sørensen, O. W.; Ernst, R. R. *J. Am. Chem. Soc.* **1985**, *107*, 6394$-$6396. Biamonti, C.; Rios, C. B.; Lyons, B. A.; Montelione, G. T. *Adv. Biophys. Chem.* **1994**, *4*, 51$-$120. Eberstadt, M.; Gemmecker, G.; Mierke, D. F.; Kessler, H. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 1671$-$1695.

(47) Delaglio, F.; Torchia, D. A.; Bax, A. *J. Biomol. NMR* **1991**, *1*, 439$-$446.

(48) Schneider, D. M.; Dellwo, M. J.; Wand, A. J. *Biochemistry* **1992**, *31*, 3645$-$3652. Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 12562$-$12566.

(49) Cook, W. J.; Jeffrey, L. C.; Carson, M.; Chen, Z.; Pickart, C. M. *J. Biol. Chem.* **1992**, *267*, 16467$-$16471.
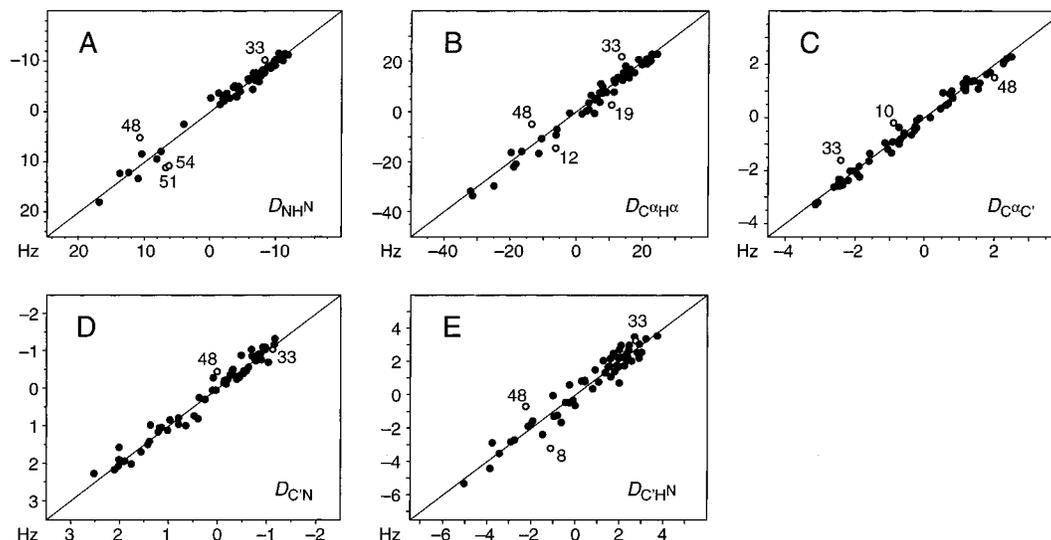
Figure 3. Plots of measured versus predicted residual dipolar couplings in ubiquitin, dissolved in the uncharged liquid crystalline phase. (A) $^1D_{NH}$, (B) $^1D_{C\alpha H\alpha}$, (C) $^1D_{C\alpha C'}$, (D) $^1D_{C'N}$, and (E) $^2D_{C'HN}$. The data represented by open circles are outliers which are due to structural differences between the solution and crystalline state, and they were not used for further analysis (see text).

strictly valid. The first includes uncertainties in the positions of the carbon and nitrogen atoms in the X-ray structure, the way the protons were added to the X-ray structure, and—last but not least—real structural differences between the protein in the crystal form and in solution. As we discuss below, the random measurement errors in the dipolar coupling are relatively small and therefore do not significantly limit the accuracy at which the various D_a values can be determined.

Reproducibility of the NMR Data. The average pairwise rmsd between $^1J_{NH} + ^1D_{NH}$ values extracted from two successive spectra recorded with the charged liquid crystal sample at 35 °C indicates random errors of 0.17, 0.75, 0.17, 0.07, and 0.67 Hz for D_{NH}, $D_{C\alpha H\alpha}$, $D_{C\alpha C'}$, $D_{C'N}$, and $D_{C'HN}$, respectively. However, the rms deviations of measured versus predicted dipolar couplings is 2.46, 5.76, 0.60, 0.31, and 1.00 Hz, respectively. The random measurement errors are even smaller in the isotropic phase which, owing to the absence of $^1H-^1H$ and other dipole−dipole splittings, offers better resolution and sensitivity. Except for the two-bond $D_{C'HN}$ interaction, this indicates that the random measurement errors contribute only a very minor fraction to the total uncertainty in r_{AB}. In the neutral liquid crystalline medium, the random measurement error is also somewhat smaller than in the charged liquid crystalline phase, but this is offset by the smaller range covered by the dipolar couplings in this medium. The possible presence of systematic errors in the measurements, which could potentially result from cross-correlation effects,[50,51] is difficult to evaluate, but such errors are expected to be very small. For example, a second measurement of $J_{CH} + D_{CH}$ splittings was made using the very different, *J*-modulated 3D CT-HSQC experiment.[52] Comparison of the two data sets yields a very small random uncertainty of 0.75 Hz for the $D_{C\alpha H\alpha}$ couplings, which cover a range of almost 120 Hz. In addition, from linear regression a slope of 1.00 is obtained and an *R* factor of 0.9992, confirming that any systematic errors are negligibly small.

Uncertainties in the X-ray Structure. When the X-ray structure of ubiquitin is calculated from the structure factors

deposited by Vijay-Kumar et al.[40] in the Protein Brookhaven Databank using Engh and Huber (E&H) parameters,[53] a structure is obtained which differs by an rmsd of 0.07 Å for the backbone atoms N, C^α, and C' of residues 1−70. The E&H refined structure gives an equally good fit to the experimental dipolar couplings. However, D_a values differ by 0.5−1.7% (Table 1). Moreover, the rmsd between the predicted dipolar couplings using the original and the E&H refined structures amounts to 50−70% of the rmsd between measured and predicted couplings using the original structure. This indicates that at least half of the observed scatter in Figure 3 results from structural uncertainties in the coordinates used. Further indication of such differences is obtained by directly comparing bond vectors of the two X-ray structures after best-fitting of the backbone atoms of residues 1−70: The corresponding N−H^N, C^α−H^α, C^α−C', C'−N, and C'−H^N vectors of residues 1−70 differ by a root-mean-square angle of 2.8°. Between the two ubiquitin 1.8 Å X-ray structures of Vijay-Kumar et al.[40] and Alexeev et al.[54] the corresponding rmsd is 3.2°. These uncertainties are expected to be even larger if the X-ray structure does not exactly represent the structure in solution, i.e., if there are real conformational differences between crystal and solution structures (see above). In addition, these comparisons do not take into account that the hydrogens are added to these structures using identical procedures. For example, it is assumed that H^N is located exactly in the N−C^α−C' plane, whereas there is evidence that the N−H^N vector may make a rms angle of up to ca. 5° with this plane.[55] Similarly, the exact location of H^α relative to the heavy atoms is not exactly known and will vary depending on the backbone and side chain torsion angles.[56] A reasonable estimate for the rms angle, $\langle\alpha\rangle$, between the internuclear vectors in the X-ray structure and the average solution orientation is about 5°.

This random difference between the X-ray structure and the average bond vector orientations in solution not only increases the uncertainty in the D_a values but also reduces the magnitude of D_a by a factor of $\langle P_2(\cos\alpha)\rangle$, where α is again the angle

(50) Tolman, J. R.; Prestegard, J. H. *J. Magn. Reson. Ser. B* **1996**, *112*, 245−252.

(51) Tjandra, N.; Bax, A. *J. Magn. Reson.* **1997**, *124*, 512−515.

(52) Ottiger, M.; Delaglio, F.; Marquardt, J. L.; Tjandra, N.; Bax, A. *J. Magn. Reson.* **1998**, *134*, 365−369.

(53) Engh, R. A.; Huber, R. *Acta Crystallogr.* **1991**, *A47*, 392−400.

(54) Alexeev, D.; Bury, S. M.; Turner, M. A.; Ogunjobi, O. M.; Muir, T. W.; Ramage, R.; Sawyer, L. *Biochem. J.* **1994**, *299*, 159−163.

(55) MacArthur, M. W.; Thornton, J. M. *J. Mol. Biol.* **1996**, *264*, 1180−1195.

(56) Karplus, P. A. *Protein Sci.* **1996**, *5*, 1406−1420.

between the solution and the model bond vectors. For small angles α, $\langle P_2(\cos\alpha)\rangle \approx P_2(\cos\langle\alpha\rangle)$. Therefore we can expect that for $\langle\alpha\rangle = 5°$ all D_a values will be reduced by about 1.2%. If we assume that $\langle\alpha\rangle$ does not differ very much among the five dipolar coupling groups, then this uniform reduction of D_a values has no significant effect on the (relative) interatomic distances, calculated below. However, if the uncertainty in the N$-$H^N bond vector orientation is larger than, for example, for the C$'-$N bond the calculated (relative) N$-$H^N bond length will be slightly overestimated.

A statistical analysis using a set of 1000 randomly distributed vectors was used to confirm the above considerations. Assuming the same D_a and R values experimentally measured for the neutral liquid crystalline ubiquitin sample, a reference set of dipolar couplings was calculated for these vectors. Subsequently, the vectors were "randomized" by addition of random errors with $\langle\alpha\rangle = 5.0°$. The reference couplings were then fitted to the "randomized" vectors. The average reduction for D_a obtained over 20 such fits is 1.2%, as predicted above. To assess the influence of the structure itself, where the bond vectors are not randomly distributed,[57] we also performed a similar analysis using the X-ray structure instead of a set of randomly distributed vectors. Based on this structure, with protons added by MOLMOL,[41] a reference set of dipolar N$-$H^N couplings was calculated for the same residues as used in the data analysis described above, and using the same alignment tensor values as obtained from the final fit of the neutral liquid crystalline ubiquitin sample. Randomly changing the orientations of the N$-$H^N bond vectors of the X-ray structure by an average of 5° and refitting the reference couplings yields an average reduction (over 100 trials) in D_a from 9.42 to 9.38 $\pm$ 0.14. This corresponds to a reduction of the average D_a of only about 0.42%. This smaller-than-expected reduction in D_a must be attributed to the fact that the distribution of N$-$H^N bond vectors is not completely random but a disproportionally high number of N$-$H^N vectors are close to perpendicular to the principal alignment axis (points in the upper right quarter of Figure 3A). On the other hand, the spread in the calculated D_a values is as high as 1.5%, again indicating that a large part of the remaining uncertainty in the D_a values listed in Table 1 is caused by uncertainties in the atom positions in the crystal structure.

Discussion

Relative Effective Bond Lengths. Although a rigorous quantitative analysis of the effect of internal motions is not possible on the basis of the limited data available from the above experiments, a number of interesting semiquantitative conclusions can be drawn, as discussed below. Values for D_a are proportional to the inverse ratio of the cubed bond lengths, scaled by the corresponding gyromagnetic ratios and the order parameter, S (eq 2). If S were accurately known, this would permit straightforward calculation of the ratios of the internuclear distances for all types of measured dipolar interactions. However, there are indications from NMR relaxation and molecular dynamics studies that the average S values are not the same for different types of internuclear vectors.[58] Analogously, we find that the relative magnitudes of different types of A$-$B dipolar couplings, D_a^{AB}, do not simply scale with $\gamma_A\gamma_B\langle r_{AB}^{-3}\rangle$. However, one can define an effective bond length for the interaction between nuclei A and B according to

$$\langle(r_{AB}^{\text{eff}})^{-3}\rangle = (D_a^{AB}/D_a^{C'N})(\gamma_C\gamma_N/\gamma_A\gamma_B)\langle(r_{C'N}^{\text{eff}})^{-3}\rangle \quad (3a)$$

Alternatively, one can assign different order parameters to the different types of interactions:

$$\mathbf{S}^{AB} = \mathbf{S}^{C'N}\langle(r_{AB}^{\text{eff}})^{-3}\rangle/\langle(r_{AB})^{-3}\rangle \quad (3b)$$

From a statistical analysis of atomic resolution X-ray structures of small compounds, Engh and Huber[53] derived $\langle r_{C'N}\rangle = 1.329$ Å and $\langle r_{C\alpha C'}\rangle = 1.525$ Å. These distances represent averages over a large set of very high resolution crystal structures, and are now widely accepted and used in default parameter sets for X-ray and NMR macromolecular structure determination. For $r_{C'N}$ and $r_{C\alpha C'}$ the averaging over vibrational stretching results in a small increase in $\langle r^{-3}\rangle$ relative to $\langle r\rangle^{-3}$ if the potential is harmonic, but this effect is largely offset by a decrease in $\langle r^{-3}\rangle$ relative to $\langle r\rangle^{-3}$ as a result of anharmonicity of the bond stretching. In either case, for pairs of heavy atoms these effects are extremely small and for C$-$C and C$-$N bonds we may safely assume $\langle r^{-3}\rangle = \langle r\rangle^{-3}$.

Using the $\langle r_{C'N}\rangle = 1.329$ Å reference value, eq 3a yields $\langle r_{C\alpha C'}\rangle = 1.526 \pm 0.014$ Å for the data set acquired in the neutral liquid crystalline phase, and $\langle r_{C\alpha C'}\rangle = 1.527 \pm 0.014$ Å for the data set acquired in the charged liquid crystalline phase. The close similarity between these $\langle r_{C\alpha C'}\rangle$ values and the value from the Engh and Huber parameter set (1.525 Å)[53] indicates that $\mathbf{S}^{C'N}$ equals $\mathbf{S}^{C\alpha C'}$, i.e., that the amplitude of the librations are essentially identical for the C$'-$N and C$^\alpha-$C$'$ bonds.

Relative effective internuclear distances for r_{NH}, $r_{C\alpha H\alpha}$, and $r_{C'HN}$, obtained using eq 3a with $\langle r_{C'N}\rangle = 1.329$ Å and $\langle r_{C\alpha C'}\rangle = 1.525$ Å as reference values, are listed in Table 2. On average, $r_{NH}^{\text{eff}} = 1.041 \pm 0.006$ Å, $r_{C\alpha H\alpha}^{\text{eff}} = 1.117 \pm 0.007$ Å, and $r_{C'HN}^{\text{eff}} = 2.055 \pm 0.016$ Å. Alternatively, using reference distances of 1.02 Å for r_{NH},[20] eq 3b indicates that, on average, $\mathbf{S}^{NH} = (0.94 \pm 0.02)\mathbf{S}^{C'N}$. Similarly, if a 1.095 Å reference distance[20] is used for $\langle r_{CH}^{-3}\rangle^{1/3}$, eq 3b yields $\mathbf{S}^{C\alpha H\alpha} = (0.94 \pm 0.02)\mathbf{S}^{C'N}$. A direct comparison of N$-$H^N and C$^\alpha-$H$^\alpha$ order parameters gives $\mathbf{S}^{C\alpha H\alpha} = (1.00 \pm 0.03)\mathbf{S}^{NH}$, i.e., the average S^2 values differ by less than 6%. This indicates that N$-$H^N and C$^\alpha-$H$^\alpha$ vectors undergo a similar degree of increased "wagging" motions relative to the backbone C$'-$N and C$^\alpha-$C$'$ bonds. However, the 6% uncertainty in the ratio of the average N$-$H^N and C$^\alpha-$H$^\alpha$ S^2 values is rather large, and therefore does not contradict experimental and molecular dynamics results which suggests that, on average, C$^\alpha-$H$^\alpha$ S^2 values are somewhat higher than N$-$H^N S^2 values.[58]

Although the uncertainty in $r_{C'HN}^{\text{eff}}$ is considerably larger than for the other internuclear distances, it is interesting to note that standard covalent geometry with a C$'-$N$-$H^N angle of 120°, $r_{NH} = 1.02$ Å, and $r_{C'N} = 1.329$ Å yields $r_{C'HN} = 2.041$ Å. Equation 3b then suggests $\mathbf{S}^{C'HN} = (0.98 \pm 0.04)\mathbf{S}^{C'N}$. Thus, $\mathbf{S}^{C'HN}$ is larger than $\mathbf{S}^{HN}$, which presumably reflects the fact that wagging motions of the N$-$H bond vector have less of an effect on the orientation of the longer C$'-$H^N bond vector. It must be pointed out, however, that this derived $\mathbf{S}^{C'HN}/\mathbf{S}^{C'N}$ ratio is quite sensitive to the C$'-$N$-$H^N angle which is not known

(57) Lee, L. K.; Rance, M.; Chazin, W. J.; Palmer, A. G. *J. Biomol. NMR* **1997**, *9*, 287$-$298.

(58) Palmer, A. G.; Case, D. A. *J. Am. Chem. Soc.* **1992**, *114*, 9059$-$9067. Brüschweiler, R.; Wright, P. E. *J. Am. Chem. Soc.* **1994**, *116*, 8426$-$8427. Bremi, T. Ph.D. Thesis, ETH Zürich, 1997. Fadel A. R.; Jin, D. Q.; Montelione, G. T.; Levy, R. M. *J. Biomol. NMR* **1995**, *6*, 221$-$226. LeMaster, D. M., Kushlan, D. M. *J. Am. Chem. Soc.* **1996**, *118*, 9255$-$9264. Nicholson, L. K.; Kay, L. E.; Torchia, D. A. In *NMR Spectroscopy and its Application to Biomedical Research*; Sarkar, S. S., Ed.; Elsevier: Amsterdam; pp 241$-$276.

Table 2. Relative Effective Bond Lengths in Ubiquitin Determined in Two Different Liquid Crystalline Phases[a]

	r_{NHN} (Å)	$r_{C\alpha H\alpha}$ (Å)	$r_{C\alpha C'}$ (Å)	$r_{C'N}$ (Å)	$r_{C'HN}$ (Å)
uncharged LC	1.041	1.118	1.526	**1.329**	2.067
	1.042	1.119	**1.525**	1.329	2.069
charged LC	1.042	1.115	1.527	**1.329**	2.043
	1.040	1.113	**1.525**	1.327	2.040

[a] The individual values are calculated using the boldface values as reference distances which were taken from Engh and Huber.[53] See section on error analysis for uncertainties in the distances listed.

accurately. A C′−N−HN angle of 119° yields $\mathbf{S}^{C'HN} = 0.96\mathbf{S}^{C'N}$; a C′−N−H^N angle of 121° yields $\mathbf{S}^{C'HN} = 0.99\mathbf{S}^{C'N}$.

Concluding Remarks

All measurements of dipolar couplings assumed that the change in splitting upon aligning the protein is entirely caused by the dipolar coupling and that anisotropy of the *J* coupling itself is negligible. Recent ab initio calculations and experimental studies of formamide indicate that the anisotropy of $^1J_{C'N}$ is relatively small and nearly axially symmetric, with its unique axis parallel to the C′−N bond, and a magnitude of −14 Hz.[59] This is about 200 times smaller than the dipolar coupling and therefore has an effect which, after scaling by the alignment tensor, is about four times smaller than the random error in the measured magnitude of the $^{13}C'-^{15}N$ dipolar coupling. Anisotropy of the one-bond $^{13}C-^{13}C$ *J* coupling in benzene is of comparable absolute magnitude to that found for $^1J_{C'N}$ in formamide,[60] but as D_{CC} is about 1.65 times larger than $D_{C'N}$, the effect of J_{CC} anisotropy is even smaller for $^{13}C-^{13}C$ than for $^{13}C'-^{15}N$. *J* couplings involving ^1H are dominated by the Fermi-contact contribution and anisotropy is entirely negligible relative to the large dipolar couplings.

The data shown for ubiquitin demonstrate that dipolar couplings can be measured with high relative accuracy when the protein is dissolved in a dilute liquid crystalline phase. Comparison of the magnitudes of the different types of dipolar interactions yields ratios for the effective internuclear distances. Whereas in other methods, such as rotational spectroscopy, gas phase electron diffraction, and X-ray or neutron crystallography, bond stretching vibrations pose the main obstacle when interpreting the data in terms of bond length, for liquid crystal NMR the wagging motions are considerably more important.[2,3]

Relative effective bond lengths for r_{NH}^{eff} and $r_{C\alpha H\alpha}^{eff}$, obtained using eq 3a with $\langle r_{C'N} \rangle = 1.329$ Å and $\langle r_{C\alpha C'} \rangle = 1.525$ Å as reference value, yielded $r_{NH}^{eff} = 1.041 \pm 0.006$ Å and $r_{C\alpha H\alpha}^{eff}$

(59) Vaara, J.; Kaski, J.; Jokisaari, J.; Diehl, P. *J. Phys. Chem. A* **1997**, *101*, 5069−5081.
(60) Kaski, J.; Vaara, J.; Jokisaari, J. *J. Am. Chem. Soc.* **1996**, *118*, 8879−8886.
(61) Kay, L. E.; Keifer, P.; Saarinen, T. *J. Am. Chem. Soc.* **1992**, *114*, 10663−10665.
(62) Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *2*, 661−665.

= 1.117 ± 0.007 Å. The dipolar couplings for these latter two interactions are, on average, 6% smaller than expected for a static structure, resulting in a 2% increase in the effective bond length over the average bond length, derived from non-NMR techniques. This corresponds to $\mathbf{S}^{C\alpha H\alpha}/\mathbf{S}^{C'N} = \mathbf{S}^{NH}/\mathbf{S}^{C'N} = 0.94$. Assuming a model where the N−H^N or C$^\alpha$−H$^\alpha$ bond vector freely diffuses in a cone, superimposed on the fluctuations of positions of the heavy backbone atoms, this corresponds to a relatively large semi-angle β of 16°.[42] As discussed in the section on error analysis, the fact that the random error in the coordinates of the protons (which are added to the crystal structure assuming idealized geometry) is presumably larger than the uncertainties in the carbon and nitrogen atom positions may to some extent be responsible for the small value of $\mathbf{S}^{NH}/\mathbf{S}^{C'N}$. However, the presence of extensive wagging motions of the N−H^N bond vectors is compatible with solid-state NMR experiments which yield dipolar couplings corresponding to ca. 1.06 Å,[22,23] or $\mathbf{S}^{NH} = (1.02/1.06)^3 = 0.89$.

The accuracy at which the relative magnitudes of the various types of D_a values can be determined is limited by small uncertainties in the crystal structure coordinates of ubiquitin and small true differences between the structures in solution and in the crystalline state. However, the accuracy of ±2% at which we were able to determine these relative magnitudes is more than sufficient to permit simultaneous use of multiple types of dipolar couplings in protein structure determination by NMR. Our data also suggest that the ultimate accuracy at which macromolecular structures can be determined by NMR when including dipolar couplings measured in the liquid crystalline state generally will be limited by the effect of internal motions on the measured dipolar couplings, and not by uncertainties in the measurements themselves. Alternatively, if the structure is accurately known on the basis of other information, the dipolar couplings contain unique information on both the amplitude and direction of internal motions.

Supporting Information Available: Two tables containing the dipolar couplings measured for human ubiquitin in two different liquid crystalline phases; one figure with plots of measured versus predicted residual dipolar couplings in the charged liquid crystalline sample, analogous to Figure 3 (7 pages, print/PDF). See any current masthead page for ordering information and Web access instructions.

JA9826791