J|A|C|S ARTICLES

Published on Web 11/05/2009

Exact Distances and Internal Dynamics of Perdeuterated Ubiquitin from NOE Buildups

Beat Vögeli,[†] Takuya F. Segawa,[†,‡] Dominik Leitz,[†] Alexander Sobol,[†] Alexandra Choutko,[†] Daniel Trzesniak,[†] Wilfred van Gunsteren,[†] and Roland Riek*[,†,§]

Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH-Hönggerberg, CH-8093 Zürich, Switzerland, and The Salk Institute, 10010 North Torrey Pines Road, La Jolla, California 92037

Received June 30, 2009; E-mail: roland.riek@phys.chem.ethz.ch

Abstract: It is proposed to convert nuclear Overhauser effects (NOEs) into relatively precise distances for detailed structural studies of proteins. To this purpose, it is demonstrated that the measurement of NOE buildups between amide protons in perdeuterated human ubiquitin using a designed ^{15}N-resolved HMQC-NOESY experiment enables the determination of $^{1}H^{N}-^{1}H^{N}$ distances up to 5 Å with high accuracy and precision. These NOE-derived distances have an experimental random error of ~0.07 Å, which is smaller than the pairwise rmsd (root-mean-square deviation) of 0.24 Å obtained with corresponding distances extracted from either an NMR or an X-ray structure (pdb codes: 1D3Z and 1UBQ), and also smaller than the pairwise rmsd between distances from X-ray and NMR structures (0.15 Å). Because the NOE contains both structural and dynamical information, a comparison between the 3D structures and NOE-derived distances may also give insights into through-space dynamics. It appears that the extraction of motional information from NOEs by comparison to the X-ray structure or the NMR structure is challenging because the motion may be masked by the quality of the structures. Nonetheless, a detailed analysis thereof suggests motions between β-strands and large complex motions in the α-helix of ubiquitin. The NOE-derived motions are, however, of smaller amplitude and possibly of a different character than those present in a 20 ns molecular dynamic simulation of ubiquitin in water using the GROMOS force field. Furthermore, a recently published set of structures representing the conformational distribution over time scales up to milliseconds (pdb: 2K39) does not satisfy the NOEs better than the single X-ray structure. Hence, the measurement of possibly thousands of exact NOEs throughout the protein may serve as an excellent probe toward a correct representation of both structure and dynamics of proteins.

Introduction

Nuclear Overhauser effects (NOEs) between hydrogens are the most prominent experimental data in solution-state NMR structure determination of biomolecules.[1−3] Since the NOE cross-relaxation rate is proportional to the inverse sixth power of the distance between the two dipolar interacting spins, the power of proton−proton NOEs lies in their ability to provide a large amount of through-space distance constraints. However, as opposed to studies with small molecules, in a routine protein structure determination upper limit distance restraints obtained from NOEs are used rather than exact average values.[1] This practice dates back to the 1980s when it proved difficult to convert NOEs into exact distances.[4] At that time, the sensitivity of spectrometers required mixing times long enough to render spin diffusion very effective. Furthermore, only the 2D version of the NOESY[4] was used, yielding severe peak overlap. Here, we revisit the use of buildup-derived NOE rates in protein structure determination and propose to translate the NOE into exact distance constraints. We show in the case of the model protein ubiquitin that the combination of 3D ^{15}N-resolved ^{1}H,^{1}H-NOESY experiments,[5] ^{2}H,^{13}C,^{15}N-labeling and more sensitive NMR instruments allows the measurement of NOE buildups in the near-linear regime resulting in a very exact determination of average distances.

A second focus of this work is the extraction of motional information from homonuclear proton−proton NOEs. In routine approaches information on motion on time scales up to nanoseconds (fast motion) is commonly obtained from measurements of the backbone amide transverse and longitudinal relaxation times and heteronuclear ^{15}N{^{1}H} NOEs.[6−8] For the

[†] Swiss Federal Institute of Technology.
[‡] Present address: Laboratory of Physical Chemistry, EPFL, Lausanne.
[§] The Salk Institute.

(1) Wüthrich, K. *NMR of Proteins and Nucleic Acid*; Wiley: New York, 1986.
(2) Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*; Wiley: New York, 2000.
(3) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G.; Rance, M.; Skleton, N. J. *Protein NMR Spectroscopy: Principles and Pratice*; Academic Press: San Diego, 2007.
(4) Kumar, A.; Wagner, G.; Ernst, R. R.; Wüthrich, K. *J. Am. Chem. Soc.* **1981**, *103*, 3654–3658.
(5) Fesik, S. W.; Zuiderweg, E. R. P. *J. Magn. Reson.* **1998**, *78*, 588–593.
(6) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546–4559.

 © 2009 American Chemical Society

characterization of slower dynamics, two types of experiments have recently gained attention. Relaxation dispersion NMR experiments accurately define the time scale on which very slow (i.e., ms-s range) internal motions take place.[9] The measurement of residual dipolar couplings (RDCs) samples bond orientations over a time scale up to milliseconds.[10,11] Here, we explore the potential of NOEs to provide information on fast (faster than the rotational correlation time of the biomolecule) and slow (slower than the rotational correlation time of the molecule) motion, since the NOE-derived distance is a time-averaged parameter covering both fast and slow motions. Due to the abundance of protons in a protein this may lead to a rather complete map of the internal motion of a biomolecule. However, the dependence of NOEs on dynamics is generally not trivial[12] and a simple way to carry out an analysis is to compare experimental data to predictions of cross-relaxation rates based on structure ensembles inherently representing dynamics.[13−17] In this respect, ubiquitin is an attractive model system because recently a structure set has been derived from extensive collections of RDCs which are supposed to represent motion on all time scales up to milliseconds.[18] In addition, a 20 ns molecular dynamics (MD) simulation of ubiquitin is available that should reflect fast motions. As shall be seen, both ensemble representations fail to satisfy the derived NOE distances better than the single X-ray or NMR structures.

Theory

Extraction of Cross-Relaxation Rates. A NOESY spectrum can be described by the longitudinal relaxation properties of the spin system. Without loss of generality, a system of three dipolar coupled $^1/_2$ spins *I*, *K*, and *S* is assumed. This leads to the following Solomon equation:[19]

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\Delta I_z(t)\\ \Delta K_z(t)\\ \Delta S_z(t)\end{bmatrix} = -\begin{bmatrix}\rho_I & \sigma_{IK} & \sigma_{IS}\\ \sigma_{IK} & \rho_K & \sigma_{KS}\\ \sigma_{IS} & \sigma_{KS} & \rho_S\end{bmatrix}\begin{bmatrix}\Delta I_z(0)\\ \Delta K_z(0)\\ \Delta S_z(0)\end{bmatrix} \quad (1)$$

where ρ_X is the autorelaxation rate of spin *X* and σ_{XY} the cross-relaxation rate between spins *X* and *Y*. If *K* is far apart from *I* and *S*, $\sigma_{IK} = \sigma_{KS} = 0$. If the initial magnetization is on spin *I* the exact analytical solution for the remaining 2-spin system *IS* is

$$\frac{\Delta I_z(t)}{\Delta I_z(0)} = \frac{1}{2}\left[\left(1 - \frac{\rho_I - \rho_S}{\lambda_+ - \lambda_-}\right)e^{-\lambda_- t} - \left(1 + \frac{\rho_I - \rho_S}{\lambda_+ - \lambda_-}\right)e^{-\lambda_+ t}\right] \quad (2.1)$$

$$\frac{\Delta S_z(t)}{\Delta I_z(0)} = -\frac{\sigma_{IS}}{(\lambda_+ - \lambda_-)}[e^{-\lambda_- t} - e^{-\lambda_+ t}] \quad (2.2)$$

with

$$\lambda_\pm = \frac{(\rho_I + \rho_S)}{2} \pm \sqrt{\left(\frac{\rho_I - \rho_S}{2}\right)^2 + {\sigma_{IS}}^2} \quad (2.3)$$

It is now shown that in most cases in a perdeuterated protein eqs 2.1−2.3 can also be used to obtain σ_{IS} for a three-spin system with σ_{IK},$\sigma_{KS} \neq 0$. For that purpose, the solution of eq 1 is expanded in a Taylor series to second order in *t*:

$$\frac{\Delta I_z(t)}{\Delta I_z(0)} = 1 - \rho_I t + \frac{1}{2}({\rho_I}^2 + {\sigma_{IS}}^2 + {\sigma_{IK}}^2)t^2 \quad (3.1)$$

$$\frac{\Delta S_z(t)}{\Delta I_z(0)} = -\sigma_{IS} t + \frac{1}{2}[(\rho_I + \rho_S)\sigma_{IS} + \sigma_{IK}\sigma_{KS}]t^2 \quad (3.2)$$

In both equations, the last terms containing σ_{IK} or σ_{KS} are the modifications to the Taylor expansion of eqs 2.1 and 2.2. In eq 3.1 this term stands for an additional loss of *I* magnetization to spin *K*, and in 3.2 it is the spin diffusion term constituting an indirect magnetization flow from *I* to *S* via *K*. Since ρ_X is typically 2 s^{-1} and the largest σ_{XY} are ca. 1.5 s^{-1} (vide infra) and fits are carried out up to $t = 90$ ms eqs 2.1 and 2.2 are good approximations for most NOEs in the amide proton network of a perdeuterated protein. Simulations show that the approximation only breaks down for amide protons in nonconsecutive residues in the α helix and loops. The impact of spin diffusion pathways on the NOE due to incomplete deuteration must also be taken into account (see also Supporting Information). Whereas the effect of spin diffusion via residual aliphatic protons in the helix is negligible, it is approximately linearly dependent on the protonation level for consecutive spin pairs in the β sheet and nonuniform for spin pairs between the β strands. The deuteration level of the sample used in this study is ca. 99% for H^α and 95% for other carbon-bound protons. Inclusion of all these additional pathways leads to an overestimation of NOE rates σ_{XY} of 10−50% in nonconsecutive residues in the α helix and loops, and an averaged overestimation of only 9% in the other spin pairs. This effect is reduced by 3% by the H/D equilibrium of 97%/3%. Representative NOE buildup curves between amides located in secondary elements are shown in the Supporting Information.

Conversion of NOE Rates into Distances. The homonuclear cross-relaxation rate is given by:[2,3]

$$\sigma_{XY} = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma^4 \hbar^2}{40\pi^2} \frac{1}{(r_{XY}^{\mathrm{rigid}})^6}[J(0) - 6J(2\omega)] \quad (4)$$

where γ is the gyromagnetic ratio of nucleus *X*, ω is the spectral frequency of the nuclei, μ_0 is the permeability in vacuum, and $\hbar$ denotes Planck's constant. r_{XY}^{rigid} is the internuclear distance in a hypothetically rigid structure, which is obtained here from a single conformer representation determined by NMR or X-ray crystallography. A simple expression for the spectral density *J* is obtained under the assumption of isotropic molecular tumbling,[12]

(7) Kay, L. E.; Torchia, D. A.; Bax, A. *Biochemistry* **1989**, *28*, 8972–8979.
(8) Korzhnev, D. M.; Billeter, M.; Arseniev, A. S.; Orekhov, V. Y. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *38*, 197–266.
(9) Mittermaier, A.; Kay, L. E. *Science* **2006**, *312*, 224–228.
(10) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. *Nat. Struct. Biol.* **1997**, *4*, 292–297.
(11) Meiler, J.; Prompers, J. J.; Peti, W.; Griesinger, C.; Brüschweiler, R. *J. Am. Chem. Soc.* **2001**, *123*, 6098–6107.
(12) Brüschweiler, R.; Roux, B.; Blackledge, M.; Griesinger, C.; Karplus, M.; Ernst, R. R. *J. Am. Chem. Soc.* **1992**, *114*, 2289–2302.
(13) Olejniczak, E. T.; Dobson, C. M.; Karplus, M.; Levy, R. M. *J. Am. Chem. Soc.* **1984**, *106*, 1923–1930.
(14) Abseher, R.; Lüdemann, S.; Schreiber, H.; Steinhauser, O. *J. Mol. Biol.* **1995**, *249*, 604–624.
(15) Lienin, S. F.; Bremi, T.; Brutscher, B.; Brüschweiler, R.; Ernst, R. R. *J. Am. Chem. Soc.* **1998**, *120*, 9870–9879.
(16) Peter, C.; Daura, X.; Van Gunsteren, W. F. *J. Biomol. NMR* **2001**, *20*, 297–310.
(17) Feenstra, K. A.; Peter, C.; Scheek, R. M.; Van Gunsteren, W. F.; Mark, A. E. *J. Biomol. NMR* **2002**, *23*, 181–194.
(18) Lange, O. F.; Lakomek, N. A.; Fares, C.; Schröder, G. F.; Walter, K. F. A.; Becker, S.; Meiler, J.; Grubmüller, H.; Griesinger, C.; De Groot, B. L. *Science* **2008**, *320*, 1471–1475.
(19) Solomon, I. *Phys. Rev.* **1955**, *2*, 559–565.

$$J(\omega) = S_{XY}^{\text{fast2}} \frac{\tau_c}{1 + (\tau_c\omega)^2} + \left((r_{XY}^{\text{rigid}})^6 \left\langle \frac{1}{r_{XY}^6} \right\rangle - S_{XY}^{\text{fast2}} \right) \frac{\tau_{\text{tot}}}{1 + (\tau_{\text{tot}}\omega)^2} \tag{5}$$

with

$$\frac{1}{\tau_{\text{tot}}} = \frac{1}{\tau_c} + \frac{1}{\tau_{\text{int}}} \tag{6}$$

where τ_c is the rotational correlation time of the molecule and τ_{int} is the correlation time for internal motion. The angled brackets denote a Boltzmann ensemble average. S_{XY}^{fast2} is an order parameter for fast internal motion,

$$S_{XY}^{\text{fast2}} \equiv (r_{XY}^{\text{rigid}})^6 \frac{4\pi}{5} \sum_{q=-2}^{2} \left\langle \frac{Y_{2q}(\theta_{XY}^{\text{mol}}, \varphi_{XY}^{\text{mol}})}{(r_{XY})^3} \right\rangle^2 \tag{7}$$

An experimentally accessible order parameter necessarily covers all time scales of the measurements and may be defined by the true cross-relaxation rate normalized to the one expected for a rigid molecule,

$$S_{XY}{}^2 \equiv \frac{\sigma_{XY}^{\text{exp}}}{\sigma_{XY}^{\text{rigid}}} \tag{8}$$

This order parameter is unitless. For macromolecules at high magnetic fields J sampled at frequencies other than zero can be neglected and $S_{XY}{}^2$ can be rewritten as

$$S_{XY}{}^2 = S_{XY}^{\text{fast2}} + \left((r_{XY}^{\text{rigid}})^6 \left\langle \frac{1}{r_{XY}^6} \right\rangle - S_{XY}^{\text{fast2}} \right) \frac{1}{1 + \tau_c/\tau_{\text{int}}} \tag{9}$$

For internal motion much faster than nanoseconds ($\tau_{\text{int}} \ll \tau_c$), $S_{XY}{}^2$ reduces to the order parameter of fast motion as defined in eq 7

$$S_{XY}{}^2 = S_{XY}^{\text{fast2}} \tag{10.1}$$

S_{XY}^{fast2} may be approximately decomposed into a radial ($S_{XY}^{\text{rad2}} > 1$) and an angular component ($S_{XY}^{\text{ang2}} < 1$).[12] Hence, S_{XY}^{fast2} can be smaller or larger than 1 depending on the exact nature of the fast internal motions. For motion much slower than the molecular tumbling ($\tau_{\text{int}} \gg \tau_c$) $S_{XY}{}^2$ becomes independent of angular coordinates,

$$S_{XY}{}^2 = (r^{\text{rigid6}}) \left\langle \frac{1}{r^6} \right\rangle \tag{10.2}$$

The simplest and most common way to extract distances from the measured cross-relaxation rate is to use eq 4 under the assumption of a rigid molecule. Motional effects are absorbed into the distance which must be replaced by an effective distance r_{XY}^{eff}:

$$\sigma_{XY} = \left(\frac{\mu_0}{4\pi} \right)^2 \frac{\gamma^4 \hbar^2}{40\pi^2} \frac{1}{(r_{XY}^{\text{eff}})^6} J^{\text{rigid}}(0) \tag{11}$$

The relationship between the distances in a rigid molecule and the effective distances can be expressed with the order parameter:

$$r^{\text{eff}} = \frac{r^{\text{rigid}}}{S_{XY}{}^{1/3}} \tag{12}$$

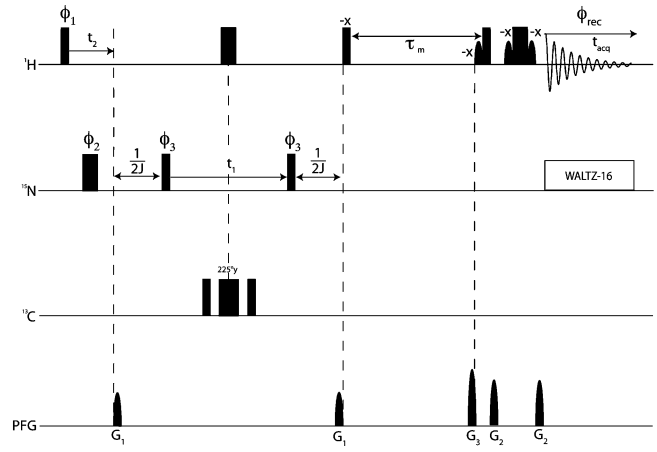

Figure 1. Experimental scheme for the ^{15}N-resolved HMQC-NOESY experiment. The radio frequency pulses on ^{1}H, ^{15}N, and ^{13}C were applied at 4.7, 116, and 110 ppm, respectively. The narrow and wide black bars indicate nonselective 90° and 180° pulses. The curved shapes on the ^{1}H line represent Gaussian-shaped selective 90° pulses truncated at the 5% level on the water resonance with a length of 1.0 ms. Pulsed field gradients along the z-axis are shown on the line marked PFG. All gradients have a smoothed square shape with a length of 1.0 ms. The individual gradient strengths are 44% for G_1, 70% for G_2, and 80% for G_3 with a maximal strength of 53 G/cm. Quadrature detection is obtained by phases ϕ_1 in the t_2 (^{1}H) dimension, and by ϕ_3 in the t_1 (^{15}N) dimension, respectively, both cycled according to the States-TPPI method.[44] Five-hundred twelve complex points were recorded for the direct (^{1}H), 256 and 22 for the indirect t_2 (^{1}H) and t_1 (^{15}N) dimensions with $t_{1,\text{max}} = 24$ ms and $t_{2,\text{max}} = 22.4$ ms, respectively. $1/2J = 5.4$ ms. ^{15}N was decoupled during acquisition using the WALTZ-16 sequence.[45] All radio frequency pulses are applied with phase x, unless a different phase is indicated in the figure: $\phi_1 = \{x, x, -x, -x\}$, $\phi_2 = \{x, -x\}$, $\phi_{\text{rec}} = \{-x, x, x, -x\}$.

Note that this holds true for motions on all time scales as well as for cases where $J(2\omega)$ cannot be neglected.

Experimental section

Sample Expression and Purification and Preparation. Ubiquitin with the human sequence was expressed recombinantly in *E. coli* in the triple-labeled medium from Silantes *E. coli*-OD2 CDN (^{2}H > 95%, ^{13}C, ^{15}N) and purified in ^1H_2O, having a mass of 9500.2 Da. This results in a nearly perdeuterated sample with exception of the exchangeable backbone amide protons and amine protons in the side chains. The sample was measured in ^1H_2O with 3% of ^2H_2O at a concentration of ~4.3 mmol/L and pH 5.8 at 284 K having a rotational correlation time of 9.3 ns.

NMR Spectroscopy. All experiments were performed on a Bruker 700 MHz spectrometer equipped with a triple resonance cryoprobe. The temperature was 284 K. All spectra were processed with the program PROSA[20] and analyzed with the program XEASY.[21] Peak intensities were determined by taking the maximal peak height rather than integrating volumes. This method is advantageous in crowded regions with partial overlap such as the NOESY diagonal. Because the sample is perdeuterated and because of the limited spectral resolution along the two indirect dimensions the line widths of the diagonal and cross peaks are rather uniform. Potential derived artifacts from this procedure would have an impact on the determined experimental random error (vide infra) and hence would also be reflected by it.

The resonances were assigned using a TROSY-HNCA and a 3D ^{15}N-resolved HMQC-NOESY experiment.[3] The 3D ^{15}N-resolved HMQC-NOESY experiment (Figure 1) was also used to measure

(20) Güntert, P.; Dötsch, V.; Wider, G.; Wüthrich, K. *J. Biomol. NMR* **1992**, *2*, 619–629.
(21) Bartels, C.; Xia, T. H.; Billeter, M.; Güntert, P.; Wüthrich, K. *J. Biomol. NMR* **1995**, *6*, 1–10.

NOE buildup rates. Spectra were acquired with the mixing times $\tau_m = 0.03$, 0.06, 0.09, and 0.20 s.

Placing the NOESY element after the HMQC element is advantageous for the extraction of the cross-relaxation rates. Relaxation during the HMQC is identical for the diagonal and every of its NOESY cross peaks since they share the same magnetization pathway. The intensity of detected ^1H_j magnetization I_{ij}^{det} originating from the initial ^1H_i magnetization with intensity I_i^{init} can be expressed as follows,

$$I_{ij}^{det}(\tau_m) = \alpha_i^{rec} \cdot I_i^{init} \cdot T_{ii}^{HMQC} \cdot T_{ij}^{NOESY}(\tau_m) \cdot T_{jj}^{WG}, \alpha_i^{rec} \leq 1 \quad (13)$$

α_i^{rec} accounts for the part of the magnetization that has recovered during the interscan delay. T_{ii}^{HMQC} denotes the loss of magnetization during the HMQC element, T_{jj}^{WG}describes the loss of magnetization during the Watergate element[22] before acquisition, and T_{ij}^{NOESY} is described by eq 2.1 if $i \neq j$ and by eq 2.2 if $i = j$, respectively. Slight ^1H_j dependence during T_{jj}^{WG} can be neglected by the assumption that it is identical for all j. In case this assumption is invalid it would be reflected in the experimental error (vide infra). Note, if the peak intensity is determined by the maximum height rather than volume measurements this procedure may translate into an additional correction term that is however absorbed by T_{jj}^{WG}. The term of interest is T_{ij}^{NOESY} which can now be extracted by the normalization of the cross-peak intensities ($i \neq j$) by the diagonal peak intensity ($i = j$) at $\tau_m = 0$:

$$\frac{I_{ij}^{det}(\tau_m)}{I_{ii}^{det}(0)} = T_{ij}^{NOESY}(\tau_m) \quad (14)$$

I_{ii}^{det} (0) must be either measured or is back-predicted. Note, this approach allows the interscan delay to be reduced without compromising on the extraction of the NOE and hence experimental time can be saved. This finding has been confirmed experimentally by a comparison of the NOE relaxation rates extracted from ^{15}N-resolved HMQC-NOESY experiments with an interscan delays of 0.5 and 2.0 s (data not shown).

T_1 and T_2 relaxation times of the backbone ^{15}N nuclear spins in ubiquitin were measured using standard pulse sequences.[3] The global correlation time τ_c was calculated from the ratio T_2/T_1[7] using the program DASHA[23] under assumption of isotropic overall tumbling. This simplification is justified because the diffusion anisotropy is $\sim$1.17[24] and the expected experimental error is maximally 5% for an individual rate.

Cross-Relaxation Rate Fitting. Using the relationship $T_{ii}^{NOESY}(\tau_m) = (\Delta I_z(\tau_m))/(\Delta I_z(0))$ the buildups of the diagonal peaks were fitted with a maximal τ_m of 90 ms. For convenience eqs 2.1 and 3.1, respectively, are approximated with the monoexponential function $e^{-\rho_i t}$ which is independent of any σ_{ij}. ρ_i, ρ_j, and $\Delta I_z(0)$ were subsequently used as fixed inputs for the fits of the cross-peak buildups using the relationship $T_{ij}^{NOESY}(\tau_m) = (\Delta S_z(\tau_m))/(\Delta I_z(0))$ and eq 2.2. The only fitting parameter then was σ_{ij}. Buildups that did not follow the expected function were omitted. Usually this is observed for weak peaks where the spectral noise becomes dominant. The experimental errors were determined by using the symmetry relationship of the ^{15}N-resolved HMQC-NOESY experiment which enables the determination of both NOEs σ_{ij} and σ_{ji} resulting in an error $\Delta\sigma_{ij} = (\sigma_{ij} - \sigma_{ji})/2$. In addition, the cross-relaxation rates are underestimated by 3% due to the H/D exchange equilibrium with the water protons/deuterons.

Cross-Relaxation Rate Prediction from a Set of Structures. The cross-relaxation rate for the RDC-derived set of structures and the MD trajectory is calculated assuming either slow intramolecular motion as

$$\sigma_{XY} = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma^4 h^2}{40\pi^2} \left(\sum_{a=1}^{N} \frac{1}{r_{XY,a}^6}\right) \tau_c \quad (15)$$

or assuming only fast intramolecular motion as

$$\sigma_{XY} = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma^4 h^2}{40\pi^2} \left(\sum_{a=1}^{N} \frac{1}{r_{XY,a}^3} \sum_{b=1}^{N} \frac{P_2(\cos \vartheta_{int,ab}^M)}{r_{XY,b}^3}\right) \tau_c \quad (16)$$

where a and b count the N structures in the set, P_2 is the Legendre polynomial of second order and $\vartheta_{int,ab}^M$ denotes the projection angle between the two internuclear vectors. For derivations see the Supporting Information.

The RDC-derived structure set is taken from the pdb deposition 2k39 containing 116 conformers.[18] In addition, the NOEs were also set in context to a 20 ns MD trajectory performed with the GROMOS software as described in Supporting Information[25−29] using 1UBQ as a starting structure. For the latter comparison, eq 16 was used for two reasons. First, the autocorrelation expression for the order parameters is not very suitable to obtain accurate results because the long-time tail of a correlation functions is generally plagued by poor statistics. Equation 16 yields therefore more accurate results. Second, the overall rotational motion of the protein is, in view of the correlation time, barely sampled within 20 ns of the MD simulations. To avoid undersampling of the overall rotational motion, a superposition of centers of mass followed by a least-squares rotational fit of all C^α atom positions to the initial structure was performed before the orientations of the distance vectors were determined.

Results

Validation of Cross-Relaxation Rates. NOE buildup curves of amide moieties were extracted from the ^{15}N-resolved HMQC-NOESY spectra with mixing times 0.03, 0.06, 0.09, and 0.20 s (the latter only if appropriate, see below). Typical buildup curves are shown for the spin network of residue Arg42 (Figure 2). The three NOE build-ups are representatives of strong (>0.5 s^{-1}), medium, and weak (<0.1 s^{-1}) cross-relaxation rates and large (22%), typical (14%), and small (5%) random errors. In addition, the time-dependent decrease of the diagonal of Arg42 is shown, which is used for the extractions of both I_{ii}^{det} (0) and the longitudinal relaxation rate ρ.

Overall, for 64 residues, the diagonal peaks were extracted and fitted to eq 3.1. Because of a large signal-to-noise ratio (typically, the diagonal peaks are 2 orders of magnitude more intense than the cross peaks), the fits were nearly perfect. No data could be obtained for residues Val5, Leu8, Glu24 and Gly53, Tyr59, Pro residues and both termini, because of signal

(22) Piotto, M.; Saudek, V.; Sklenar, V. *J. Magn. Reson.* **1992**, *2*, 661–665.
(23) Orekhov, V. Y.; Nolde, D. E.; Golovanov, A. P.; Korzhnev, D. M.; Arseniev, A. S. *Appl. Magn. Reson.* **1995**, *9*, 581–588.
(24) Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 12562–12566.
(25) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; Van Gunsteren, W. F. *J. Phys. Chem.* **1999**, *A 103*, 3596–3607.
(26) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656–1676.
(27) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Hermans, J. Interaction Models for Water in Relation to Protein Hydration. In *Intermolecular Forces*; Pullman B., Ed.; Reidel: Dordrecht, 1981; pp 331−342.
(28) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–342.
(29) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

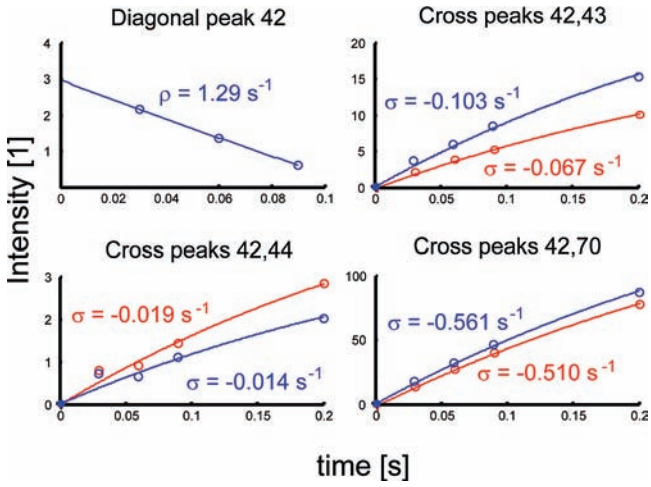

Figure 2. Peak intensities and fits involving H^N of residue Arg42. Top left panel shows the decay of the diagonal peak. The buildups of the cross peaks between residues 42 and 43 (top right), 42 and 44 (bottom left), and 42 and 70 (bottom right) are shown for the transfers originating from and terminating at Arg42 in red and blue, respectively. The intensity scales are arbitrarily chosen but identical for all buildups.

absence or spectral overlap in the NOESY spectra. Almost all extracted ρ values are between 1 and 4 s^{-1} with an average of 2 s^{-1} (see Table S1 in Supporting Information).

217 NOE build ups between amide protons were fitted to eq 3.2 using I_{ii}^{det} (0) and ρ extracted from the diagonal curves (see Table S2 in Supporting Information). The largest NOE rates are ~1.5 s^{-1}. Because of the symmetry of the ^{15}N-resolved HMQC-NOESY experiment, 77 NOEs could be determined in both directions. The presence of two buildup curves per spin pair enabled on the one hand an improvement in accuracy of the cross-relaxation rates by averaging. On the other hand, an experimental error could be determined that includes simplifications made as discussed above. On average the error is ~0.05 s^{-1} and approximately proportional to the absolute value of the rate indicating that the experimental error is small and not dominated by spectral noise. To check the validity of eqs 3.1 and 3.2 the NOE build ups were fitted using 3 time points (30, 60, and 90 ms) and 4 time points (additionally 200 ms) if they could be fit reasonably well. Linear regression between the three and four point fits yields a slope of 1.007 and Pearson's correlation coefficient of 1. Significant changes were only observed for small rates. Errors are equally, positively and negatively, distributed as expected for noise-dominated errors. In the following, averaged rates obtained from the two fits are used. Twenty-seven curves could only be fit with 3 points (30, 60, and 90 ms). These are typically weak buildups where eq 2.2 is not strictly valid because spin diffusion is likely to cause an appreciable distortion of the curve. Indeed, when the NOE cross-relaxation rates are translated into distances and the distances are compared with the 3D structures almost all of these NOEs are underestimating the distances (vide infra). Similarly, 19 curves could only be fitted with 4 points because they are typically extremely weak NOEs whose cross peaks in the first 3 points are buried in the noise and spin diffusion brings them above noise level after 200 ms mixing time. Most of them are overestimated by more than 50%. It is therefore obvious that strong spin diffusion can be easily identified by the shape and fit quality of the buildup curve. In parallel, the presence of spin diffusion can be calculated from the available structure. At the given mixing times, spin diffusion between amide protons in perdeuterated ubiquitin is significantly present between nonconsecutive residues in loops (typically >5 Å apart) and in helical secondary structure (typically <5 Å apart). Indeed, all NOE build ups classified to be strongly influenced by spin diffusion fall in these two categories. Simulations of buildup curves further show that for all other spin pairs eqs 2.1−2.3 are good approximations (see Figure S1 in Supporting Information) overestimating the rates on average by 9%. This effect corrects partially for the underestimation of the rate due to H/D equilibrium of 97%/3% leaving ca. 5% overestimation. For spin pairs in nonconsecutive residues in the helix, a rate correction factor of 0.6 may be used if no information (i.e., 3D structure) about spin diffusion pathways is available. In this study, an individual theoretical correction to the experimental cross-relaxation rate is calculated (Table S3 in Supporting Information).

Determination of the Rotational Correlation Time τ_c. To convert the cross-relaxation rates into effective distances (eq 4) as discussed below, the rotational correlation time τ_c of the molecule must be determined. If the extraction of dynamics from the cross relaxation rates is anticipated τ_c should be determined accurately since an error in τ_c has a direct impacts on the order parameter S^2 (eqs 4−6). For ubiquitin the τ_c was determined conventionally by T1/T2 and T1/T1ρ ^{15}N-relaxation measurements[3,7,30] and confirmed by measurements of cross correlated relaxation between chemical shift anisotropy of ^{15}N and the ^{15}N−^{1}H dipole−dipole coupling. Using this procedure a τ_c of 9.3 ± 1 ns was obtained for ubiquitin at a concentration of ~4.3 mmol/L at pH 5.8 and at 284 K. This is a rather large value when compared to a published value of ubiquitin at a concentration of 2 mM at 288 K[30] and by taking into account the temperature-related viscosity difference of pure water. Since the spectral quality of ubiquitin is excellent and the resonance frequencies are close to published values, we attribute the difference between experimental (9.3 ns) and predicted values (6 ns) to the higher viscosity of water at high protein concentration. Even in the case there would be partial aggregation (such as partial dimerization) in the sample (which is not supported by dynamic light scattering; data not shown) the structural analysis followed below would not change significantly because aggregation-induced errors (for example, a partial dimerization may result in a stronger anisotropy in tumbling) would be already included in the error calculations.

Extraction of Distances. The cross-relaxation rates are converted into effective distances assuming a hypothetically rigid molecule following eq 4 (see Table S2 in Supporting Information). These NOE-derived distances are compared to distances extracted from an X-ray structure at 1.8 Å resolution (pdb code: 1UBQ)[31] and an NMR structure (pdb code: 1D3Z).[32] In the X-ray structure the amide protons were placed at ideal positions with 1.01 Å H−N bond length using the program MolMol.[33] The distances from the 10 conformers of the NMR structure were linearly averaged (note: although 10 conformers represent the 3D structure it is regarded as a single structure, since every single conformer fulfills all the experimental distances and the 10 conformers represent the precision of the structure). Since

(30) Chang, S. L.; Tjandra, N. *J. Magn. Reson.* **2005**, *174*, 43–53.
(31) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531–544.
(32) Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 6836–6837.
(33) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51–55.

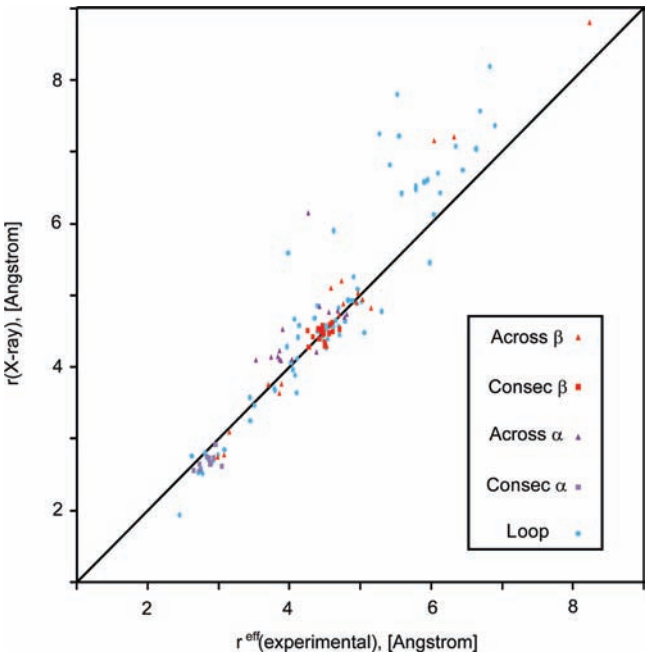

Figure 3. Correlation plot showing H^N-H^N distances obtained from the X-ray structure 1UBQ[31] with protons placed in ideal positions versus effective distances extracted from experimental cross-relaxation rates. Distances between two spins in the β sheet are red, α helix are purple, and involving one or two spins in a loop are blue. Distances between two spins of consecutive (nonconsecutive) residues in a secondary structural element are triangular (square). The slope of the black line is 1. An expansion showing error bars is provided in Figure S2 in Supporting Information.

in the NMR structure the H−N bonds have an unphysical distance of 0.98 Å,[34,35] a second NMR distance set was extracted from a structure with ideally positioned protons with H−N bond lengths of 1.01 Å.

In Figure 3 distances obtained from the X-ray structure 1UBQ are plotted versus the effective experimental distances. Correlation coefficients r and slopes s for specific secondary structure elements are listed in Table 1. Virtually the same results are obtained with the NMR structure (1D3Z) (Table 1) and therefore not discussed further. Taking the 3D structure as standard, the NOE-derived effective distances up to 5 Å are in excellent agreement with the distances in the 3D structure as demonstrated by the clustering along the slope of 1 in Figure 3. In contrast, weaker NOEs translating into effective distances >5 Å appear to result in underestimated distances. If the NOE-derived distances are discussed from a secondary structural point of view, it is observed that NOE-derived distances in β-sheets are in better agreement with the 3D structure than those in the α-helix. In detail, NOE-derived distances between consecutive residues in the α-helix are overestimated (slope 0.95) and in β-sheets nearly exact (slope 1.00). For distances between nonconsecutive residues in secondary structural elements all NOE-derived distances are underestimated. The effect is equally strong in the α-helix (slope 1.05) as in the β-sheets (1.05). Compared to the secondary structural elements, distances involving at least one loop residue show generally less correlation with the 3D structure, which may be in part due to conformational artifacts in the structures and in part due to motional effects. In addition, cross-relaxation rates between nonconsecutive residues in the α-helix and in the loops are subject to spin diffusion resulting in a distance underestimation of about 10%. Spin diffusion between amides was calculated using the 3D structure of ubiquitin. Subsequently, spin diffusion-corrected cross-relaxation rates were obtained from the NOESY experiments. Corresponding distances are in excellent agreement with the distances of the 3D structure. The slopes for all secondary elements fall within 1.00 ± 0.03 with exception of the residues in the helix (see Table S4 in the Supporting Information). Here, the slope is 0.94 and the corrections cannot account for the discussed distance underestimation from the experimental data. This observation points to dynamical behavior which will be discussed in detail below.

Conclusively, distances between amide protons in perdeuterated ubiquitin can be accurately determined. Distances shorter than 5 Å have a pairwise rmsd (root-mean-square deviation) of 0.23 and 0.25 Å with those obtained from a single NMR and X-ray structure at 1.8 Å resolution (pdb codes: 1D3Z and 1UBQ). This translates into an averaged error of 5%. Since the pairwise rmsd between distances from X-ray and NMR structures are of similar size (0.15 Å), the experimental error of the NOE-derived distances is considerably smaller than 5%. Hence, there are errors both from the NOESY experiments as well as from the structures. Furthermore, dynamics is influencing the NOE. It is therefore important to estimate the experimental error in a direct manner. Since the ^{15}N-resolved NOESY spectra are symmetric, for many spin pairs two cross-relaxation rates σ_{ij} and σ_{ji} could be determined. Conversions of these cross-relaxation rates into a distance yield a random error of ~0.07 Å (note, that this error does not take into account possible systematic errors). If NOEs of distances longer than 5 Å are included in the analysis, the pairwise rmsd doubles to 0.54 Å (X-ray) and 0.53 Å (NMR) and the error is ~7%. A slight improvement is obtained when the NOE-derived distances are corrected for spin-diffusion (0.45 Å for X-ray or 0.44 Å for NMR). Overall, highly accurate distances in perdeuterated ubiquitin from NOEs were established. The reason for this accuracy lies in the fact that the distance is proportional to the sixth power root of the cross-relaxation rate, which very effectively reduces experimental errors, motional and spin-diffusion effects. For illustration, a rate overestimation of 5% results in less than 1% underestimation of the distance.

Extraction of Order Parameters. If a single 3D structure is a true average representation of the conformational ensemble of ubiquitin, a comparison between cross-relaxation rates extracted from the NOE build ups and those calculated from the structure enables the extraction of some aspects of through space motional disorder described by the order parameter S_{XY}^2 (see Theory). According to eq 5, the order parameter S_{XY}^2 is proportional to the ratio of experimental and calculated NOE cross-relaxation rates (note: the same information is in principle also contained in the NOE-derived and structure-derived distance comparison discussed above; see also eq 7). Figure 4 shows correlation plots of calculated rates based on various structures versus the experimental rates. Table S5 in the Supporting Information presents order parameters S_{XY}^2. Because errors in the cross-relaxation rates affect the orders parameters proportionally, the experimental cross-relaxation rates were corrected for the H/D equilibrium and spin diffusion using the 3D X-ray structure as a basis. The order parameters also depend on the structure from which r^{rigid} is calculated. The pairwise rmsd between S_{XY}^2 based on the X-ray and NMR structures is 0.15 for consecutive residues and 0.29 for nonconsecutive residues.

(34) Case, D. A. *J. Biomol. NMR* **1999**, *15*, 95–102.
(35) Yao, L. S.; Vögeli, B.; Ying, J. F.; Bax, A. *J. Am. Chem. Soc.* **2008**, *130*, 16518–16520.

Table 1. Slopes *s* and Pearson's Correlation Coefficients *r* between Effective Experimental and Extracted H^N-H^N Distances in Ubiquitin

atom pair[a]	all	consecutive β strand[b]	between β strands	consecutive α helix	nonconsecutive α helix	loops[b]
s	1.04	1.00	1.05	0.95	1.05	1.05
r^2	0.87	−0.61	0.92	0.01	0.00	0.85
r	0.93		0.96	0.10	0.07	0.92

[a] All distances are taken from the X-ray structure 1UBQ.[23] [b] All pairs of which at least one atom is in a loop.

Since the experimental error based on the symmetry relationship is 0.21 (averaged error 0.11) for consecutive residues and 0.11 (averaged error 0.09) for nonconsecutive residues, respectively, the order parameter for consecutive residues is largely independent of the structure used as input for consecutive residues, whereas structural noise may influence the order parameters for nonconsecutive residues. The latter order parameters must therefore be interpreted with care.

The X-ray structure or the NMR structure used as reference structures are not true average representations of the conformational ensemble. Therefore, the exact order parameters obtained by NOEs cannot be deduced. Nevertheless, a motional network map might be established. Figure 5 shows a ribbon plot of ubiquitin where the color at residue *i* is coded according to the spin diffusion corrected order parameter between *i* and *i+1*. Order parameters between nonconsecutive residues are represented by sticks colored according to the same code. For through-space vectors between amides of consecutive residues, all order parameters S_{XY}^2 with an error smaller than 0.10 are between 0.40 and 1.15 with most of the largest values at the termini. Thirty percent of the values fall below 0.7. Values larger than 1.3 are obtained for residues 21/22, 22/23, 43/44 and 50/51. For order parameters S_{XY}^2 of nonconsecutive residues, reliable values are between 0.5 and 1.7. Many other values are not reliable because they have no or a large error or are strongly structure dependent.

To avoid the danger of overinterpretation, only average values in specific secondary elements are compared. Averaged order parameters S_{XY}^2 are calculated by inverting the slopes in Figure 4 (Table 2). In the β-sheet, order parameters S_{XY}^2 between consecutive residues cluster around 1, whereas those of spin pairs between two β-strands are much smaller than 1. Even larger differences are observed between consecutive and nonconsecutive residues in the α helix. Order parameters S_{XY}^2 between consecutive residues are clearly smaller than 1, whereas order parameters S_{XY}^2 between nonconsecutive residues are much larger than 1. Order parameters S_{XY}^2 between spin pairs involving loop residues are generally much smaller than 1. Although as mentioned above these values are not reliable due to structural and experimental uncertainties, the extracted order parameters suggest that very different types of motion are present in ubiquitin. Order parameters larger than 1 are dominated by distance fluctuations and those smaller than 1 are dominated by angular fluctuations.[12] Hence, along a β-strand in ubiquitin motion appears to be highly restricted, whereas interstrand motions are present and are of angular nature. In the α-helix, it is conceivable to speculate that motions between amide protons of two consecutive residues are primarily an effect of angular fluctuations of the relatively short H^N-H^N vectors. In contrast, fluctuations of vectors between nonconsecutive amides are also affected by distance changes. These considerations evoke a picture of the α-helix with twisting along the backbone and stretching along the helix axis. In the loops, strong motional effects are expected, and order parameters smaller than 1 point to orientationally dominated fluctuations.

It has been shown that positions of amide protons are an important factor in high-precision measurements of H^N-H^α scalar couplings and RDCs.[36,37] To analyze impacts from such structural uncertainties, Table S6 in the Supporting Information lists averaged order parameters between consecutive and nonconsecutive residues for β-sheets and α-helices. Overall, order parameters based on the X-ray and NMR structures are similar. Generally, consecutive vectors have smaller values for the NMR structure (−8%). A comparison of the NMR structure 1D3Z with the control structure in which the protons are replaced to ideal positions at 1.01 Å H−N bond lengths shows that generally the same order parameters are extracted. However, consecutive order parameters are clearly smaller due to the relatively small spin pair distances. As expected, the effect is most pronounced in β-sheets (−5%). This is a direct consequence of unphysically short H−N bond lengths that falsely increase the amide−amide distances. Hence, since many structures deposited in the protein data bank (PDB) are calculated with 0.98 Å H−N bond lengths, it is advisible to rescale the bond lengths to 1.015 Å H−N[34,35] before extraction of order parameters from NOEs.

Comparison to a Set of Structures Covering Motion. The use of order parameters in describing the motion of a protein is limited because many dynamical models could explain the NMR observables including the NOE of interest. Furthermore, the extraction of the order parameters above is biased by the assumption that the structure used is a true average representation of the conformational ensemble and because some motions may be missed by the NOE through cancellation effects. Therefore, an ensemble representation of the structure that covers motion would be superior to a single static structure. Such a set of structures is supposed to be the recently published RDC-derived set of 116 structures of ubiquitin (pdb code: 2k39).[18] To provide an estimate of the conformational space sampled by this representation, Table S7 in the Supporting Information presents order parameters of H^N-N bonds. Apart from a uniform scaling due to small fluctuations, these order parameters are only affected by angular motion and are strictly smaller than 1.[34,35] Averaged over the residues relevant for our study, the order parameter is 0.76. Because all time scales are included, it is lower than the values obtained from Lipari−Szabo modelfree analysis which covers only subnanosecond time scales.[6] Order parameters obtained from modelfree analysis of ^{15}N relaxation data at 288 K and an H^N-N distance of 1.02 Å presented in reference are on average 0.83 (kindly provided by Nico Tjandra).[30] Bonds for which this order parameter is substantially larger than the one obtained from the RDC-derived structure set are undergoing considerable motions on the slow time scale.

Because RDCs are differently sensitive to the time scales of motion than NOEs, it is not possible to calculate the exact cross-correlation rates from an RDC-derived structure. Hence, for a

(36) Vögeli, B.; Ying, J. F.; Grishaev, A.; Bax, A. *J. Am. Chem. Soc.* **2007**, *129*, 9377–9385.
(37) Vögeli, B.; Yao, L.; Bax, A. *J. Biomol. NMR* **2008**, *41*, 17–28.

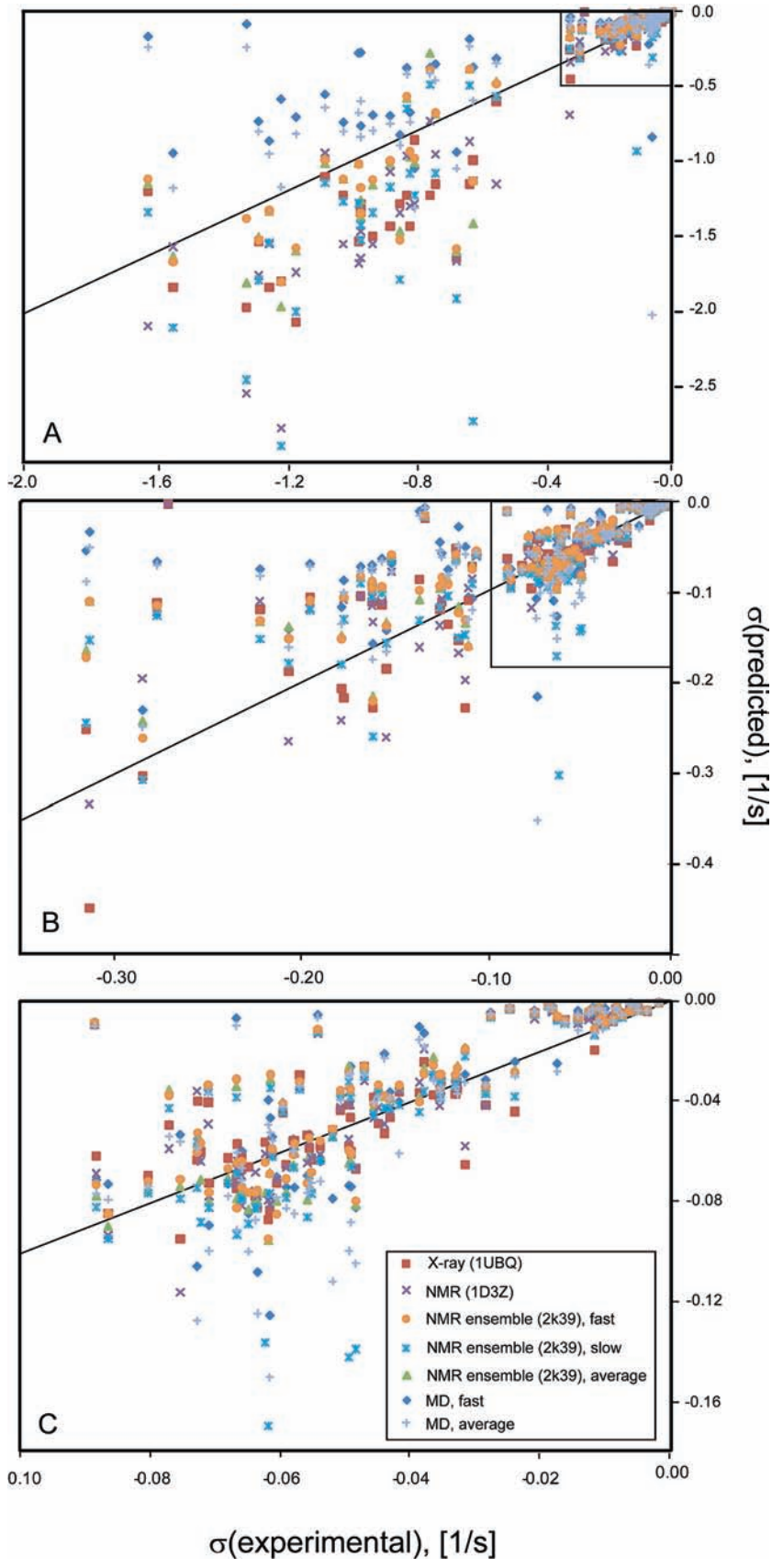

Figure 4. Correlation plots showing H^N-H^N NOE cross-relaxation rates predicted from structures versus experimental cross-relaxation rates. (A) All rates, (B) expansion of the medium and weak rates, and (C) expansion of the weak rates only. Values are obtained from the X-ray structure 1UBQ (red square), the NMR structure 1D3Z (purple cross), the NMR RDC-derived structure set 2k39 assuming only fast motion (orange circle), assuming only slow motion (blue star), and linearly averaged distances in 2k39 (green triangle), and the MD simulation assuming only fast motion (dark blue diamond) and linearly averaged distances (bright blue diamond). The slope of the black lines is 1. A plot including error bars is provided in Figure S3 in Supporting Information.

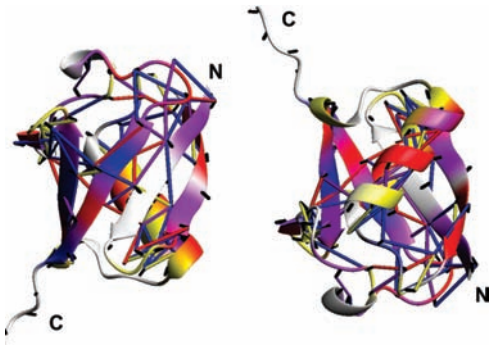

Figure 5. Ribbon representation of ubiquitin depicting order parameters. Order parameters S_{XY}^2 between spins of consecutive residues are color coded on the secondary element ribbon and those between nonconsecutive residues on connecting sticks. The color code is: yellow <0.7; red 0.7−0.9; purple 0.9−1.1; blue >1.1 and white if no value is available. The H−N bonds are shown in black. (Left) View on the β sheet and (right) view on the α helix. The figure was prepared with the program MolMol.[33]

comparison between the RDC-derived structure set with the measured NOEs, the cross-relaxation rates were calculated twice assuming either exclusively fast or exclusively slow (slower than molecular tumbling) motions (Table 2) with the true value between these two extremes. The slow motion-based rates are larger than rates calculated in the presence of fast motion and are in principle too large since fast motion is always present in a biomolecule.[6−8] In addition, the rates were also calculated based on linearly averaged distances (Table 2) to determine whether deviations between calculated and experimental rates originate from structural or motional differences.

As shown in Table 2, the RDC-set of structures reflects the motion between consecutive residues in β-strands very well with the inverse slope between calculated and experimental rates fairly close to 1. While the fast motions bring the calculated rates closer to the experimental values, the averaged distances give a lower value (−10%). A very different picture emerges for spin pairs across β-strands. The rates calculated from the RDC-set are underestimated by 47%. Even if all motions present were assumed to be of slow nature, the cross-relaxation rates would still be too small. By using the averaged distances from the set of structures, the rates are also underestimated (56%) indicating that on average the β-strands may be placed too far apart from one another in the RDC bundle. In contrast, the static X-ray and NMR structures overestimate these rates by 30−40%, which we attributed above to the lack of motions. Hence, neither the X-ray structure nor the RDC-derived structures represent the inter β-strand structure and motion of ubiquitin adequately. Furthermore, since the correlation coefficients of the RDC-bundle listed in Table 2 are similar to those for the single structures, no improvement by the introduction of a set of structures is obtained to explain the experimental NOEs, although the RDC set of structures should cover the motions of ubiquitin. In the α-helix, rates between consecutive residues are underestimated by the RDC-derived set of structures but to a smaller degree than by the static structures. Again, this appears to be a consequence of the mean structure rather than of motions since distance averaging yields virtually the same slope. For the nonconsecutive helical residues, both X-ray and NMR structures as well as the RDC-derived structural ensemble representation predict the experimental rates well, although the RDC set of structures is in slightly better agreement than the X-ray structure. Again, this appears to be an effect of the mean structure (Table 2). The only significant improvement of the RDC-derived bundle of conformers over the single structure is found for cross-relaxation rates between spins located in loops. Incorporation of fast motion lifts the inverse slope from 0.87 (as obtained with averaged distances) to 0.92. As mentioned above, these observations may be due to large errors in the X-ray structure, for example, due to crystal packing artifacts. Overall, correlation coefficients for ensemble-based rates are similar to the static structure-based rates. Hence, on the basis of the NOE analysis, the RDC-derived set of structures does not obviously outperform the X-ray structure even though it should reflect motional effects.

Another way to gauge the effects of motions is achieved by extraction of order parameters as represented by the RDC-derived structure set. Order parameters between amides S_{XY}^2 are calculated with eq 7 where the brackets are replaced by the average over the conformers. Table S5 in the Supporting Information lists the values assuming the presence of fast motions only. Although most order parameters S_{XY}^2 are close to 1 (±0.05), there are some large deviations. Spin pairs strongly dominated by distal motion are res17/res18 ($S_{XY}^{\text{ensemble2}} = 1.07$), 34/35 (1.09), 4/64 (1.08), 6/69 (1.10), 21/57 (1.09), 44/68 (1.41), 45/68 (1.14), 47/68 (1.11), 50/68 (1.08), 55/58 (1.08), and 62/66 (1.08). However, for these spin pairs, the RDC-derived representation does not fit the experimentally derived NOE rates better than the static structures. The calculated rates are typically too low for the distances averaged over the set of structures and the order parameter pulls them closer to the experimental values, whereas both static structures do not need a correction term that takes motion into account. One may argue that part of the motion in the structure set takes place on the slow time scale, which would further increase the calculated rates. However, even if only the presence of slow motion is assumed most of the cross relaxation rates are still smaller than those from the static NMR and X-ray structures. Some spin pairs are dominated by angular fluctuations. These are res9/res10 (0.76), 47/48 (0.80), 62/63 (0.84), 63/64 (0.88), 70/71 (0.84), 71/72 (0.80), and 10/12 (0.87). For the amide−amide distances between residues 9/10, 62/63, 63/64, and 70/71 the RDC-derived set of structures correctly accounts for the motional effect highlighting that in general large motion-based modifications of rates are physically possible. Overall, however, the RDC-derived structure representation does not explain the experimental NOE data better than the single X-ray structure.

Comparison to a Molecular Dynamic Simulation. The extracted NOE data can also be compared to cross-relaxation rates calculated from molecular dynamic simulations following the procedures described above for the RDC-derived set of conformations. Here, this analysis is applied to the 20 ns MD simulation of ubiquitin. Obviously, the ensemble of conformers of interest is time-resolved along the time trajectory and the motions present are only of fast nature because of the 20 ns time limit of the simulation. Hence, NOE cross-relaxation rates from the MD ensemble can be calculated using eq 16 (Table 2). In all types of structural elements, the MD-derived NOE rates are significantly larger than the experimentally derived values (see Table 2). This discrepancy may originate from the small sampling time of the MD simulation that excludes slow motion as well as may sample ns-motions inadequately, which could decrease the MD-derived rates considerably. However, the MD simulation seems also to fail to represent adequately the structure and dynamics on the fast time scale. Support for this is based by the local order parameters of H^N-N bonds, which averaged over the relevant residues is 0.68 and hence

Table 2. Inverted Slopes 1/*s* and Pearson's Correlation Coefficients *r* between Spin-Diffusion and H/D Equilibrium-Corrected Experimental and Calculated H^N-H^N Cross-Relaxation Rates σ of Ubiquitin[a]

structure		all[b]	all consec[b]	all nonconsec	consec β strand	between β strands	consec α helix	nonconsec α helix	loops[b,c]
1UBQ	1/*s*	0.71	0.71	0.71	0.98	0.67	0.72	0.77	0.77
	r^2	0.94	0.92	0.91	−0.27	0.97	0.03	0.36	0.91
	r	0.97	0.96	0.96		0.98	0.17	0.60	0.95
1D3Z[d]	1/*s*	0.66	0.65	0.80	0.91	0.78	0.70	0.73	0.65
	r^2	0.93	0.92	0.85	−0.01	0.89	−4.78	0.39	0.90
	r	0.96	0.96	0.92		0.94		0.62	0.95
1D3Z, ideal[e]	1/*s*	0.66	0.65	0.75	0.97	0.72	0.70	0.73	0.63
	r^2	0.92	0.91	0.85	−0.32	0.90	−3.86	0.39	0.89
	r	0.96	0.95	0.92		0.95		0.61	0.94
2k39, fast[f]	1/*s*	0.85	0.77[i]	1.47	0.87[i]	1.52	0.79	0.87	0.94
	r^2	0.91	0.90	0.90	−0.80	0.92	−1.01	0.63	0.92
	r	0.95	0.95	0.94		0.96		0.80	0.96
2k39, slow[g]	1/*s*	0.64	0.57[i]	1.22	0.78[i]	1.27	0.66	0.79	0.67
	r^2	0.84	0.81	0.91	−0.72	0.94	−0.77	0.63	0.84
	r	0.92	0.90	0.95		0.97		0.79	0.92
2k39, average[h]	1/*s*	0.81	0.74[i]	1.56	0.81[i]	1.64	0.79	0.85	0.87
	r^2	0.89	0.89	0.85	−1.01	0.86	−1.38	0.62	0.91
	r	0.95	0.94	0.92		0.93		0.79	0.95
MD, fast[f]	1/*s*	2.08	2.04	2.13	0.88	2.13	1.30	1.16	2.78
	r^2	0.52	0.31	0.79	−0.96	0.84	−3.56	0.04	0.35
	r	0.72	0.56	0.89		0.92		0.21	0.59
MD, average[h]	1/*s*	1.61	1.59	1.85	0.80	1.89	1.12	1.09	1.89
	r^2	0.46	0.23	0.81	−0.97	0.86	−1.75	0.04	0.23
	r	0.68	0.48	0.90		0.93		0.21	0.48

[a] For the 1UBQ and 1D3Z structures the inverted slope 1/*s* is also the averaged order parameter S_{XY}^2. [b] Spin pair 34−35 omitted. [c] All pairs of which at least one atom is in a loop. [d] H^N placed as in 1D3Z pdb deposition with $r_{HN} = 0.98$ Å. [e] H^N placed at ideal position with $r_{HN} = 1.01$ Å. [f] $\sigma \propto \langle (P_2(\cos \vartheta_{int,ab^M}))/(r_a^3 r_b^3) \rangle$. [g] $\sigma \propto \langle (1)/(r^6) \rangle$. [h] $\sigma \propto (1)/(\langle r \rangle^6)$. [i] Corrected for changes caused by $r_{HN} = 1.04$ Å as used in 2k39 instead of 1.01 Å as used in this study.

somewhat smaller than the one from the RDC-derived set (0.76) and clearly smaller than the experimental one obtained from the modelfree analysis of ^{15}N relaxation data (0.83) (Table S7, Supporting Information).[30] Furthermore, a detailed comparative analysis of the through space order parameters S_{XY}^2 obtained from the RDC-derived set of structures versus those from the MD simulation do not show any correlation but a linear regression with a slope *s* of 1.00 (see Table S3 in the Supporting Information). In addition, separating S_{XY}^2 into the angular and distal contributions reveal that the averaged Legendre polynomials are only slightly correlated ($r = 0.65$, $s = 1.00$), whereas $\langle (1)/(r^3) \rangle^2$ is even less so ($r = 0.53$, $s = 0.28$). Therefore, the generally lower rates predicted from the MD simulation originate from longer averaged distances rather than from motional effects. Indeed, rate calculations based only on the average distances effectively avoiding impacts from motional effects are larger than ensemble-derived rates but still too small. Because the correlation coefficients between calculated cross-relaxation rates and experimentally derived NOE rates are similar for the X-ray structure, and both set of structures (i.e., RDC set, and MD ensemble), it appears that the conformational dispersion present in the set of ubiquitin structures do not reflect the motions probed by NOEs well.

Discussion

A method is presented that yields very accurate proton−proton distances for up to 5 Å in perdeuterated proteins from NOEs. The error is about 0.1 Å. Such distances can be extracted by assuming a simple 2-spin model. Extraction of motional information from NOEs is more challenging and corrections for spin diffusion are required. For example, in an α-helix, rates between nonconsecutive residues can be corrected by a uniform factor of 0.6 or the contribution of spin diffusion can be calculated for each spin pair if the structure is known. Once the NOE cross-relaxation rates are determined accurately, the extraction of motions from NOEs is possible by several approaches of which each has its caveat: (i) A comparison between a static structure such as the X-ray structure or the NMR structure with NOE-derived distances may give insights into motion (Figure 5). However, it appears that the crystal structure or the NMR structure are not true average representatives of the protein structural ensemble and structural noise may therefore mask much of this motion. Furthermore, since under certain circumstances radial ($S_{XY}^{rad2} > 1$) and angular ($S_{XY}^{ang2} < 1$) motions can cancel each other out in their contributions to the NOE, the comparison between a structure and NOEs may in part underestimate motions. Nonetheless, a detailed analysis with the protein ubiquitin suggests little motion along β-strands, motion between β-strands, and larger complex motions in its α-helix. These observations are in agreement with recent studies on ubiquitin[18,38] and the similarly folded GB3.[39,40] In both proteins, the β-sheet has been shown to undergo a collective rotational fluctuation along the polypeptide chain. (ii) The NOE cross-relaxation rates may be used to validate an MD simulation or a differently derived set of conformers that is believed to reflect motion. For ubiquitin, neither the MD simulation covering fast motions nor the RDC-derived set of structures covering motions up to ms are in better agreement than the static X-ray structure (Table 2). (iii) A cross-relaxation rate-constrained MD simulation that is constrained through ensemble-averaging would result in an MD simulation that fulfills the experimental restraints.[41,42] (iv) A simultaneous structure and dynamic calculation using exact NOEs could be determined by the

(38) Lakomek, N. A.; Fares, C.; Becker, S.; Carlomagno, T.; Meiler, J.; Griesinger, C. *Angew. Chem., Int. Ed.* **2005**, *44*, 7776–7778.
(39) Bouvignies, G.; Bernado, P.; Meier, S.; Cho, K.; Grzesiek, S.; Bruschweiler, R.; Blackledge, M *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13885–13890.
(40) Vögeli, B.; Yao, L. *J. Am. Chem. Soc.* **2009**, *131*, 3668–3678.
(41) Torda, A. E.; Scheek, R. M.; Van Gunsteren, W. F. *J. Mol. Biol.* **1990**, *214*, 223–235.

use of true ensemble averaging through the structure calculation process.[43] Such an approach requests, however, the measurements of NOEs between all types of protons resulting in a dense network of exact average distances. Furthermore, because of the internal motion-dependent averaging procedure needed to translate NOE into distances (eqs 10.1 and 10.2) additional input data such as the extent of fast internal motion is required (for example in the case of H^N-H^N NOEs H^N-N heteronuclear Lipari−Szabo order parameters could be used to get knowledge about subnanosecond motion). Alternatively, the lack of knowledge about which protocol to use for the conversion of the NOE into a distance might be pushed into the error calculation by the use of lower and upper distance restraints, which would then cover the presence of both extreme regimes, that is, only fast internal motion or only slow internal motion. Once successful, such a distance network would thereby provide a compact through-space sampling of protein structure and dynamics including both backbone as well as side chains, and possibly would result in the extraction of correlated motion.[39,40] Preliminary experiments in this regard are in progress.

Conclusion

NOEs in NMR experiments are influenced both by the conformation and the dynamics of a biomolecule. Here, it is demonstrated that the determination of NOE buildup rates results in precise average distances, which can be used to get a very detailed structural and possibly dynamical picture of a protein. Since a large amount of NOEs can be collected, the measurement of exact NOEs may enable the establishment of a comprehensive representation of a protein characterizing both its 3D structure as well as its motion.

Acknowledgment. We thank Dr. Jason Greenwald for the preparation of the ubiquitin sample. This work was supported by the SNF.

Supporting Information Available: Derivations of eq 15 and 16; description of simulation of the impact of spin diffusion on apparent cross-correlation rates; details on the MD simulation; figure showing simulated buildup curves with and without spin diffusion; figure showing a correlation plot between distances obtained from the X-ray structure versus effective distances with error bars; figure showing a correlation plot between cross-relaxation rates obtained from various structures and experimental cross-relaxation rates with error bars; table presenting autorelaxation rates ρ and initial intensities $\Delta I_z(0)$; table presenting experimental and predicted cross-relaxation rates, experimental effective distances, and distances obtained from structures; table presenting apparent cross-relaxation rates obtained from fits of simulated buildup curves; table presenting slopes and correlation coefficients between effective experimental and extracted distances after correcting the experimental cross-relaxation rates for spin diffusion and the H/D equilibrium; table presenting order parameters S_{XY}^2 of H^N-H^N spin pairs; table presenting order parameters S_{XY}^2 of H^N-H^N spin pairs for specific secondary structure elements normalized to those obtained from 1UBQ. This material is available free of charge via the Internet at http://pubs.acs.org.

JA905366H

(42) Lindorff-Larsen, K.; Best, R. B.; DePristo, M. A.; Dobson, C. M.; Vendruscolo, M. *Nature* **2005**, *433*, 128–132.
(43) Clore, G. M.; Schwieters, C. D. *J. Mol. Biol.* **2006**, *355*, 879–886.
(44) Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1989**, *85*, 393–399.
(45) Shaka, A. J.; Keeler, J.; Frenkiel, T.; Freeman, R. *J. Magn. Reson.* **1983**, *52*, 335–338.