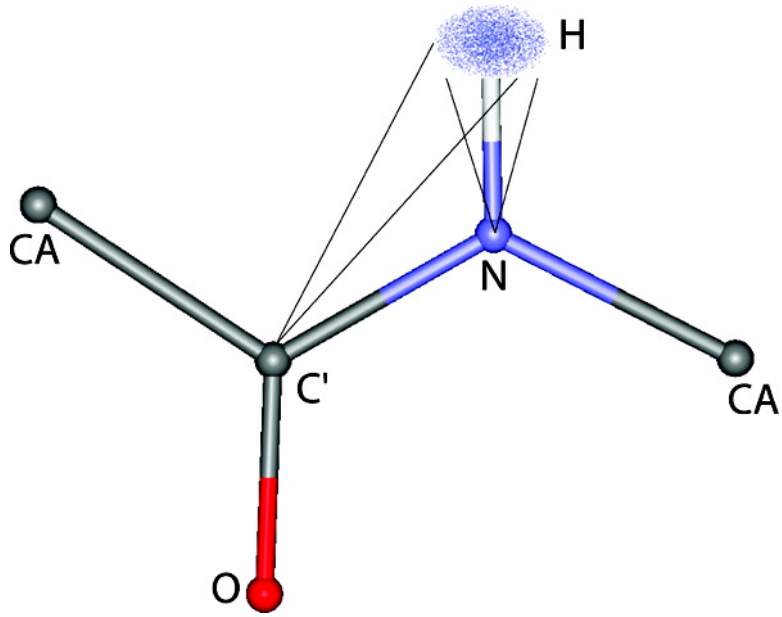
Communication

NMR Determination of Amide N#H Equilibrium Bond Length from Concerted Dipolar Coupling Measurements

Lishan Yao, Beat Vo#geli, Jinfa Ying, and Ad Bax

J. Am. Chem. Soc., **2008**, 130 (49), 16518-16520 • DOI: 10.1021/ja805654f • Publication Date (Web): 17 November 2008



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
COMMUNICATIONS

Published on Web 11/17/2008

NMR Determination of Amide N−H Equilibrium Bond Length from Concerted Dipolar Coupling Measurements

Lishan Yao, Beat Vögeli, Jinfa Ying, and Ad Bax*

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

Received July 20, 2008; E-mail: bax@nih.gov

Quantitative interpretation of ^{15}N NMR relaxation rates in terms of motion hinges on the value of the ^{15}N−^{1}H internuclear distance. Neutron diffraction of small model peptides suggested a value of 1.020−1.024 Å.[1] From an NMR relaxation perspective, spectral density terms resulting from ultrafast motions such as zero-point librations do not significantly contribute to relaxation and simply attenuate the effective internuclear ^{15}N−^{1}H dipolar coupling. Their impact is therefore similar to a lengthening of the N−H bond, and an increased effective bond length can be used to remove the effect of the nearly uniform normal mode librations and vibrations from the modelfree order parameters, often extracted from such data.[2] Internal motions also scale the internuclear dipolar couplings observed in solid-state and liquid-crystal NMR, with both technologies potentially providing access to quantitative analysis of the amplitude and direction of motions integrated over the entire submillisecond time scale. Although the equilibrium N−H bond length is known to be sensitive to hydrogen bonding, DFT calculations indicate this dependence to be weak for normal, high barrier hydrogen bonds, with an approximate 0.01 Å decrease in r_{NH}^{eq} when $r_{H\cdots O}$ increases from 1.8 to 2.5 Å, in agreement with prior computational results (Supporting Information).[2,3]

Quantum mechanics defines the lowest vibrational states of any molecular bond. A detailed description of the impact of normal mode vibrations (bond stretching) and librations (bending) on NMR-observable parameters has been presented by Case.[2] Different definitions of the bond length are required for different purposes (Figure 1A): The equilibrium bond length, r_{NH}^{eq}, corresponds to the internuclear distance of lowest energy. The average $<r_{NH}>$ distance includes averaging over the lowest vibrational bond stretching mode, which due to anharmonicity of the potential is calculated to be ca. 0.015 Å longer than r_{NH}^{eq}. From an NMR relaxation perspective, where the strength of the dipolar coupling scales with r_{NH}^{-3}, the pertinent bond length $r_{eff} = <r_{NH}^{-3}>^{-1/3}$. As motions of the N−H bond vector caused by processes other than the universal and largely structure-independent zero-point normal mode librations and vibrations also scale the ^{15}N−^{1}H dipolar interaction, it has remained difficult to establish experimentally precise baseline values for $<r_{NH}>$ and r_{eff}. Here, we report concerted measurement and analysis of ^{13}C′−^{13}C$^{\alpha}$, ^{13}C′−^{1}H^{N} and ^{15}N−^{1}H^{N} residual dipolar couplings (RDCs), measured at very high precision in perdeuterated, amide-protonated forms of five mutants of the small and previously well-studied protein GB3. These mutants, K19AD47K, K19ED40N, K19AT11K, and two which include a C- and N-terminal His-tag, K19EK4A-C-His$_6$, and K19EK4A-N-His$_6$, previously were shown to map the five-dimensional alignment space when dissolved in a liquid crystalline Pf1 suspension.[4] Using an improved stability, iterative implementation[5] of the DIDC method,[6] alignment tensors are derived independent of a starting structure, that is, with no adverse impact from coordinate uncertainty in the reference model. Alternate iterative approaches can also be used for this purpose.[7,8] Excluding the invisible N-terminus, four residues identified by iterative DIDC as highly dynamic, and five residues for which resonance overlap interfered with accurate measurement of ^{13}C′−^{13}C$^{\alpha}$, ^{13}C′−^{1}H^{N}, or ^{15}N−^{1}H^{N} RDCs in one or more of the five mutants, leaves 45 sets of 3 RDCs per mutant for analysis.

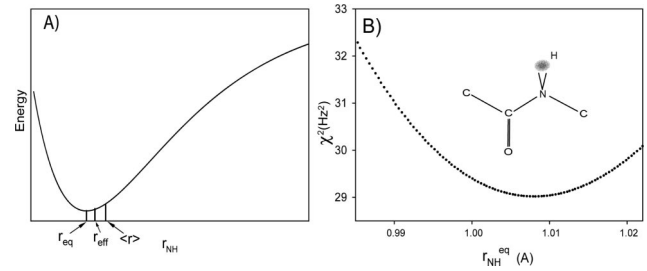


Figure 1. Amide N−H bond length in peptide groups. (A) Bond energy versus r_{NH} bond length and definitions of r_{NH}^{eq}, $<r_{NH}>$, and $r_{eff}(<r_{NH}^{-3}>^{-1/3})$. (B) Total χ^2 fitting error of C$^{\alpha}$−H^{N} RDCs to GB3 structure (see text) as a function of r_{NH}^{eq}, with $\chi^2=\sum_i\sum_j (RDC_{ij}^{pred} - RDC_{ij}^{exp})^2$, where i is the alignment index and j is the residue index. Values for $<r_{NH}>$ and r_{eff} follow directly from r_{NH}^{eq} (see Figure S2). The H^{N} cloud in the inset represents the proton distribution generated using eq S28.

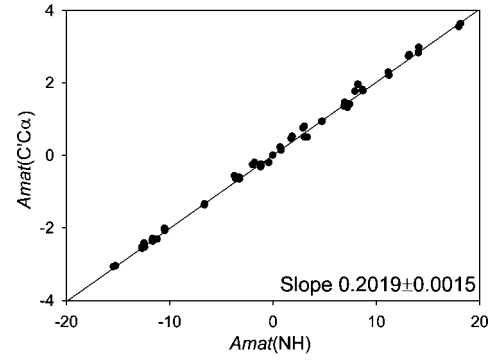


Figure 2. Alignment matrix elements obtained using iterative DIDC[5] from C$^{\alpha}$−C′ RDCs (y axis) versus values obtained from N−H RDCs. Matrix elements result from five mutants, each measured twice with slightly different HNCO experiments and different GB3 and Pf1 concentrations.

Assuming uniform C′−C$^{\alpha}$ and N−H^{N} bond lengths, the alignment tensor coefficients derived from the iterative DIDC method can be written in matrix format, Amat, and their ratio (Figure 2)

$$\frac{\mathrm{Amat(C'C^{\alpha})}}{\mathrm{Amat(NH)}} = \frac{\gamma_C\gamma_C S_{C'C\alpha}\left\langle \frac{1}{r_{C'C^{\alpha}}^{3}} \right\rangle}{\gamma_N\gamma_H S_{NH}\left\langle \frac{1}{r_{NH}^{3}} \right\rangle} = 0.202 \qquad (1)$$

 10.1021/ja805654f

Table 1. Average Dynamics Parameters and Amide N−H Bond Lengths (Å)

	r_{NH}^{eq} [a]	r_{eff} [b]	$<r_{NH}>$ [c]	S'_{NH}	σ_{tr} [g]	$S'_{C'HN}$ [h]
this work [d]	1.008 ± 0.006	1.015 ± 0.006	1.023 ± 0.006	0.910 ± 0.016	9.9° ± 0.9°	0.981 ± 0.004
others			1.022[e]	0.938[f]	8.2°[f]	0.994[f]

[a] r_{eq} is the equilibrium bond length. [b] $r_{eff} = <r_{NH}^{-3}>^{-1/3}$, where <...> denotes the bond stretching average. [c] $<r>$ is the stretching-averaged bond length with the difference between r_{eq} and $<r>$ resulting from anharmonicity. [d] Uncertainty in the averaged value based on jackknife analysis. [e] $<r_{NH}>$ from neutron scattering.[1] [f] Value for N−H zero-point libration in *N*-methylacetamide, obtained from MP2 quantum calculation, divided by the corresponding $S_{C\alpha C'}$ value of 0.990. [g] $\sigma_{tr} = (\sigma_u^2 + \sigma_v^2)^{1/2}$. [h] $S'_{C'HN}$ is the C′−H^N order parameter relative to the C$^\alpha$−C′−N backbone.

represents the ratio of the strengths of the $^{13}C'-^{13}C^\alpha$ and $^{15}N-^1H^N$ dipolar interaction constants, which falls very close to the value previously reported[9] using ubiquitin RDCs and its X-ray structure. Besides the assumption of uniform C′−C$^\alpha$ and N−H^N bond lengths, two additional approximations are used in eq 1: First, all 45 included residues have uniform S_{NH} as well as $S_{C'C^\alpha}$ order parameters. Second, angular internal motion of C′−Cα and N−H^N vectors, which determines S_{NH} and $S_{C'C^\alpha}$, is assumed to be axially symmetric. Using simulated data, we previously have shown[5] that the impact of these approximations upon the derived Amat values, used to calculate the ratio in eq 1, is much smaller than the uncertainty caused by RDC measurement error. The very tight correlation observed in Figure 2 also confirms the validity of these assumptions.

With values for the average order parameters $S_{C'C^\alpha}$ and S_{NH} being unknown, the RDC data do not uniquely define $<r_{NH}^{-3}>$ and $<r_{C\alpha C'}^{-3}>$. The brackets, < >, denote vibrational averaging, where, for C$^\alpha$−C′ to a very good approximation,[2] $<r_{C\alpha C'}> = <r_{C\alpha C'}^{-3}>^{-1/3} = 1.525$ Å.[10] By defining $S'_{NH} = S_{NH}/S_{C'C\alpha}$, a value of

$$S'_{NH}<r_{NH}^{-3}> = 0.872 \pm 0.006\ \text{Å}^{-3} \qquad (2)$$

is then obtained from the Amat ratios (Figure 2). Note that S'_{NH} defines the order of the N−H vectors relative to the average order parameter of the C$^\alpha$−C′ vectors, which is taken as a proxy for the motion of the entire backbone, including C′−N and N−C$^\alpha$ vectors. Fluctuations of the peptide plane about the $C^\alpha_{i-1}-C^\alpha_i$ vector, so-called γ motions,[11−14] have a larger impact on the orientations of N_i-H_i vectors than on the backbone $C^\alpha_{i-1}-C'_{i-1}$, $C'_{i-1}-N_i$ and $N_i-C^\alpha_i$ vectors. Together with zero-point librations of the N−H vector, these motions therefore dominate the impact on S'_{NH}. The problem of finding the average N−H bond length and angular motion is now reduced to solving the positional distribution of the H^N nuclei relative to the backbone frame.

Considering that N−H bond stretching motions occur at higher frequencies than librational motions, librational and stretching motions to first order can be considered uncorrelated, allowing factorization of the positional distribution of the proton associated with any given N−H vector:

$$\rho(r, u, v) \approx \rho(u, v)\ \rho_1(r) \qquad (3a)$$

where $\rho_1(r)$ denotes bond stretching (Supporting Information, eq S25) and

$$\rho(u, v) = \frac{1}{2\pi\sigma_u\sigma_v} \exp(-u^2/2\sigma_u^2) \exp(-v^2/2\sigma_v^2) \qquad (3b)$$

separates the angular motion into in-plane (σ_v) and out-of-plane (σ_u) components, described by two independent Gaussian distributions. Bond stretching motions impact averaging of the inverse cubic dependence of $^{15}N-^1H$ dipolar couplings on r_{NH}. On the basis of the narrow range of infrared N−H stretching frequencies, the amplitudes of such motions are quite uniform, and their impact can be calculated (Supporting Information) to yield:

$$r_{eff} \equiv <r_{NH}^{-3}>^{-1/3} \approx r_{eq} + 0.007\ \text{Å} = <r_{NH}> - 0.008\ \text{Å} \qquad (4)$$

In our study, the time-averaged N−H orientations are taken from the iterative DIDC analysis of the very high precision RDC data, recorded on the perdeuterated mutants of GB3. High accuracy of these N−H orientations is exemplified by exceptionally good fits to the $^3J_{HNH\alpha}$ Karplus curve (rmsd 0.33 Hz; Supporting Information). Three independent parameters σ_u, σ_v, and $<r_{NH}>$ are then determined by fitting C′−H^N and $^{15}N-^1H$ RDCs data to eq 3a.

We employ a two-stage procedure when fitting the positional distribution of each amide proton to the 10 RDCs (N_i-H_i and $C'_{i-1}-H_i$ RDCs for each of the five GB3 mutants): First, the axially symmetric internal motion model is applied, assuming uniform $\sigma_u = \sigma_v$ values for all 45 residues included in the analysis. Whereas in this model the internal motion would not be separable from a change in bond length if only N−H RDCs were considered, the two-bond C′−H^N RDCs lift this degeneracy because the N−H bond length also impacts the C′−H^N vector orientations. Second, after a global value for r_{NH}^{eq} has been determined, r_{NH}^{eq} is fixed at this value and the constraint $\sigma_u = \sigma_v$ is removed. This then makes it possible to evaluate the N−H motional asymmetry by simultaneously fitting σ_u and σ_v to the five sets of C′−H^N and $^{15}N-^1H$ RDCs. Below, we describe these two stages in more detail.

In the first stage, the amide protons of the previously determined GB3 structure (PDB entry 2OED)[15] are replaced by axially symmetric Gaussian distributions centered around the N−H vector orientations obtained from the above iterative DIDC analysis of the very high precision RDC data. Since the orientation of the molecular frame obtained from iterative DIDC analysis is arbitrary, it first needs to be overlaid with that of the experimental 2OED structure. For this purpose, the $^{13}C'-^{13}C^\alpha$ RDCs are best-fit to the 2OED coordinates. Then, using all 45 DIDC-accepted N−H vectors, the optimal value of $<r_{NH}>$ follows from a one-dimensional search: For each value of $<r_{NH}>$, an axially symmetric Gaussian distribution of 33 635 N−H vectors is generated where the width is defined by S'_{NH}, obtained from $S'_{NH}\ <r_{NH}^{-3}> = 0.872$ Å^{-3}. The $\rho_1(r)$ distribution in this ensemble (and thereby $<r_{NH}>$, as well as r_{eff} and r_{NH}^{eq}, cf. eq 4) is described by standard quantum statistical mechanics (Supporting Information, eq S24), and a singular value decomposition (SVD) fit of the C′−H^N RDCs to this ensemble yields χ^2.

The $<r_{NH}>$ obtained from this search (Figure 1B), 1.023 Å, is in excellent agreement with neutron scattering data (Table 1) and also agrees remarkably well with a prior NMR analysis of less precise RDCs, which neglected the effect of N−H bond stretching and libration but used a three-parameter model for motions of rigid peptide planes.[13] The $<S'_{NH}> = 0.910$ value we obtain when using $<r_{NH}> = 1.023$ Å is smaller than the 0.938 value expected for zero-point motion alone (Table 1). Backbone motions other than those affecting the reference C$^\alpha$−C′ vectors are responsible for this difference. For example, γ motions around the C$^\alpha$−C$^\alpha$ axis[11−14] have little impact on the C$^\alpha$−C′ axis due to the small angle (~20°) between these axes, but strongly affect the N−H vectors, oriented nearly perpendicular to $C^\alpha_{i-1}-C^\alpha_i$.

Table 2. Average Dynamics Parameters of Individual Secondary Structure Fragments[a]

	β1	β2	β3	β4	α1
S'_{NH}[b]	0.923 ± 0.029	0.867 ± 0.032	0.924 ± 0.025	0.931 ± 0.011	0.900 ± 0.026
η_{NH}	0.016 ± 0.026	0.048 ± 0.026	0.024 ± 0.020	0.017 ± 0.011	0.025 ± 0.045
in-plane[c]	8 ± 3°	9 ± 1°	7 ± 3°	7 ± 1°	9 ± 5°
out-of plane [c]	10 ± 1°	15 ± 3°	11 ± 2°	10 ± 1°	12 ± 3°

[a] Extracted from five independent alignment N−H and C′−H^N RDCs fittings, comprising residues Y3−N8 (β1), K13−K19 (β2), E42-D46 (β3), T51-T55 (β4), and (A23-D36) α1; values include the effect of zero-point motions. [b] The error is the standard deviation within each element of secondary structure, obtained from standard jackknife analysis. [c] Root-mean-square amplitude of motion.

In the second stage, we explore the motional asymmetry of the N−H bond vector. Using the $\rho_1(r)$ function that corresponds to $<r_{NH}> = 1.023$ Å, the N−H and C′−H^N RDCs are fit to ensembles of structures where the two parameters S'_{NH} and η (asymmetry of motion parameter)[5,6] define the ellipsoidal cone-shaped Gaussian distributions of the proton positions, corresponding to independent σ_u and σ_v values (eq 3b). Rather than fitting individual S'_{NH} and η values for each residue, which is adversely affected by error propagation, N−H vectors in each of the four β strands as well as the single α-helix are treated as having the same S'_{NH} and η, and a two-dimensional grid search for the S'_{NH} and η values yielding the best fit to C′−H^N and N−H RDCs is carried out. Results are summarized in Table 2. The principal axes of the ellipsoidal cones are defined such that positive η values correspond to out-of-plane motion being larger than in-plane motion, as is found for all five fragments. To provide a more intuitive picture of these motions, the corresponding rms angular fluctuations are also listed in Table 2. In agreement with earlier analyses,[5,12,16] we find the largest motional amplitude in solvent-exposed Igg-recognition strand β2.

The value $<r_{NH}> = 1.023$ Å obtained from our study is equivalent to an effective bond length $r_{eff} = 1.015$ Å. This is considerably smaller than the $r_{eff} = 1.041$ Å value obtained when a rigid model is used to relate $^{15}N-^{1}H$ RDCs to $^{13}C^{\alpha}-^{13}C'$ and $^{13}C'-^{15}N$ RDCs, where both the zero-point librations and the impact of larger angular fluctuations about the $C^{\alpha}-C^{\alpha}$ axis (γ motions) are absorbed into r_{eff}. For analysis of ^{15}N NMR relaxation data, the nearly uniform sub-picosecond zero-point N−H librations simply scale both the ^{15}N R_1 and R_2 relaxation rates by a factor of 0.863,[2] and their effect on the ^{15}N chemical shift anisotropy (CSA) tensor is already absorbed in CSA values observed by solid state or liquid crystal NMR. They therefore may be factored out by increasing r_{eff} by 2.5% to 1.040 Å. Note that variations in hydrogen bonding causes site-to-site deviations from this value as large as ±0.005 Å (Supporting Information, Figure S5). The effect of γ motions should not be absorbed into r_{eff} as such motions are relaxation active and can vary residue by residue.

The RDCs used in the present study were obtained for perdeuterated mutants of GB3, yielding narrower line widths and removal of very small systematic RDC errors resulting from cross-correlated relaxation.[17] The new RDCs are therefore more accurate than values used in prior GB3 studies.[5,12] Analysis of these new RDC values in terms of N−H order parameters, in the exact same manner as previously carried out for the protonated GB3 mutants,[5] yields values that lack the $S > 1$ outliers obtained previously. Excepting the three highly mobile residues K10-L12, RDC-derived S values agree remarkably well (pairwise rmsd 0.03) with those obtained from ^{15}N relaxation[16] (Supporting Information, Figure S4).

Acknowledgment. This work was supported in part by the Intramural Research Program of the NIDDK, NIH, and by the Intramural AIDS-Targeted Antiviral Program of the Office of the Director, NIH. We thank D.A. Torchia for useful discussions.

Supporting Information Available: Theoretical analysis of positional distribution of H^N under zero-point motion; calculated impact of H-bonding on r_{NH}^{eq}; plot of experimental $^3J_{HNH\alpha}$ vs DIDC-derived dihedral angle; tables with RDC values. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Kvick, A.; Alkaraghouli, A. R.; Koetzle, T. F. *Acta Crystallogr., B* **1977**, *33*, 3796–3801.
(2) Case, D. A. *J. Biomol. NMR* **1999**, *15*, 95–102.
(3) Guo, H.; Karplus, M. *J. Phys. Chem.* **1992**, *96*, 7273–7287.
(4) Yao, L. S.; Bax, A. *J. Am. Chem. Soc.* **2007**, *129*, 11326–11327.
(5) Yao, L.; Vogeli, B.; Torchia, D. A.; Bax, A. *J. Phys. Chem. B* **2008**, *112*, 6045–6056.
(6) Tolman, J. R. *J. Am. Chem. Soc.* **2002**, *124*, 12020–12030.
(7) Lakomek, N. A.; Carlomagno, T.; Becker, S.; Griesinger, C.; Meiler, J. *J. Biomol. NMR* **2006**, *34*, 101–115.
(8) Lange, O. F.; Lakomek, N. A.; Fares, C.; Schroder, G. F.; Walter, K. F. A.; Becker, S.; Meiler, J.; Grubmuller, H.; Griesinger, C.; de Groot, B. L. *Science* **2008**, *320*, 1471–1475.
(9) Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 12334–12341.
(10) Engh, R. A.; Huber, R. *Acta Crystallogr., A* **1991**, *47*, 392–400.
(11) Lienin, S. F.; Bremi, T.; Brutscher, B.; Bruschweiler, R.; Ernst, R. R. *J. Am. Chem. Soc.* **1998**, *120*, 9870–9879.
(12) 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.
(13) Bernado, P.; Blackledge, M. *J. Am. Chem. Soc.* **2004**, *126*, 4907–4920.
(14) Bremi, T.; Bruschweiler, R. *J. Am. Chem. Soc.* **1997**, *119*, 6672–6673.
(15) Ulmer, T. S.; Ramirez, B. E.; Delaglio, F.; Bax, A. *J. Am. Chem. Soc.* **2003**, *125*, 9179–9191.
(16) Hall, J. B.; Fushman, D. *J. Biomol. NMR* **2003**, *27*, 261–275.
(17) de Alba, E.; Tjandra, N. *J. Magn. Reson.* **2006**, *183*, 160–165.

JA805654F