Published in final edited form as:

J Am Chem Soc. 2008 December 10; 130(49): 16518–16520. doi:10.1021/ja805654f.

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

Abstract

The N-H bond length in backbone peptide groups of the protein GB3 has been studied by liquid crystal NMR, using five structurally conserved mutants of this protein. In the absence of additional information, the impact of dynamic fluctuations of the N-H vector orientation on the ^{15}N - ^1H dipolar interaction cannot be separated from a change in N-H bond length. However, a change in N-H bond

E-mail: E-mail: bax@nih.gov.

Supporting Information Available: Theoretical analysis of positional distribution of H^{N} under zero-point motion; calculated impact of H-bonding on $r_{\text{NH}}^{\text{eq}}$; plot of experimental $^3J_{\text{HNH}\alpha}$ vs DIDC-derived dihedral angle; tables with RDC values.

length directly impacts the orientation of C'-H vectors in the peptide group, and simultaneous analysis of $^{13}\text{C}'\text{-}^1\text{H}^{\text{N}}$ and $^{15}\text{N}\text{-}^1\text{H}^{\text{N}}$ residual dipolar couplings, measured under five different alignment orientations, permits model-free determination of the average equilibrium N-H bond length in GB3, yielding $r_{\text{NH}}^{\text{eq}} = 1.008 \pm 0.006 \text{ \AA}$. Anharmonicity of the bond stretching results in a slightly longer time-averaged bond length $\langle r_{\text{NH}} \rangle = 1.015 \pm 0.006 \text{ \AA}$, and an effective bond length $r_{\text{eff}} = \langle r_{\text{NH}}^{-3} \rangle^{-1/3} = 1.023 \pm 0.006 \text{ \AA}$ pertinent for NMR relaxation analysis, not including the impact of zero-point or other angular fluctuations in N-H orientation. Using a reference frame defined by the backbone C $^{\alpha}$ -C' vectors of the protein, angular fluctuations for N-H vectors in elements of secondary structure are approximately 1.5 fold larger for out-of-plane fluctuations than motions within the peptide plane and not much larger than anticipated on the basis of quantum mechanical analysis of their zero-point librations.

Quantitative interpretation of ^{15}N NMR relaxation rates in terms of motion hinges on the value of the $^{15}\text{N}\text{-}^1\text{H}$ internuclear distance. Neutron diffraction of small model peptides suggested a value of 1.020-1.024 $\text{\AA}$.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 $^{15}\text{N}\text{-}^1\text{H}$ 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 model-free 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 sub-ms 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 $\text{\AA}$ decrease in $r_{\text{NH}}^{\text{eq}}$ when $r_{\text{H}\cdots\text{O}}$ increases from 1.8 to 2.5 $\text{\AA}$, 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 (Fig. 1A): The equilibrium bond length, $r_{\text{NH}}^{\text{eq}}$, corresponds to the internuclear distance of lowest energy. The average $\langle r_{\text{NH}} \rangle$ distance includes averaging over the lowest vibrational bond stretching mode, which due to anharmonicity of the potential is calculated to be ca 0.015 $\text{\AA}$ longer than $r_{\text{NH}}^{\text{eq}}$. From an NMR relaxation perspective, where the strength of the dipolar coupling scales with r_{NH}^{-3} , the pertinent bond length equals $r_{\text{eff}} = \langle r_{\text{NH}}^{-3} \rangle^{-1/3}$. 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 $^{15}\text{N}\text{-}^1\text{H}$ dipolar interaction, it has remained difficult to establish experimentally precise baseline values for $\langle r_{\text{NH}} \rangle$ and r_{eff} . Here, we report concerted measurement and analysis of $^{13}\text{C}'\text{-}^{13}\text{C}^{\alpha}$, $^{13}\text{C}'\text{-}^1\text{H}^{\text{N}}$ and $^{15}\text{N}\text{-}^1\text{H}^{\text{N}}$ 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-His₆ and K19EK4A-N-His₆, previously were shown to map the five-dimensional alignment space when dissolved in a liquid crystalline Pf1 suspension.⁴ Using an improved stability, iterative implementation⁵ of the DDC method,⁶ alignment tensors are derived independent of a starting structure, *i.e.* 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 DDC as highly dynamic, and five residues for which resonance overlap interfered with accurate measurement of $^{13}\text{C}'\text{-}^{13}\text{C}^{\alpha}$, ^{13}C

$^1\text{H}^{\text{N}}$ or ^{15}N - $^1\text{H}^{\text{N}}$ RDCs in one or more of the five mutants, leaves 45 sets of three RDCs per mutant for analysis.

Assuming uniform $\text{C}'\text{-C}^\alpha$ 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 (Fig. 2)

$$\frac{\text{Amat}(\text{C}'\text{C}^\alpha)}{\text{Amat}(\text{NH})} = \frac{\gamma_{\text{C}}\gamma_{\text{C}'}S_{\text{C}'\text{C}^\alpha}\left\langle\frac{1}{r_{\text{C}'\text{C}^\alpha}^3}\right\rangle}{\gamma_{\text{N}}\gamma_{\text{H}}S_{\text{NH}}\left\langle\frac{1}{r_{\text{NH}}^3}\right\rangle} = 0.202 \quad (1)$$

represents the ratio of the strengths of the $^{13}\text{C}'$ - $^{13}\text{C}^\alpha$ and ^{15}N - $^1\text{H}^{\text{N}}$ dipolar interaction constants, which falls very close to the value previously reported⁹ using ubiquitin RDCs and its X-ray structure. Besides the assumption of uniform $\text{C}'\text{-C}^\alpha$ 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 $S_{\text{C}'\text{C}^\alpha}$ order parameters. Second, angular internal motion of $\text{C}'\text{-C}^\alpha$ and N-H^{N} vectors, which determines S_{NH} and $S_{\text{C}'\text{C}^\alpha}$, is assumed to be axially symmetric. Using simulated data, we previously have shown⁵ 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 Fig. 2 also confirms the validity of these assumptions.

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

$$S'_{\text{NH}} \langle r_{\text{NH}}^{-3} \rangle = 0.872 \pm 0.006 \text{ \AA}^{-3} \quad (2)$$

is then obtained from the Amat ratios (Fig. 2). Note that S'_{NH} defines the order of the N-H vectors relative to the average order parameter of the $\text{C}^\alpha\text{-C}'$ vectors, which is taken as a proxy for the motion of the entire backbone, including $\text{C}'\text{-N}$ and N-C^α vectors. Fluctuations of the peptide plane about the $\text{C}^\alpha_{i-1}\text{-C}^\alpha_i$ vector, so-called γ motions,¹¹⁻¹⁴ have a larger impact on the orientations of $\text{N}_i\text{-H}_i$ vectors than on the backbone $\text{C}^\alpha_{i-1}\text{-C}'_{i-1}$, $\text{C}'_{i-1}\text{-N}_i$ and $\text{N}_i\text{-C}^\alpha_i$ 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:

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

where $\rho_1(r)$ denotes bond stretching (Supporting Info. Eq S25) and

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

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} . Based on 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:

$$r_{\text{eff}} \equiv \langle r_{\text{NH}}^{-3} \rangle^{-1/3} \approx r_{\text{eq}} + 0.007 \text{ \AA} = \langle r_{\text{NH}} \rangle - 0.009 \text{ \AA} \quad (4)$$

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 $^3J_{\text{HNH}\alpha}$ Karplus curve (rmsd 0.33 Hz; Supporting Information). Three independent parameters σ_u , σ_v and $\langle r_{\text{NH}} \rangle$ are then determined by fitting C'-H^N and ^{15}N - ^1H RDCs data to eq 3.

We employ a two-stage procedure when fitting the positional distribution of each amide proton to the ten RDCs ($\text{N}_i\text{-H}_i$ and $\text{C}'_{i-1}\text{-H}_i$ 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 $r_{\text{NH}}^{\text{eq}}$ has been determined, $r_{\text{NH}}^{\text{eq}}$ 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 $^{13}\text{C}'$ - $^{13}\text{C}^\alpha$ RDCs are best-fit to the 2OED coordinates. Then, using all 45 DIDC-accepted N-H vectors, the optimal value of $\langle r_{\text{NH}} \rangle$ follows from a one-dimensional search: For each value of $\langle r_{\text{NH}} \rangle$, an axially symmetric Gaussian distribution of 33,635 N-H vectors is generated where the width is defined by S'_{NH} , obtained from $S'_{\text{NH}} \langle r_{\text{NH}}^{-3} \rangle = 0.872 \text{ \AA}^{-3}$. The $\rho_1(r)$ distribution in this ensemble (and thereby $\langle r_{\text{NH}} \rangle$, as well as r_{eff} and $r_{\text{NH}}^{\text{eq}}$, 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 $\langle r_{\text{NH}} \rangle$ obtained from this search (Fig. 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 $\langle S'_{\text{NH}} \rangle = 0.910$ value we obtain when using $\langle r_{\text{NH}} \rangle = 1.023 \text{ \AA}$ is smaller than the 0.938 value expected for zero-point motion alone (Table 1). Backbone motions other than those affecting the reference C^α-C' vectors are responsible for this difference. For example, γ motions around the C^α-C^α axis¹¹⁻¹⁴ have little

impact on the C^α-C' axis due to the small angle (~20°) between these axes, but strongly affect the N-H vectors, oriented nearly perpendicular to C^α_{i-1}-C^α_i.

In the second stage, we explore the motional asymmetry of the N-H bond vector. Using the $\rho_1(r)$ function that corresponds to $\langle r_{\text{NH}} \rangle = 1.023 \text{ \AA}$, 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 $\beta 2$.

The value $\langle r_{\text{NH}} \rangle = 1.023 \text{ \AA}$ obtained from our study is equivalent to an effective bond length $r_{\text{eff}} = 1.015 \text{ \AA}$. This is considerably smaller than the $r_{\text{eff}} = 1.041 \text{ \AA}$ value obtained when a rigid model is used to relate ¹⁵N-¹H RDCs to ¹³C^α-¹³C' and ¹³C'-¹⁵N RDCs, where both the zero-point librations and the impact of larger angular fluctuations about the C^α-C^α axis (γ motions) are absorbed into r_{eff} . For analysis of ¹⁵N NMR relaxation data, factoring out only the nearly uniform effect of zero-point N-H librations² increases r_{eff} by 2.2% to 1.037 $\text{\AA}$. This latter value is appropriate when characterizing backbone dynamics from ¹⁵N relaxation data, but note that site-to-site deviations from this value can be as large as $\pm 0.005 \text{ \AA}$ (Supporting Information Fig. 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 ¹⁵N relaxation¹⁶ (Supporting Information Fig. S4).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

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.

References

1. Kwick A, Alkaraghoul AR, Koetzle TF. Acta Crystallogr Sect B-Struct Commun 1977;33:3796–3801.
2. Case DA. J Biomol NMR 1999;15:95–102. [PubMed: 10605083]
3. Guo H, Karplus M. J Phys Chem 1992;96:7273–7287.
4. Yao LS, Bax A. J Am Chem Soc 2007;129:11326–11327. [PubMed: 17718572]

5. Yao L, Vogeli B, Torchia DA, Bax A. *J Phys Chem B* 2008;112:6045–6056. [PubMed: 18358021]
6. Tolman JR. *J Am Chem Soc* 2002;124:12020–12030. [PubMed: 12358549]
7. Lakomek NA, Carlomagno T, Becker S, Griesinger C, Meiler J. *J Biomol NMR* 2006;34:101–115. [PubMed: 16518697]
8. Lange OF, Lakomek NA, Fares C, Schroder GF, Walter KFA, Becker S, Meiler J, Grubmuller H, Griesinger C, de Groot BL. *Science* 2008;320:1471–1475. [PubMed: 18556554]
9. Ottiger M, Bax A. *J Am Chem Soc* 1998;120:12334–12341.
10. Engh RA, Huber R. *Acta Crystallogr A* 1991;47:392–400.
11. Lienin SF, Bremi T, Brutscher B, Bruschweiler R, Ernst RR. *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. [PubMed: 16172390]
13. Bernado P, Blackledge M. *J Am Chem Soc* 2004;126:4907–4920. [PubMed: 15080696]
14. Bremi T, Bruschweiler R. *J Am Chem Soc* 1997;119:6672–6673.
15. Ulmer TS, Ramirez BE, Delaglio F, Bax A. *J Am Chem Soc* 2003;125:9179–9191. [PubMed: 15369375]
16. Hall JB, Fushman D. *J Biomol NMR* 2003;27:261–275. [PubMed: 12975584]
17. de Alba E, Tjandra N. *J Magn Reson* 2006;183:160–165. [PubMed: 16949845]

Figure 1.

Amide N-H bond length in peptide groups. (A) Bond energy *versus* r_{NH} bond length, and definitions of $r_{\text{NH}}^{\text{eq}}$, $\langle r_{\text{NH}} \rangle$, and r_{eff} ($\langle r_{\text{NH}}^{-3} \rangle^{-1/3}$). (B) Total χ^2 fitting error of $\text{C}^\alpha\text{-H}^{\text{N}}$ RDCs

to GB3 structure (see text) as a function of $r_{\text{NH}}^{\text{eq}}$, with $\chi^2 = \sum_i \sum_j (\text{RDC}_{ij}^{\text{pred}} - \text{RDC}_{ij}^{\text{exp}})^2$, where i is the alignment index and j is the residue index. Values for $\langle r_{\text{NH}} \rangle$ and r_{eff} follow directly from $r_{\text{NH}}^{\text{eq}}$ (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 DIDC⁵ from C ^{α} -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.

Table 1
Average dynamics parameters and amide N-H bond lengths (Å).

	$r_{\text{NH}}^{\text{eqd}}$	r_{eff}^b	$\langle r_{\text{NH}} \rangle^c$	S'_{NH}	σ_{tr}^g	$S'_{\text{C'HN}}^h$
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 $r_{\text{eff}} = \langle r_{\text{NH}}^{-3} \rangle^{-1/3}$, where $\langle \dots \rangle$ denotes the bond stretching average;
- ^c $\langle r \rangle$ is the stretching-averaged bond length with the difference between r_{eq} and $\langle r \rangle$ resulting from anharmonicity;
- ^d Uncertainty in the averaged value based on jackknife analysis;
- ^e $\langle r_{\text{NH}} \rangle$ from neutron scattering¹;
- ^f Value for N-H zero point libration in N-methylacetamide, obtained from MP2 quantum calculation, divided by the corresponding $S_{\text{C'HN}}$ value of 0.990.
- ^g $\sigma_{\text{tr}} = (\sigma_u^2 + \sigma_v^2)^{1/2}$.
- ^h $S'_{\text{C'HN}}$ is the C'-H^N order parameter relative to the C^α-C'-N backbone.