Hydrogen Bond Geometry and $^2H/^1H$ Fractionation in Proteins

Peter M. Bowers and Rachel E. Klevit*

Contribution from the Department of Biochemistry, Box 357742, University of Washington, Seattle, Washington 98195-7742

Received July 30, 1999. Revised Manuscript Received November 18, 1999

Abstract: Measurement of $^2H/^1H$ equilibrium exchange (Φ) is commonly applied as an indirect probe of hydrogen bonds in small molecules. To expand the limited set of protein Φ measurements and to gain insight into the putative correlation between low Φ and hydrogen bond strength, we report measurements for two proteins, *src* SH3 and ubiquitin. A hydrogen bond network in *Gallus gallus src* SH3 domain, involving residues 30, 47, and 50, contributes ~2.0 kcal/mol to native state stability and provides an excellent system in which to test the purported connection between hydrogen bonding networks and protium enrichment. All observed sites in *src* SH3 had fractionation factors greater than unity. In particular, the backbone protons associated with the hydrogen bond network at residues Glu30, Ser47, and Thr50 all exhibit moderate deuterium enrichment. Measured fractionation values in ubiquitin range from $\Phi = 1.52$ for the amide of residue Ser20 to a remarkably low $\Phi = 0.29$ for the amide of Thr9. The majority (75%) of backbone amides fall between $0.9 < \Phi < 1.3$, with an average $\Phi = 1.07$, closely matching the average $\Phi = 1.11$ previously determined by another triple resonance method, and consistent with the fractionation observed in other weak hydrogen bonding amide systems. A survey of protein sites exhibiting low fractionation show a conservation of hydrogen bonding geometry. Our data, in combination with other studies, suggest that $^1H/^2H$ fractionation at protein *backbone* amides is a product of the complex three dimensional and static protein hydrogen bonding environment that restricts or enhances specific vibrational modes and is largely independent of hydrogen bonding strength.

A myriad of NMR techniques have been developed to examine hydrogen (H) bonds in proteins, essential for catalytic function and stabilization of the native state. Protection from chemical exchange with solvent, backbone $^1J_{C'N}$ scalar coupling, ^2H quadrupolar coupling, and isotropic and anisotropic chemical shift data provide indirect information regarding H-bond character.[1−5] The detection of $^{3h}J_{NiC'j}$- scalar coupling in proteins and $^2J_{NN}$-scalar coupling between DNA Watson−Crick base-pair partners provided the first direct evidence, independently demonstrated by Compton-scattering experiments, for the partial covalent character of H-bonds.[6−8] A consistent picture, correlating (^NJ) scalar coupling, (δ) chemical shift, H-bond length, and H-bond strength in proteins has emerged.[7,9]

Measurement of $^2H/^1H$ equilibrium exchange is commonly applied as an indirect probe of H-bonds in small molecules and has been reviewed in detail.[1,10] An equilibrium constant of isotope exchange is known as a fractionation factor (Φ).

$$X{-}H + D_{solvent} \leftrightarrow X{-}D + H_{solvent}$$

$$\Phi = ([X{-}D]/[X{-}H])_{solute} / ([D]/[H])_{solvent}$$

A $\Phi < 1.0$ corresponds to an equilibrium preference for protium, relative to solvent content, whereas a $\Phi > 1.0$ reflects an equilibrium preference for deuterium. The relative vibrational energies of the deuterium (X−D) and protium (X−H) bonds (a function of the mass of the nuclei) in the solvent and solute determine the extent of isotope enrichment. If the difference between the protium and deuterium zero-point vibrational energies is greater in the solute than in the solvent, deuterium will enrich at the solute site to minimize the total energy of the system. Interactions that alter the vibrational energy of solute X−H/D bonds affect the extent of isotopic enrichment at a site. Formation of strong H-bonds slows in-line stretching motions, leading to a characteristic enrichment for protium ($\Phi < 1$). The primary effect of weak H-bonds ($\Delta G < 2$ kcal·mol^{-1}) is to restrict off-line bending motions, leading to moderate enrichment for deuterium at a solute site. Steric interactions that restrict specific vibrational modes are observed to have similar effects on isotope enrichment.[11]

Fractionation factors have been determined in several proteins by measuring peak volumes of NMR resonances for a series of matched samples dissolved in buffers with defined $^2H_2O/^1H_2O$ ratios.[12−16] Although the majority of measured protein sites do

(1) Hvidt, A.; Nielsen, S. O. *Adv. Prot. Chem.* **1966**, *21*, 287−385.
(2) Juranic, N.; Ilich, P. K.; Macura, S. *J. Am. Chem. Soc.* **1995**, *117*, 405−410.
(3) LiWang, A. C.; Bax, A. *J. Magn. Reson.* **1997**, *127*, 54−64.
(4) Tessari, M.; Vis, H.; Boelens, F.; Kaptein, R.; Vuister, G. W. *J. Am. Chem. Soc.* **1997**, *119*, 8985−8990.
(5) Tjandra, N.; Bax, A. *J. Am. Chem. Soc.* **1997**, *119*, 8076−8082.
(6) Cordier, F.; Grzesiek, S. *J. Am. Chem. Soc.* **1999**, *121*, 1601−1602.
(7) Cornilescu, G.; Hu, J.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 2949−2950.
(8) Dingley, A. J.; Grzesiek, S. *J. Am. Chem. Soc.* **1998**, *120*, 8293−8297.
(9) Harris, T. K.; Mildvan, A. S. *Proteins: Struct., Funct., and Genet.* **1999**, *35*, 275−282.
(10) Hibbert, F.; Emsley, J. *Hydrogen bonding and chemical reactivity*; Academic Press: New York, 1990; Vol. 26.
(11) Guo, H.; Kresge, A. J. *J. Chem. Soc. Perk.* **1997**, *2*, 295−298.
(12) Loh, S. N.; Markley, J. L. In *Measurement of amide hydrogen D/H fractionation factors in proteins by NMR spectroscopy*; Angeletti, R., Ed.; Academic Press: San Diego, 1993; p 517.
(13) Loh, S. N.; Markley, J. L. *Biochemistry* **1994**, *33*, 1029−1036.
(14) Bowers, P. M.; Klevit, R. E. *Nat. Struct. Biol.* **1996**, *3*, 522−31.
(15) LiWang, A. C.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 12864−12865.
(16) Khare, D.; Alexander, P.; Orban, J. *Biochemistry* **1999**, *38*, 3918−3925.

10.1021/ja9927016 CCC: $19.00
Published on Web 02/01/2000

not exhibit a significant isotopic preference, certain backbone and side chain sites have fractionation factors that differ significantly from unity ($0.3 < \Phi < 1.5$). Extensive studies of protein catalytic triads and small molecules have been performed in which protium enrichment has been shown to be a function of strong H-bond formation, leading to the suggestion that low *backbone* Φ may correlate with strong H-bonds.[17−19] Calorimetric analysis of model compounds and proteins, site-directed mutagenesis studies, and NMR measurements indicate that protein backbone amide H-bonds are weak ($0.0 < \Delta G < 1.5$ kcal·mol^{-1}·bond^{-1}), and the observation of low fractionation factors in proteins was consequently unexpected.[6,7,20−22]

To expand the limited set of protein Φ measurements and to gain insight into the putative correlation between low Φ and H-bonds in proteins, we report measurements for two proteins, *src* SH3 and ubiquitin. A H-bond network in *Gallus gallus src* SH3 domain, involving residues 30, 47, and 50, is crucial for native state stability,[23] contributing 1.8−2.5 kcal/mol to native state stability and providing an excellent system in which to test the purported connection between H-bonding networks and protium enrichment. All measured amide protons of *src* SH3 exhibit moderate enrichment for deuterium, as expected for weak H-bond formation (Figure 1). Although a majority of the sites in ubiquitin also exhibit enrichment for deuterium, a few sites exhibit significant enrichment for protium. Alternate explanations for the observation of protium enrichment in protein, including incomplete solvent T_1 relaxation coupled with rapid solvent exchange or dipolar relaxation, do not appear to be responsible for observations of low Φ, as previously demonstrated.[16] In sites exhibiting the lowest fractionation factors, the geometry of H-bond networks is conserved. Our data indicate that protium enrichment at the protein backbone amide may be a product of the complex three-dimensional and static protein H-bonding environment that restricts or enhances specific vibrational modes, independent of H-bonding strength and experimental artifacts. Complementary H-bond studies, including $^{3h}J_{NiC'j}$- scalar coupling measurements in ubiquitin and protein G, and the effect of steric interactions on specific vibrational modes and Φ in small molecules suggest that protium enrichment at backbone amide protons is not a simple function of H-bond strength.

Fractionation factors for ubiquitin have previously been determined using an indirect protocol that involves collection of a triple resonance HA(CA)CO NMR spectrum of a single protein sample dissolved in 50% $^2H_2O/^1H_2O$.[15] To compare values among the set of proteins investigated to date, we measured the Φ values in ubiquitin using the $^1H-^{15}N$ SE-HSQC protocol employed in all other studies.[24] Solution conditions were matched to the earlier ubiquitin study to allow direct comparison of the two NMR methods. A gradient water flip-back solvent suppression scheme was used, avoiding complications associated with direct detection of amide protons in fast chemical exchange with the solvent.[25] The interscan relaxation delay between transients was set to ≥2.0s (for both the ubiquitin and SH3 data sets) to allow complete protein T_1 relaxation.[15,16] A total of 45 ubiquitin amide resonances were well-resolved and clearly identified from previous chemical shift assignments. Measured fractionation values range from $\Phi = 1.52$ for the amide of residue Ser20 to a remarkably low $\Phi = 0.29$ for the amide of Thr9 (Figure 1). The majority (75%) of backbone amides fall between $0.9 < \Phi < 1.3$, with an average $\Phi = 1.07$, closely matching the average $\Phi = 1.11$ determined by the triple resonance method, and consistent with the fractionation observed in other weak H-bonding amide systems.[11] The close correspondence of the average values measured indicates that the effects of varying isotope composition effects on viscosity and/or nuclear relaxation across the sample set do not generate a

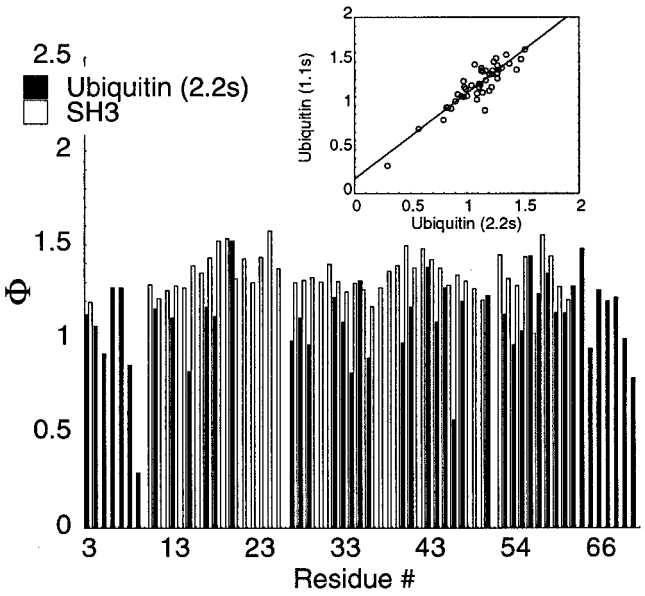

Figure 1. Plot of measured fractionation values in ubiquitin with a 2.2 s versus 1.1 s interscan relaxation delay. Final protein samples were 3-4 mM in a volume of 500 μL buffered in 30 mM Tris-d_{11} (Cambridge) and 30 mM deuterated (Cambridge)/0.1% acetate. Lyophilized ubiquitin was dissolved and divided into six equal parts, lyophilized, and dissolved again into solution with a fractional H_2O content of 0.0, 0.1, 0.4, 0.6, 0.80, and 0.9 and lyophilized. SH3 domain was divided equally into seven samples with fractional H_2O content of 0.0, 0.1, 0.3, 0.5, 0.6, 0.85, and 0.95 and lyophilized. Samples were redissolved into buffer with the same isotopic content, heated at 45 °C at pH 9.0 for 10 hours. The pH of the samples was then adjusted to pH 6.0 for the SH3 and pH 5.0 for ubiquitin (using a 1 M HCl/DCl solution matching the fractional solvent content of each sample). Care was taken to insure that protein concentrations did not vary over the range of samples, and that isotopic concentrations were accurate. Integrated methyl peak intensities from 1D spectra were found to vary by less than 2% across a sample set. A sensitivity enhanced $^1H-^{15}H$ HSQC spectrum of the 100% ^2H_2O sample for each protein was collected to confirm that exchange had reached equilibrium. All spectra were collected on a Bruker 750 MHz DMX spectrometer at 30 °C. 1D proton spectra with 8K complex points and SW of 7 KHz were collected for all samples utilizing WATERGATE water suppression. $^1H-^{15}N$ SE-HSQC 2D data sets were collected with $256(t_1) \times 2K$ complex points. Data sets were zero-filled to give $2K \times 2K$ matrices, weighted in both dimensions by a sine function shifted 90°, and baseline corrected in both dimensions using a second order polynomial function. NMRPipes was used for NMR data processing and manipulation.[28] The program PIPP was used for measurement of peak intensities and volumes.[29] Fractionation factors were determined as previously described.[14]

(17) Cleland, W. W.; Kreevoy, M. M. *Science* **1994**, *264*, 1887−1890.
(18) Kreevoy, M. M.; Liang, T. M. *J. Am. Chem. Soc.* **1980**, *102*, 3315−3322.
(19) Lin, J.; Westler, W. M.; Cleland, W. W.; Markley, J. L.; Frey, P. A. *Proc Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14664−8.
(20) Shirley, B. A.; Stanssens, P.; Hahn, U.; Pace, C. N. *Biochemistry* **1992**, *31*, 725−732.
(21) Wintrode, P. L.; Makhatadze, G. I.; Privalov, P. L. *Proteins: Struct., Funct., and Genet.* **1994**, *18*, 246−253.
(22) Cornilescu, G.; Ramirez, B. E.; Frank, M. K.; Clore, G. M.; Gronenborn, A. M.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 6275−6279.
(23) Grantcharova, V.; Riddle, D. S.; Santiago, J. V.; Baker, D. *Nat. Struct. Biol.* **1998**, *5*, 714−720.
(24) Palmer III, A. G.; Cavanagh, J.; Wright, P. E.; Rance, M. *J. Magn. Reson.* **1991**, *93*, 151.
(25) Sklenar, V.; Piotto, M.; Leppik, R.; Saudek, V. *J. Magn. Reson. A* **1993**, *102*, 241−245.

Table 1. Low Backbone Φ Values Obtained in Four NMR Fractionation Studies

PDB	H^N	CO	*F*	R_N (Å)	∠(NHO) (deg)	∠(COH) (deg)
1BR7	T9	T7	0.29/0.32	3.26	156	113
1BR7	T9	T9	0.29/0.32	2.83	96	78
1POH	S31	D69	0.74/0.63[a]	3.02	162	130
1POH	S31	S31	0.74/0.63[a]	2.87	97	74
1PGA	T49	D46	0.82	3.03	155	112
1PGA	T49	T49	0.82	3.02	89	76
1STN	T120	D77	0.30	2.89	152	122
1STN	T120	T120	0.30	2.82	93	71

[a] Indicates the side chain serine (−OγH) hydroxyl proton.

directional perturbation in Φ, as previously proposed.[15] Despite the correspondence in average fractionation values, there is no correlation between individual residue values obtained using the HA(CA)CO and HSQC methods. Several explanations for the striking discrepancy are possible. The triple resonance protocol requires multiple correction factors, determined from the X-ray structure of ubiquitin and solution NMR measurements of ubiquitin, that could affect the apparent Φ of sites differentially. Second, the triple resonance approach relies on measurements on a single sample. Errors or uncertainties in the isotopic ratio of the solvent in the single sample would affect the apparent Φ values, although this effect should produce a systematic error rather than affect individual values differentially.

Liwang and Bax have suggested that rapid solvent exchange at amide sites coupled with a relaxation delay shorter than that of the solvent could lead to an increase in signal intensity (decrease in the apparent Φ) proportional to the solvent content. This effect represents the most probable effect leading to anomalous observation of low fractionation factors utilizing the HSQC technique. To determine if incomplete T_1 relaxation was responsible for the observation of low fractionation factors at backbone amide sites, a second series of $^1H-^{15}N$ SE-HSQC spectra was collected for ubiquitin in which the interscan relaxation delay was set to 1.1 s. In direct contrast to the prediction, shortening the relaxation delay increased the apparent fractionation factor in all but one residue, Lys11. The average value in ubiquitin increases from $\Phi = 1.07$ to 1.21 with a decrease in the relaxation delay from 2.2 s to 1.1 s. The site with the lowest fractionation factor, Thr9, has a value of $\Phi = 0.32$, slightly larger than the value of $\Phi = 0.29$ obtained in the 2.2 s data set. A plot of fractionation factors at each residue in ubiquitin with the 1.1 s versus 2.2 s relaxation delay has a correlation coefficient of $R = 0.91$, demonstrating that Φ measurements are reproducible and that doubling the relaxation time has only small effects on observed fractionation values (Figure 1, inset). The results are not surprising: the flip-back solvent suppression scheme avoids excitation of the H_2O resonance, and few of the amide protons in ubiquitin exchange rapidly with the solvent under these solution conditions. No correlation between low fractionation factors and solvent exchange rates has been observed in previous studies.[14,16] Incomplete protein T_1 relaxation is likely to be responsible for the observed increase in Φ with a shorter relaxation delay. The effect of shortening the interscan delay on the observed Φ results from longer T_1 relaxation rates as the $^2H_2O/^1H_2O$ solvent ratio (and ubiquitin deuteration) increases across the sample set. We conclude that incomplete nuclear relaxation and chemical exchange with the solvent have only modest effects on Φ measurements using the current experimental approach and do not account for the observations of low Φ of backbone amide protons.[15,16]

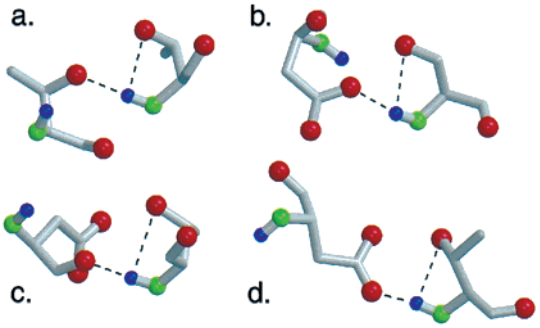

Figure 2. H-bond networks in as well (a) ubiquitin (1BR7) between residues Thr 7 and Thr 9, (b) *Escherichia coli* HPr (1POH) between residues Asp69 and Ser31, (c) the immunoglobulin G binding bomain of protein G (1PGA) between residues Asp46 and Thr49, and (d) in S. nuclease (1STN) between residues Thr120 and Asp77, identified by protium enrichment. Predicted H-bonds are indicated by dashed lines.

Calculations aimed at understanding the low Φ observed in proteins led to the proposal that H-bonding networks involving charged H-bond acceptors will exhibit low Φ.[26] Indeed, the sites of lowest Φ in two variants of HPr and in staphylococcal nuclease are in such H-bonding networks.[14,15] Residues Glu30, Ser47, and Thr50 in *src* SH3 are likewise involved in a charged H-bond network that contributes significantly to the stability of the SH3 domain,[23] providing an excellent opportunity to explore the purported connection between H-bonds networks and Φ. $^1H-^{15}N$ sensitivity-enhanced HSQC spectra of SH3 were collected for samples equilibrated over a range of $^2H_2O/^1H_2O$ concentrations. Of 63 backbone amide protons in SH3, 49 were well-resolved, and fractionation factors were determined (Figure 1). Despite the presence of fast-exchanging amide protons, all observed sites had fractionation factors greater than unity. All but three residues fractionation factors fall between $1.19 < \Phi < 1.57$, with an average $\Phi = 1.33$. The backbone protons associated with the H-bond network at residues Glu30, Ser47, and Thr50 all exhibit deuterium enrichment. These data suggest that H-bond networks alone are not sufficient for protium enrichment and confirm that SE-HSQC experimental technique does not contribute to the anomalous observation of low Φ in proteins.

The H-bond network in ubiquitin, involving residues Thr7 and Thr9 ($\Phi = 0.29$), is reminiscent of H-bond networks in other proteins identified by protium enrichment (Table 1). Residues 1 through 16 of human ubiquitin form a β-hairpin with a turn centered at residues Thr9 and Gly10. The amide proton of Thr9 ($\Phi = 0.29$) forms an H-bond with the γ-OH of Thr7 and a potential H-bond with the hydroxyl oxygen of its own side chain (Figure 2). The amide proton of Thr120 in *S.* nuclease, which forms an interstrand H-bond with the side chain of Asp77 and a potential H-bond with the side chain hydroxyl of its own side chain, has a fractionation factor of $\Phi = 0.30$.[13] In *ec*HPr, the amide proton of Ser31 ($\Phi = 0.74$) forms an interstrand H-bond with the side chain of Asp69 and a potential intraresidue H-bond with its own side chain γ-OH ($\Phi = 0.63$).[14] The backbone amide site in protein G with the lowest fractionation factor is involved in a similar H-bond network, in which the amide proton of Thr49 ($\Phi = 0.82$) H-bonds with the

(26) Edison, A. S.; Weinhold, F.; Markley, J. L. *J. Am. Chem. Soc.* **1995**, *117*, 9619−9624.
(27) Hennig, M.; Geierstanger, B. H. *J. Am. Chem. Soc.* **1999**, *121*, 5123−5126.
(28) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277−293.
(29) Garrett, D. S.; Powers, R.; Gronenborn, A. M.; Clore, G. M. *J. Magn. Reson.* **1991**, *95*, 214−220.

backbone carbonyl of Asp46, as well as forming a potential H-bond with its own side chain hydroxyl (Figure 2). No H-bonding networks that parallel this arrangement of H-bond acceptors and donor were observed in SH3. In each instance, the angles for potential H-bonds [$\angle$(N−H···O)] formed between the amide proton of a serine or threonine (Ser31, Thr120, Thr9, Thr49) with its own hydroxyl oxygen are conserved (Table 1). Likewise, the H-bonding angles to the interresidue H-bond acceptor (Ser31-Asp69, Thr120-Asp77, Thr9-Thr7, Thr49-Asp46) are also strictly conserved (Table 1). A pairwise comparison of the four nuclei (two oxygen H-bond acceptors and amide nitrogen and proton) in each H-bond network reveals a RMSD of 0.25 Å.

The H-bond distances (R_{NO}) (2.82−3.26 Å) in the networks identified by protium enrichment do not fall within the range expected for low-barrier (<2.55 Å) H-bonds.[9] $^{3h}J_{NiC'j}$- scalar couplings, which correlate with the length and the partial covalent character of the H-bond, have been measured in ubiquitin and Protein G.[22,27] The H-bond networks identified by protium enrichment were not found to have large $^{3h}J_{NiC'j}$- scalar couplings expected for strong or low-barrier H-bonds. In addition, no overall correlation between H-bond distance (determined from X-ray structures) and fractionation factor was observed, as reported in previous studies.[13,14]

We propose that the arrangement of H-bond donors and acceptors may enhance or hinder specific vibrational modes of the amide bond vector, enhancing protium enrichment at specific backbone sites. It was previously observed that slight protium enrichment in α-helical structure relative to β-sheet secondary structure likely stems from differences in the spatial organization of H-bond donors and acceptors associated with regular α-helical and β-sheet secondary structure, and is unrelated to the relative strength of these two classes of H-bonds.[13−16] Indeed, H-bond distances, $^{3h}J_{NiC'j}$- scalar coupling constants, and thermodynamic analysis of proteins universally suggest that β-sheet H-bonds are somewhat stronger, despite the greater enrichment for deuterium seen in all protein backbone Φ measurements.[9,21,22] It was previously observed that a systematic preference for protium at threonine residues detected in other studies may stem in part from their ability to form intraresidue out-of-plane side chain H-bonds (Figure 2).[13] Steric interactions, unique to protein structure, may also contribute to fractionation.[11] Our data, together with others, suggest that $^{1}H/^{2}H$ fractionation at protein *backbone* amides is a result of H-bonding geometry, which may account for the observation of exceptionally low fractionation values for some residues not involved in strong H-bonds.

JA9927016