Hydrogen Bonding and Equilibrium Protium−Deuterium Fractionation Factors in the Immunoglobulin G Binding Domain of Protein G†

Devesh Khare, Patrick Alexander, and John Orban*

Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, 9600 Gudelsky Drive, Rockville, Maryland 20850

Received November 13, 1998; Revised Manuscript Received February 10, 1999

ABSTRACT: Protium−deuterium fractionation factors (ϕ) were determined for more than 85% of the backbone amide protons in the IgG binding domains of protein G, G_{B1} and G_{B2}, from NMR spectra recorded over a range of H_2O/D_2O solvent ratios. Previous studies suggest a correlation between ϕ and hydrogen bond strength; amide and hydroxyl groups in strong hydrogen bonds accumulate protium ($\phi < 1$), while weak hydrogen bonds accumulate deuterium ($\phi > 1$). Our results show that the α-helical residues have slightly lower ϕ values (1.03 ± 0.05) than β-sheet residues (1.12 ± 0.07), on average. The lowest ϕ value obtained (0.65) does not involve a backbone amide but rather is for the interaction between two side chains, Y45 and D47. Fractionation factors for solvent-exposed residues are between the α-helix and β-sheet values, on average, and are close to those for random coil peptides. Further, the difference in ϕ_{av} between α-helix and solvent-exposed residues is small, suggesting that differences in hydrogen bond strength for intrachain hydrogen bonds and amide···water hydrogen bonds are also small. Overall, the enrichment for deuterium suggests that most backbone···backbone hydrogen bonds are weak.

The role of intrachain hydrogen bonding in the stabilization of proteins has been a long-standing question in biochemistry that is still actively debated (*1*−*5*). One major difficulty is evaluation of the energetic contribution of amide···water versus intramolecular amide···amide hydrogen bonds. Depending on individual interpretations of experiments, the free energy contribution per amide···amide hydrogen bond appears to vary from slightly destabilizing (*4*) to strongly stabilizing (*6*). The early work of Schellman (*7*) and Kauzmann (*8*) suggested that hydrophobic interactions play a dominant role in protein folding and stability and that hydrogen bonding provides only minimal stability. This view is generally supported in more recent articles (*1*, *4*). However, on the basis of mutagenesis of backbone···side chain and side chain···side chain hydrogen bonds in ribonuclease T1, Shirley et al. (*2*) concluded that hydrogen bonding and hydrophobic interactions make comparable contributions to protein stability. Other mutagenesis studies have led to similar conclusions (*9*). Protium−deuterium fractionation studies with proteins may potentially resolve some of these contradictory results.

The fractionation factor, ϕ, is defined as the equilibrium constant for the reaction

$$\mathrm{NH} + \mathrm{D} \leftrightarrows \mathrm{ND} + \mathrm{H} \qquad (1)$$

Therefore

$$\phi = ([\mathrm{ND}]/[\mathrm{NH}])_{\mathrm{protein}}/([\mathrm{D}]/[\mathrm{H}])_{\mathrm{solvent}} \qquad (2)$$

where NH is any exchangeable amide proton in the protein and D and H represent bulk deuterium and protium, respectively, in D_2O, H_2O, and HOD solvents. A ϕ value of >1 indicates a preference for deuterium over protium at the exchangeable site, while a ϕ value of <1 indicates a preference for protium over deuterium, relative to the solvent ratio. The physical basis of the measurement stems from the difference in vibrational energy between the N−H and N−D bonds at a particular site in the protein. Due to its larger mass, the vibrational energy minimum for deuterium is lower than that for protium in both the protein and solvent. If this energy difference is larger in the protein site than in the solvent, then the protein site will be enriched in deuterium ($\phi > 1$) to lower the total energy of the system. Theoretical calculations of fractionation factors in model systems suggest that ϕ is >1 for most hydrogen bonds in peptides but can be significantly <1 when hydrogen-bonded networks involving charged groups are present (*10*).

Fractionation studies with both small molecules and proteins indicate correlation between ϕ and hydrogen bond strength with low ϕ values (<1) observed for "strong" hydrogen bonds (*11*, *12*). Initial fractionation measurements in proteins focused on enzyme active sites, providing important mechanistic information about the role of low-barrier hydrogen bonds in catalysis (*13*−*15*). More recently, detailed measurements of backbone amide fractionation factors have been taken with the goal of understanding the role of hydrogen bonding in protein stability. Backbone amide fractionation studies have been carried out on three protein folds to date, staphylococcal nuclease (*16*, *17*), histidine-containing protein (HPr) (*18*), and ubiquitin (*19*). The ranges of main chain ϕ values obtained are 0.3−1.5 ($\phi_{av} = 0.84 \pm 0.19$), 0.70−1.22 ($\phi_{av} = 0.92 \pm 0.12$), and 1.01−

† Supported by NSF Grant MCB-92-19309.

* To whom correspondence should be addressed. Phone: (301) 738-6221. Fax: (301) 738-6255. E-mail: john@iris11.carb.nist.gov.

10.1021/bi9827114 CCC: $18.00
Published on Web 03/12/1999

1.21 (ϕ_{av} = 1.11 ± 0.05), respectively. Thus, there is variation in the range and average of main chain ϕ values in the three systems. These results also lead to quite different conclusions regarding the contribution of main chain hydrogen bonding to protein stability. It is apparent that more work is required to obtain a better understanding of the relation between fractionation factors and hydrogen bonding.

To expand further on these fractionation studies of proteins, we report here protium−deuterium fractionation factor measurements for the G_{B1} and G_{B2} immunoglobulin G (IgG) binding domains of streptococcal protein G. These two domains are small 56-amino acid folding units that differ in sequence by six residues, have similar thermal stabilities (*20*), and provide an excellent system for fractionation studies. Both G_{B1} and G_{B2} have been extensively characterized thermodynamically (*20*), kinetically (*21*), and structurally (*22*−*27*) and have well-resolved $^1H-^{15}N$ HSQC (heteronuclear single-quantum correlation) spectra. In addition, the hydrogen exchange properties of backbone amides in G_{B1} and G_{B2} are well-known (*28*, *29*). This is important, since fractionation measurements are carried out under equilibrium conditions where no further H−D exchange takes place. For both G_{B1} and G_{B2}, main chain NH ϕ values were obtained for more than 85% of the residues compared with 60% for staphylococcal nuclease, 67−79% for HPr, and 73% for ubiquitin. Our results are compared with fractionation data for these systems as well as with mutational and structural data from X-ray crystallography and NMR spectroscopy. The results presented here show an overall enrichment for deuterium in main chain NHs, suggesting that most backbone···backbone hydrogen bonds are weak. We further show that the difference in ϕ_{av} between helix and solvent-exposed residues is small, suggesting that differences in intramolecular amide···amide versus amide···water hydrogen bond strengths are also small. Finally, the lowest ϕ value observed is for an Asp···Tyr side chain···side chain hydrogen bond. On the basis of comparisons with the other three proteins studied to date, the lowest ϕ values observed are typically for hydrogen bonds between acidic residues and either the backbone or hydroxyl-containing side chains. Therefore, calculation of the total contribution of hydrogen bonding to stability by extrapolating from these types of interactions may overestimate this contribution.

MATERIALS AND METHODS

Sample Preparation. ^{15}N-labeled protein G samples were expressed and purified as described previously (*20*, *29*). Samples for fractionation experiments were prepared using the guidelines suggested by Loh and Markley (*16*). In particular, care was taken to ensure that (1) sample concentrations were uniform, (2) H_2O/D_2O ratios were accurate, and (3) H−D exchange had gone to completion. For each domain, a stock solution of protein in 100 mM sodium acetate-d_3/0.05% (w/v) sodium azide at pH 5.4 was used to prepare 11 samples with equal volumes, one of which was used as a control of exchange and pH conditions. After lyophilization, samples were exchanged in specific ratios of H_2O/D_2O solvent, relyophilized, and redissolved in their respective solvents (0.45 mL). Mole fractions of H_2O of 0.100, 0.200, 0.300, 0.400, 0.500, 0.600, 0.700, 0.800, 0.900, and 0.980 were used for both domains. The pH of each sample was adjusted to 7.4 ± 0.1 with a Sentron microelectrode using the appropriate ratios of NaOH/NaOD (1−2 μL), and samples were equilibrated for 60 h at 30 °C. Under these conditions, complete H−D exchange was achieved for the G_{B1} and G_{B2} control samples in 100% D_2O within 50 and 10 h, respectively, consistent with previous hydrogen exchange measurements taken on these domains (*29*). The pH was then lowered to 5.3 by adding the appropriate HCl/DCl ratios (5−6 μL). A protein concentration of 2.2 mM was used for G_{B1} and G_{B2} domains, and NMR samples were stored in a desiccator containing drierite.

NMR Spectroscopy. NMR spectra were recorded on a Bruker AMX-500 spectrometer and processed on Silicon Graphics workstations using FELIX 2.30 (Molecular Simulations, Inc.). One-dimensional ^1H NMR spectra were recorded at 4 °C using a WATERGATE solvent suppression scheme (*30*) with a sweep width of 6024 Hz, 4K complex points, and a relaxation delay of 5.0 s. Data were apodized with a 90°-shifted squared sine-bell window, Fourier transformed, and baseline corrected with a third-order polynomial. Two-dimensional $^1H-^{15}N$ HSQC spectra (*31*−*33*) were collected at 30 °C using a gradient water flip-back scheme for solvent suppression (*34*). Sine-bell *Z*-axis gradients were used with the following parameters: G_1 2.5 ms, 9 G cm^{-1}; G_2 1.0 ms, 9 G cm^{-1}; and G_3 0.4 ms, 4.2 G cm^{-1}. ^{15}N decoupling during t_2 was achieved using a GARP sequence (*35*). Spectral widths of 6024 and 1700 Hz were utilized in the ^1H and ^{15}N dimensions, respectively, with 512 t_1 increments of 2K complex points and 32 transients per increment. Data were processed using 90°-shifted squared sine-bell functions in both dimensions with zero-filling to give 2K × 2K matrixes. A fifth-order polynomial baseline correction was used in F_2. The sampling delay (d0) was set to half of a dwell time so that no baseline correction was necessary in F_1 (*36*). The total time for data collection per HSQC experiment was about 25 h.

Peak intensities in HSQC spectra as a function of D_2O/H_2O concentrations need to be interpreted with caution. As the D_2O/H_2O ratio is increased, amide proton T_1's get longer as cross-relaxation becomes less efficient. Average T_1's for main chain amide protons in protein G were determined using inversion−recovery experiments and ranged from 0.8 s in 98% H_2O to 1.4 s in 10% H_2O. This affect will attenuate peaks at the high D_2O/H_2O ratios more than at low D_2O/H_2O ratios if the relaxation delay is too short and will lead to an increase in the apparent ϕ value. An additional effect concerning the interaction between the solvent magnetization and the protein magnetization also needs to be considered, as pointed out recently by LiWang and Bax (*19*). In particular, the use of solvent presaturation will attenuate signals of NHs in rapid exchange with water. The degree of attenuation is proportional to the H_2O/D_2O ratio and this will lead to a decrease in ϕ. Through dipolar interactions, other NHs that are not necessarily solvent exposed will also be attenuated, with a corresponding decrease in ϕ. The use of a relaxation delay shorter than the T_1 of water (∼2−3 s) will have a similar effect on ϕ. To minimize these types of problems, we used a gradient water flip-back HSQC scheme which does not perturb the water magnetization (*34*) and a relatively long relaxation delay (∼5.3 s).

Determination of Fractionation Factors. Fractionation factors were determined from one- and two-dimensional NMR spectra using methods analogous to those described

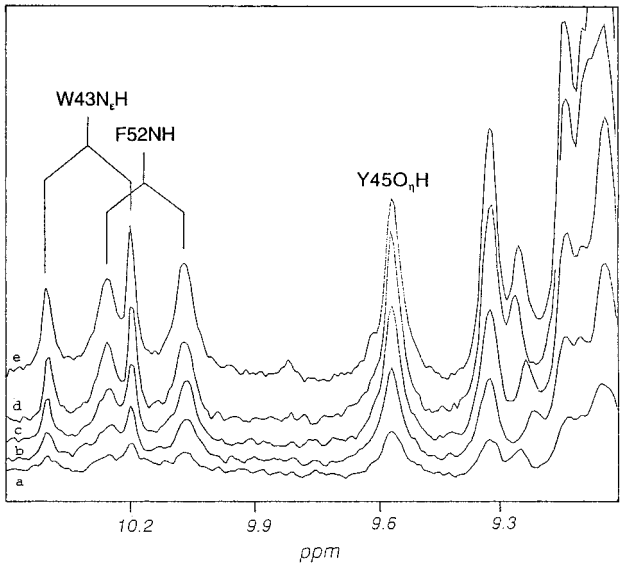

FIGURE 1: Stack plot of one-dimensional ^{1}H NMR spectra showing the Y45 O$_\eta$H resonance as a function of solvent content with the following H_2O/D_2O ratios: (a) 20/80, (b) 40/60, (c) 60/40, (d) 80/20, and (e) 98/2.

previously (*16*−*18*). In one-dimensional spectra, the peak area of the Y45 O$_\eta$H signal was normalized to that of nonexchangeable upfield-shifted aliphatic peaks to correct for small variations (typically <1−2%) in protein concentration from sample to sample. In two-dimensional spectra, cross-peak volumes were normalized internally using the sum of peak volumes for residues with no apparent isotopic preference. The normalization factors obtained from this approach were very similar to those from one-dimensional data and did not change the pattern of fractionation factors observed in G_{B1} and G_{B2}. The fractionation factor at each exchangeable site, ϕ, was obtained by a linear least-squares fit to the equation

$$1/y = C[\phi(1 - x)/x + 1] \quad (3)$$

where y is the peak integral, x is the mole fraction of H_2O in the sample, and C is a normalization parameter (equal to 1/peak volume at 100% intensity). Each $1/y$ point was weighted by y^2 in the fitting procedure.

RESULTS

All backbone amide ^{1}H and ^{15}N resonances have been assigned previously (*22*, *29*). The resolution in two-dimensional ^{1}H−^{15}N HSQC spectra is such that backbone amide ϕ values were obtained for more than 85% of the residues. In addition, fractionation data were obtained for the side chain hydroxyl group of Y45 from one-dimensional ^{1}H NMR spectra using the ^{15}N-labeled samples. Room-temperature one-dimensional ^{1}H NMR spectra recorded using gradient solvent suppression showed a broad resonance at ∼9.6 ppm which was not apparent in earlier spectra acquired using presaturation methods (*22*). At 4 °C, this signal sharpens, is well-resolved, and has NOE cross-peaks to Y45 C$_\epsilon$H and D47 C$_\alpha$H, consistent with assignment as the Y45 O$_\eta$H resonance. One-dimensional ^{1}H NMR spectra for a range of H_2O/D_2O concentrations are shown in Figure 1.

Representative plots of normalized peak volume/area versus fractional H_2O content are shown for three exchangeable sites in G_{B1} (Figure 2A). For the backbone NH of N35, there is no isotope preference and a straight line is observed. In contrast, F52 has a distinct preference for deuterium while Y45 O$_\eta$H is enriched with protium and curved lines on either side of the diagonal are observed. These data can also be represented as linear plots of $1/y$ versus $(1 - x)/x$ (eq 3), where the ϕ values can be obtained from the slope of the fit (Figure 2B).

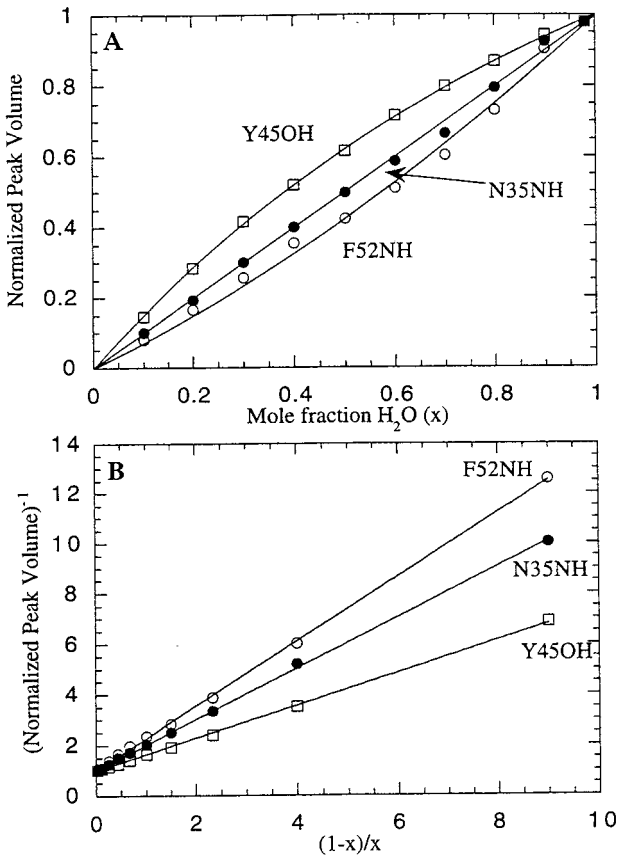

FIGURE 2: (A) Representative plots of normalized peak volume vs the mole fraction of H_2O (x) for Y45 O$_\eta$H (□) and the backbone amides of N35 (●) and F52 (○). (B) Linear plots in the form of eq 3 for the same data.

The fractionation factors are shown for each residue in Table 1. The lowest ϕ value observed in either domain is for the Y45 side chain hydroxyl group ($\phi_{av} = 0.65$). In contrast, the range of backbone amide ϕ values is 0.82−1.28 in G_{B1} and 0.93−1.36 in G_{B2} with an average value of 1.08 ± 0.08 over all residues. Of the 50 backbone amide sites for which fractionation factors could be obtained, 44 have ϕ_{av} values that are ≥1. On average, a slightly lower fractionation factor is observed for α-helical residues ($\phi_{av} = 1.03 \pm 0.05$) than for β-sheet residues ($\phi_{av} = 1.13 \pm 0.06$). This can also be seen from Figure 3 which shows the backbone NH fractionation factors mapped onto the tertiary structure of protein G.

The fractionation results are similar for the two domains which differ in only six amino acids. In this respect, the G_{B1} and G_{B2} data serve as near-duplicate sets of experiments. From these results, we estimate the error level in the determination of ϕ values to be <10%. Only two residues, I6V and V54, have ϕ values with a standard deviation of more than 0.10 between the two domains. Indeed, I6V is the only variant position with a significant difference in ϕ

Table 1: Fractionation Factors for G_{B1} and G_{B2}

residue	ϕB1	ϕB2	ϕ_{av} (±SD)	residue	ϕB1	ϕB2	ϕ_{av} (±SD)
M1[a]	–	–	–	V29A[d]	0.95	1.06	1.01 (0.08)
T2[a]	–	–	–	F30	–[a]	0.95	0.95 (–)
Y3	1.19	1.24	1.22 (0.04)	K31	1.06	1.12	1.09 (0.04)
K4	1.13	1.12	1.13 (0.01)	Q32	0.98	1.09	1.04 (0.08)
L5	1.09	1.15	1.12 (0.04)	Y33	–[a]	0.95	0.95 (–)
I6V[d]	1.10	1.36	1.23 (0.18)	A34	1.07	1.06	1.07 (0.01)
L7I[d]	1.11	1.14	1.13 (0.02)	N35	1.02	0.97	1.00 (0.04)
N8	1.07	1.16	1.12 (0.06)	D36	1.05	1.08	1.07 (0.02)
G9	1.12	1.10	1.11 (0.01)	N37[a]	–	–	–
K10	0.96	1.00	0.98 (0.03)	G38	0.97	1.11	1.04 (0.09)
T11	0.99	0.98	0.99 (0.01)	V39	1.17	1.08	1.13 (0.06)
L12	0.87	0.96	0.92 (0.06)	D40	1.08	1.12	1.10 (0.03)
K13	1.04	1.17	1.11 (0.09)	G41	0.97	1.07	1.02 (0.07)
G14	1.06	1.14	1.10 (0.06)	E42V[d]	1.14	1.20	1.17 (0.04)
E15	1.06	1.08	1.07 (0.01)	W43	1.07	1.06	1.07 (0.01)
T16	1.05	1.06	1.06 (0.01)	T44	1.11	1.11	1.11 (0.00)
T17	1.06	1.06	1.06 (0.00)	Y45	–[a]/0.64[b]	1.09[c]/0.65[b]	1.09 (–)[c]/0.65 (0.01)[b]
T18	1.08	1.05	1.07 (0.02)	D46	1.11	1.19	1.15 (0.06)
E19K[d]	1.05	1.03	1.04 (0.01)	D47	1.12	1.05	1.09 (0.05)
A20	1.04	1.12	1.08 (0.06)	A48	0.99	1.03	1.01 (0.03)
V21	1.09	1.12	1.11 (0.02)	T49	0.82	0.93	0.88 (0.08)
D22[a]	–	–	–	K50	1.06	1.04	1.05 (0.01)
A23	1.00	1.09	1.05 (0.06)	T51	1.03	1.16	1.10 (0.09)
A24E[d]	1.04	0.99	1.02 (0.04)	F52	1.28	1.21	1.25 (0.05)
T25[a]	–	–	–	T53	1.10	1.17	1.14 (0.05)
A26	0.97	1.09	1.03 (0.08)	V54	1.05	1.33	1.19 (0.20)
E27[a]	–	–	–	T55	1.14	1.24	1.19 (0.07)
K28	1.05	1.14	1.10 (0.06)	E56	1.14	1.27	1.21 (0.09)

[a] Not determined due to fast exchange or peak overlap. [b] Fractionation factors for Y45 $O_\eta H$. [c] Fractionation factor for the Y45 backbone amide. [d] Variant residues are labeled with the letter code for both domains (G_{B1} to G_{B2}).

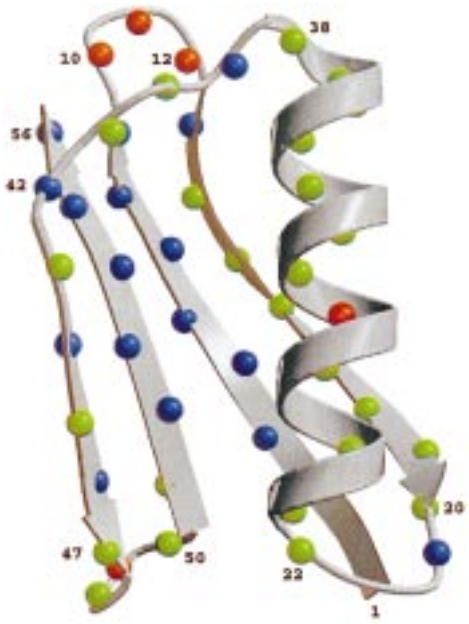

FIGURE 3: Three-dimensional structure of protein G (PDB file 1PGA) showing the ϕ_{av} values for individual backbone amides. Spheres at the backbone nitrogens are colored as follows: $\phi_{av} <$ 1.00 (red), ϕ_{av} values from 1 to 1.10 (green), and $\phi_{av} >$ 1.10 (blue). The diagram was generated using MOLSCRIPT (*45*) in combination with RASTER3D (*46*, *47*).

between G_{B1} and G_{B2}. The reasons for these differences are not clear at present, although it is interesting to note that I6V and V54 are hydrogen bonded to each other in the parallel-stranded region of the β-sheet between strands β1 and β4. However, the differences do not appear to correlate with any changes in hydrogen bonding distances or angles between the G_{B1} and G_{B2} structures (see below).

DISCUSSION

Hydrogen Bonding in Protein G. The lowest backbone ϕ values in protein G are located in the helix (F30 and Y33), $\beta1-\beta2$ turn (K10, T11, and L12), and $\beta3-\beta4$ turn (T49). There do not appear to be any obvious reasons why F30, Y33, and T49 have lowered ϕ values from inspection of the three-dimensional structures. The other three residues (K10, T11, and L12) are in a network of interactions with charged side chains, and such interactions are thought to be important in lowering ϕ values on the basis of recent theoretical calculations (*10*). This hydrogen bond network consists of stabilizing N- and C-capping interactions with the E56 and K13 side chains, respectively (Figure 4). The γ-carboxylate of E56 is hydrogen bonded to the NH of K10 and also to the NH of D40 but with poorer geometry. These interactions occur in solution as well as in the crystal structures (*37*). In addition, a bound water molecule is within hydrogen bonding distance of T11 NH, L12 NH, and the E56 γ-carboxylate group. This water molecule is present in all four X-ray structures of the IgG binding domain of protein G (*25*−*27*). Both direct hydrogen bonding of the E56 side chain to K10 and water-mediated hydrogen bonding to T11 and L12 lead to lower-than-average fractionation factors. These ϕ values are not as low as previously observed for some backbone···side chain interactions, however (*17*, *18*). Our results suggest that water-mediated hydrogen bonds between side chains and the polypeptide backbone can also contribute to lowering ϕ and, in this case, contribute just as much as direct backbone···side chain hydrogen bonds. To our knowledge, this is the first report of backbone NH fractionation factors for such water-mediated hydrogen bonds.

A similar hydrogen bond network is present in HPr where a loop is stabilized by N- and C-capping interactions with

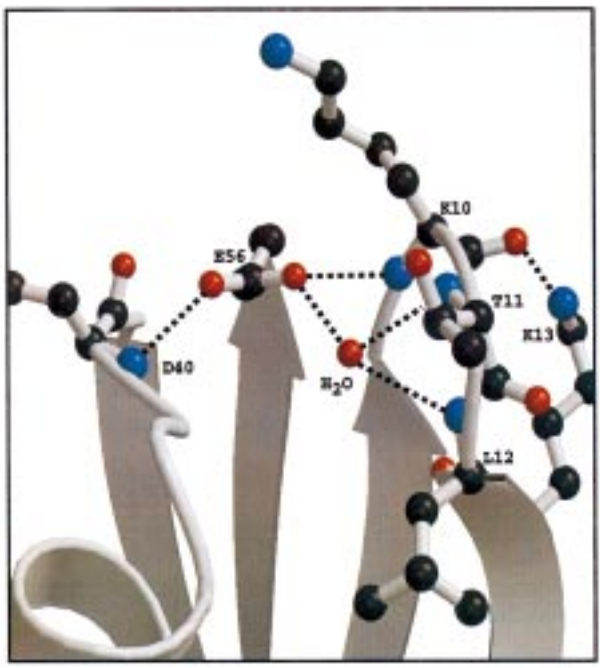

FIGURE 4: Hydrogen bond network involving the E56 side chain, a conserved water molecule, and residues in the $\beta1-\beta2$ turn. The figure was generated as described in the legend of Figure 3.

Asp and Lys groups, respectively (*18*), and low fractionation factors (0.74−0.79 in *Bacillus subtilis* HPr) are observed for backbone NHs in the loop which form hydrogen bonds with the Asp side chain. In the case of *Escherichia coli* HPr, the Asp carboxylate forms an additional strong hydrogen bond with a Ser OH group ($\phi = 0.63$) in the loop, somewhat analogous to the interaction between E56 and bound water in protein G. Further, a hydrogen bond between an Asp carboxylate and a backbone NH is responsible for the lowest ϕ value (0.30) detected in staphylococcal nuclease (*17*). It is interesting to note that, in both HPr and staphylococcal nuclease, these low backbone NH ϕ values correspond to stabilizing hydrogen bond interactions between an acidic residue in a loop and a backbone NH in another loop. In contrast, the E56 side chain in protein G is located on strand $\beta4$ and is sandwiched between two loops, the $\beta1-\beta2$ turn and the $\alpha-\beta3$ loop, forming hydrogen bonds with both.

Notably, the lowest fractionation factor observed in this study comes from a side chain hydroxyl group rather than from a backbone amide. The Y45 $O_\eta H$ is hydrogen bonded to the D47 carboxylate group in all four X-ray structures of protein G with an H···O bond distance of 1.7 Å and an O−H···O bond angle of 175°. Recent NMR data are also consistent with the formation of this hydrogen bond in solution (*37*). The relatively low ϕ value observed for Y45 $O_\eta H$ ($\phi_{av} = 0.65$) is similar in magnitude to the fractionation factor obtained for the hydrogen bond between an Asp carboxylate and a Ser hydroxyl group in HPr (see above). Our data suggest that this hydrogen bond interaction is relatively strong and is consistent with earlier studies which showed that Y45 has a high pK_a (*37*). However, this is probably not a low-barrier hydrogen bond since the O···O distance is greater than 2.5 Å (*14*). Analogous observations have also been made for ketosteroid isomerase where ϕ values of <1 are determined for Tyr···Asp and Tyr···enolate hydrogen bonds at the active site (*38*).

Comparison with Fractionation Results for Other Proteins. A key question in these studies is how do the present results compare with previous studies of backbone NH fractionation. A number of observations can be made from inspection of Figure 5. First, the average backbone fractionation value over all residues in protein G ($\phi_{av} = 1.08 \pm 0.08$) is very similar to that for ubiquitin ($\phi_{av} = 1.11 \pm 0.05$), and the range of ϕ values is relatively narrow in both proteins. In contrast, the average ϕ is lower ($\phi_{av} = 0.84 \pm 0.19$) and the range is broader for staphylococcal nuclease, while the HPr results ($\phi_{av} = 0.92 \pm 0.12$) fall between these two extremes. Second, our results indicate slight enrichment for deuterium at α-helical ($\phi_{av} = 1.03 \pm 0.05$) and β-sheet ($\phi_{av} = 1.13 \pm 0.06$) sites in G_{B1} and G_{B2} on average, consistent with the results of LiWang and Bax for ubiquitin (*19*). In comparison, the ϕ_{av} figures for helical residues in HPr (0.88 ± 0.11) and particularly staphylococcal nuclease (0.79 ± 0.10) indicate significant protium enrichment. β-Sheet residues show a small preference for protium in staphylococcal nuclease ($\phi_{av} = 0.91 \pm 0.15$) and little or no preference in HPr ($\phi_{av} = 0.98 \pm 0.11$). Third, although the absolute values differ, the trend that α-helix ϕ values tend to be slightly smaller on average (by 0.07−0.12) than β-sheet ϕ values holds for all the proteins studied so far. This includes more than 300 measurements at individual backbone amide sites. Fourth, there is no apparent correlation between ϕ and residue type in this study. Loh and Markley (*17*) reported that the six Thr residues in staphylococcal nuclease gave particularly low ϕ values ($\phi_{av} = 0.59 \pm 0.19$ in unligated nuclease). In contrast, ϕ values were measured for nine Thr residues in protein G, and the average fractionation factor obtained ($\phi_{av} = 1.07 \pm 0.09$) was not significantly different from the average obtained over all residues. One Thr residue (T49) did give a relatively low fractionation factor ($\phi_{av} = 0.88$), however. Finally, the average ϕ value of solvent-exposed residues (1.05 ± 0.05) lies between those for α-helical and β-sheet residues. This is consistent with the random coil fractionation factor of 1.1 (*39*) and is comparable with the results obtained for ubiquitin (*19*). In contrast, solvent-exposed residues in staphylococcal nuclease and HPr have ϕ values that are, on average, higher than those for both α-helical and β-sheet residues.

The differences in these fractionation data lead to differing conclusions regarding the stability of intrachain hydrogen bonds. On the basis of the staphylococcal nuclease results, Loh and Markley (*17*) speculate that intramolecular hydrogen bonds between helical residues ($\phi_{av} = 0.79$) may be stronger than amide···water hydrogen bonds ($\phi_{av} = 0.98$). Our results, as well as those of LiWang and Bax (*19*), show that the difference in ϕ between α-helical hydrogen bonds ($\phi_{av} = 1.03$) and amide···water hydrogen bonds ($\phi_{av} = 1.05$) is smaller than in the staphylococcal nuclease study and may not be significant. Moreover, the isotopic preference is for deuterium in G_B and ubiquitin, on average, which would suggest that backbone···backbone hydrogen bonds are generally weak. This is consistent with recent calculations of fractionation factors in peptide clusters which suggest that backbone···backbone interactions typically have ϕ values of >1 but can have values significantly <1 when extended hydrogen bond networks and charged side chains are involved (*10*).

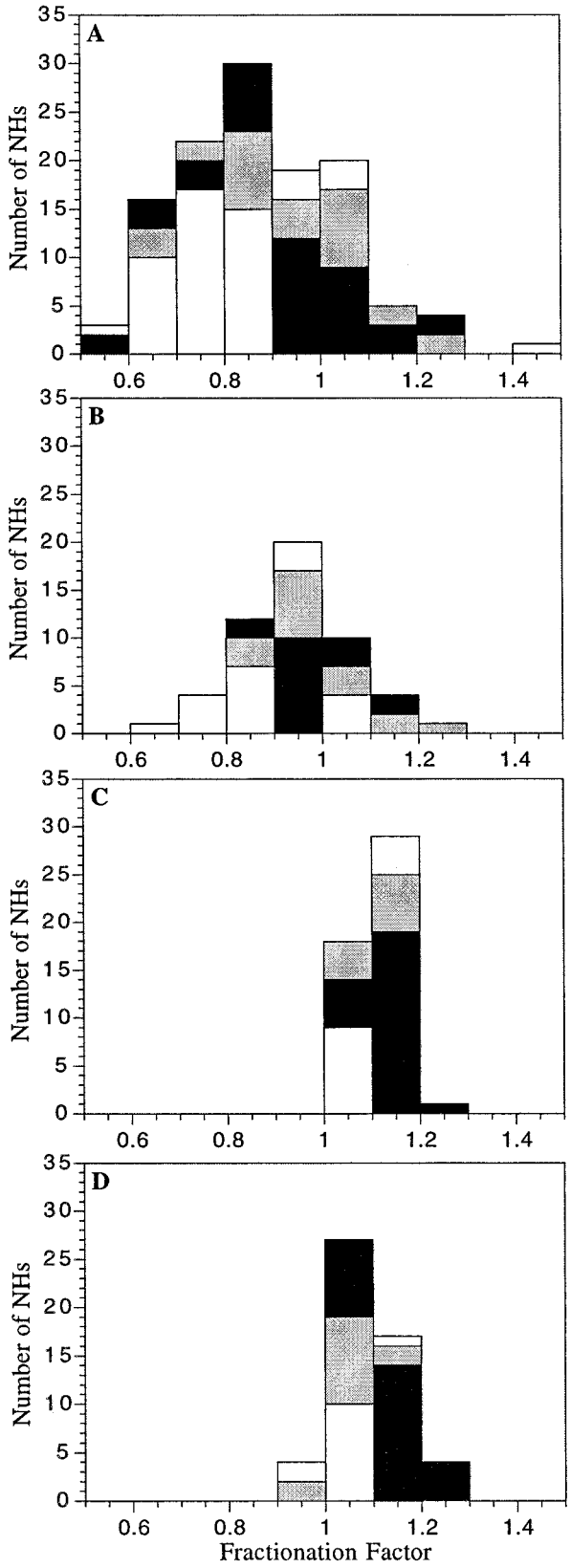

FIGURE 5: Distribution of backbone amide fractionation factors in (A) staphylococcal nuclease (*17*), (B) HPr (*18*), (C) ubiquitin (*19*), and (D) protein G. Fractionation factors are sorted as follows: α-helix (white), β-sheet (black), and solvent-exposed (gray). Fractionation factors are for both ligated and unligated staphylococcal nuclease in panel A and for the ecS31A mutant in panel B. The protein G data depicted in panel D are for the average of the G_{B1} and G_{B2} data.

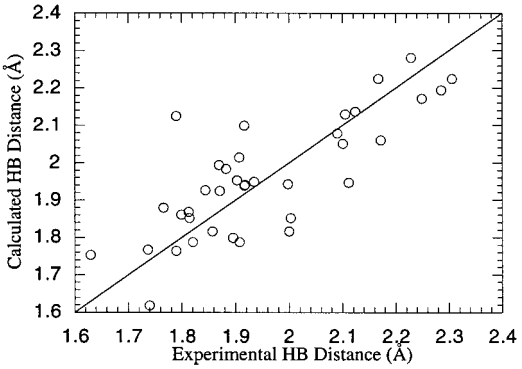

FIGURE 6: Plot of hydrogen bond distances calculated from eq 8 in ref *41* vs the experimental hydrogen bond distances obtained by averaging distances from four X-ray crystal structures of protein G (PDB files 1PGA, 1PGB, 1PGX, and 1IGD).

It has been suggested that the pattern of fractionation factors in a protein may be correlated to the tertiary structure and that strong hydrogen bonds may be conserved at specific sites in similar folding topologies (*18*). For example, low ϕ values are apparently conserved at specific sites in the *B. subtilis* and *E. coli* HPr proteins which have similar folds but only 30% sequence identity. Along these lines, it is interesting to compare protein G and ubiquitin. These two proteins have very low sequence identity (<20%) but similar α/β folds (*40*). While the overall pattern of fractionation is similar with comparable average values for the α-helix and β-sheet secondary structure elements, the few sites of protium enrichment in G_B have no such enrichment in ubiquitin. Indeed, ubiquitin appears to have no sites of protium enrichment in the polypeptide backbone. The residue at position 45 is Phe instead of Tyr, and there is no hydrogen bond network between the $\beta1-\beta2$ turn and the $\alpha-\beta3$ loop. This result suggests that conserved sites of protium enrichment may be a function of not only the folding topology but also the level of sequence identity. In this case, the sequence identity between G_B and ubiquitin may be too low to observe such conserved hydrogen bonds.

Relation between ϕ and Hydrogen Bond Geometry. Plots of ϕ versus crystallographic hydrogen bond distance and/or N−H···O bond angle (not shown) show no apparent correlation, consistent with previous reports (*17*−*19*). One possible reason for the absence of a correlation might be small differences between the solution and crystal forms. However, a comparison of fractionation factors with hydrogen bond distances obtained from chemical shift data (*41*) also failed to show any correlation. Indeed, the hydrogen bond distances derived from NH chemical shifts correlate reasonably well with hydrogen bond distances obtained from the X-ray structures (Figure 6). Wishart et al. (*41*) have shown that the average upfield shift of NHs in α-helices and downfield shift of NHs in β-sheets relative to random coil values corresponds with the ∼0.1 Å difference in crystallographic average hydrogen bond distances between these secondary structural motifs (*42*). In other words, the differences between α-helical and β-sheet hydrogen bond distances seen in X-ray structures are probably also present in solution. Most likely, the differences in ϕ are due to subtle changes in hydrogen bond geometry that are below the limits of detection in the experimental methods used. A recent theoretical study (*10*) suggests that fractionation factors may

be correlated with small differences in the N−H bond distance.

Role of Hydrogen Bonding in Protein Stability. In the limited number of proteins studied to date, the lowest observed ϕ values seem to involve acidic side chains in hydrogen bonding arrangements with either the backbone or other side chains such as Ser or Tyr. The implication is that these types of interactions can form hydrogen bonds that are stronger than regular intrachain (backbone···backbone) hydrogen bonds. Therefore, calculating the total contribution of hydrogen bonding to stability by extrapolating from side chain···side chain or side chain···backbone interactions may overestimate the hydrogen bonding contribution. From such calculations on RNase T1, it was concluded that hydrogen bonds provide about 1.3 kcal/mol per hydrogen bond and that hydrogen bonding and hydrophobicity contribute approximately equally to protein stability (*2*). In the RNase T1 study, three of the 17 hydrogen bonds mutated involve carboxylate group acceptors and two of these have Tyr hydroxyl groups as hydrogen bond donors. In addition, more than half of the residues mutated involve OH···O hydrogen bond interactions that are thought to be quite strong (*43*). Indeed, it would be interesting to measure fractionation factors for this protein. Similar results were obtained in HPr where a hydrogen bond between a Ser OH and Asp carboxylate provides at least 1.3−1.6 kcal/mol of stabilization and corresponds with the lowest observed ϕ value in that system (*18*, *44*). In summary, the hydrogen bonding contribution to stability determined from the RNase T1 mutagenesis work should probably be considered an upper limit and is likely to be lower on the basis of more recent fractionation studies.

CONCLUSIONS

Our results indicate that both α-helical and β-sheet residues in protein G exhibit deuterium enrichment ($\phi > 1$) on average, suggesting that most intrachain hydrogen bonds are weak. Fractionation factors for solvent-exposed residues are between the α-helix and β-sheet values, on average, and are close to those for random coil peptides. The difference in ϕ_{av} values between α-helix and solvent-exposed residues is smaller than in staphylococcal nuclease. This may mean correspondingly smaller differences in hydrogen bond strength for helical hydrogen bonds and peptide···water hydrogen bonds than those proposed by Loh and Markley (*17*). Also, there appears to be no correlation between ϕ and residue type. Overall, the results presented here are similar to those obtained recently for ubiquitin (*19*). In contrast, backbone NH fractionation studies on staphylococcal nuclease and HPr show a preference for protium on average. Part of the reason for this variance may be due to differences in experimental procedures. Although the absolute values vary, fractionation factors in α-helices tend to be smaller than in β-sheets in all four proteins. Another common feature in these fractionation studies is that the lowest ϕ values seem to be associated with hydrogen bonds between acidic side chains and either the backbone or hydroxyl-containing residues such as Ser or Tyr. Therefore, calculating the total contribution of hydrogen bonding to stability by extrapolating from these types of side chain···side chain or side chain···backbone interactions probably overestimates the hydrogen bond contribution to stability.

ACKNOWLEDGMENT

We thank Drs. Andy LiWang, Mike Gilson, and Jim Stivers for critical reading of the manuscript and helpful comments.

REFERENCES

1. Dill, K. A. (1990) *Biochemistry 29*, 7133−7155.
2. Shirley, B. A., Stanssens, P., Hahn, U., and Pace, C. N. (1992) *Biochemistry 31*, 725−732.
3. Honig, B., and Cohen, F. E. (1996) *Folding Des. 1*, R17−R20.
4. Honig, B., and Yang, A.-S. (1995) *Adv. Protein Chem. 46*, 27−58.
5. Dill, K. A., Bromberg, S., Yue, K., Fiebig, K. M., Yee, D. P., Thomas, P. D., and Chan, H. S. (1995) *Protein Sci. 4*, 561−602.
6. Makhatadze, G. I., and Privalov, P. L. (1993) *J. Mol. Biol. 232*, 639−659.
7. Schellman, J. A. (1955) *C. R. Trav. Lab. Carlsberg 29*, 223−229.
8. Kauzmann, W. (1959) *Adv. Protein Chem. 14*, 1−63.
9. Schildbach, B. A., Milla, M. E., Jeffery, P. D., Raumann, B. E., and Sauer, R. T. (1995) *Biochemistry 34*, 1405−1412.
10. Edison, A. S., Weinhold, F., and Markley, J. L. (1995) *J. Am. Chem. Soc. 117*, 9619−9624.
11. Kreevoy, M. M., Liang, T. M., and Chiang, K. C. (1977) *J. Am. Chem. Soc. 99*, 5207−5209.
12. Kreevoy, M. M., and Liang, T. M. (1980) *J. Am. Chem. Soc. 102*, 3315−3322.
13. Schowen, K., and Schowen, R. L. (1982) *Methods Enzymol. 87*, 551−606.
14. Cleland, W. W. (1992) *Biochemistry 31*, 317−319.
15. Cleland, W. W., and Kreevoy, M. M. (1994) *Science 264*, 1887−1890.
16. Loh, S. N., and Markley, J. L. (1993) in *Techniques in Protein Chemistry* (Angeletti, R., Ed.) pp 517−524, Academic Press, San Diego, CA.
17. Loh, S. N., and Markley, J. L. (1994) *Biochemistry 33*, 1029−1036.
18. Bowers, P. M., and Klevit, R. E. (1996) *Nat. Struct. Biol. 3*, 522−531.
19. LiWang, A. C., and Bax, A. (1996) *J. Am. Chem. Soc. 118*, 12864−12865.
20. Alexander, P., Fahnestock, S., Lee, T., Orban, J., and Bryan, P. (1992) *Biochemistry 31*, 3597−3603.
21. Alexander, P., Orban, J., and Bryan, P. (1992) *Biochemistry 31*, 7243−7248.
22. Orban, J., Alexander, P., and Bryan, P. (1992) *Biochemistry 31*, 3604−3611.
23. Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T., and Clore, M. G. (1991) *253*, 657−661.
24. Lian, L.-Y., Derrick, J. P., Sutcliffe, M. J., Yang, J. C., and Roberts, G. C. K. (1992) *J. Mol. Biol. 228*, 1219−1234.
25. Gallagher, T. D., Alexander, P., Bryan, P., and Gilliland, G. (1994) *Biochemistry 33*, 4721−4729.
26. Achari, A., Hale, S. P., Howard, A. J., Clore, G. M., Gronenborn, A. M., Hardman, K. D., and Whitlow, M. (1992) *Biochemistry 31*, 10449−10457.
27. Derrick, J. P., and Wigley, D. B. (1994) *J. Mol. Biol. 243*, 906−918.
28. Orban, J., Alexander, P., and Bryan, P. (1994) *Biochemistry 33*, 5702−5710.
29. Orban, J., Alexander, P., Bryan, P., and Khare, D. (1995) *Biochemistry 34*, 15291−15300.
30. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
31. Bodenhausen, G., and Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185−189.

32. Bax, A., Ikura, M., Kay, L. E., Torchia, D. A., and Tschudin, R. (1990) *J. Magn. Reson. 86*, 304−318.
33. Norwood, T. J., Boyd, J., Heritage, J. E., Soffe, N., and Campbell, I. D. (1990) *J. Magn. Reson. 87*, 488−501.
34. Grzesiek, S., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 12593−12594.
35. Shaka, A. J., Barker, P. B., and Freeman, R. (1985) *J. Magn. Reson. 64*, 547−552.
36. Bax, A., Ikura, M., Kay, L. E., and Zhu, G. (1991) *J. Magn. Reson. 91*, 174−178.
37. Khare, D., Alexander, P., Antosiewicz, J., Bryan, P., Gilson, M., and Orban, J. (1997) *Biochemistry 36*, 3580−3589.
38. Zhao, Q., Abeygunawardana, C., Gittis, A. G., and Mildvan, A. S. (1997) *Biochemistry 36*, 14616−14626.
39. Englander, S. W., and Poulsen, A. (1969) *Biopolymers 7*, 379−393.
40. Kraulis, P. J. (1991) *Science 254*, 581.
41. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) *J. Mol. Biol. 222*, 311−333.
42. Baker, E. N., and Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97−179.
43. Schulz, G. E., and Schirmer, R. H. (1979) *Principles of Protein Structure*, pp 33−36, Springer-Verlag, New York.
44. Hammen, P. K., Scholtz, J. M., Anderson, J. W., Waygood, E. B., and Klevit, R. E. (1995) *Protein Sci. 4*, 936−944.
45. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
46. Bacon, D. J., and Anderson, W. F. (1988) *J. Mol. Graphics 6*, 219−220.
47. Merritt, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. D50*, 869−873.

BI9827114