Comparison of Protein Backbone Entropy and β-Sheet Stability: NMR-Derived Dynamics of Protein G B1 Domain Mutants

Martin J. Stone,* Sonia Gupta, Neil Snyder, and Lynne Regan†

Department of Chemistry
Indiana University, 800 East Kirkwood Avenue
Bloomington, Indiana 47405-0001

Received August 21, 2000

The stability of folded proteins is influenced by the intrinsic propensities of the amino acids to adopt particular secondary structural conformations. Although the statistically observed and experimentally measured propensities correlate well for both α-helices and β-sheets, the physicochemical basis of these propensities remains a topic of discussion.[1−3] A correlation between β-sheet propensities and the intrinsic amide-solvent hydrogen exchange rates of the amino acids[4] led to the suggestion that higher β-sheet propensity results from the obstruction of solvent-backbone hydrogen bonds in the unfolded states of proteins, thus enthalpically stabilizing the folded relative to the unfolded state. Alternatively, β-sheet propensities can be explained by the influence of local steric interactions on the number of conformations accessible to a given residue in a β-sheet versus a random coil.[5] This suggests a strong entropic component to the observed energetic differences in β-sheet formation.

In this communication, we report experimental estimates of backbone entropy for three β-sheet mutants of a small (56 amino acid) protein domain, the B1 domain of *Streptococcal* protein G.[6] The B1 domain (T_m = 89 °C), which consists of a four-stranded β-sheet packed against a single α-helix, has been used previously as a model for determination of intrinsic β-sheet propensities by measurement of the stability of the folded domain after substitution of a surface position in the β-sheet (residue 53) with each of the 20 natural amino acids;[2,3] background mutations were also introduced to minimize cross-strand interactions. In the present study we have used the A53, M53, and T53 mutants described by Smith et al.[2], chosen for their wide range of stabilities ($\Delta\Delta G_{folding}$ relative to A53 at 30 °C) of 0, −4.2, and −8.9 kJ.mol^{-1}, respectively. We estimated the backbone entropy of each mutant using an established relationship between entropy and NMR-derived[7] order parameters for backbone NH groups in proteins.[8,9] The order parameter (S^2), describing the amplitude of angularly restricted internal motion, is obtained by fitting NMR relaxation data to the Lipari−Szabo dynamics formalism.[10] For the diffusion-in-a-cone model of NH vector motions, the conformational entropy of an isolated NH group is:

$$S_{conf} = k_B \ln[\pi(3 - \sqrt{(1 + 8S)})] \quad (1)$$

for $S^2 > 1/64$; k_B is Boltzmann's constant.[8] A major limitation of this approach is that the total conformational entropy of NH groups in a protein backbone may be calculated as the sum of such terms only if the motions of different NH groups are assumed to be independent. Since complete independence of motions is unlikely, it is conservative to consider the calculated entropy to be an upper limit (in magnitude). In addition, this approach is insensitive to conformational entropy contributions derived from translational motion or from rotational motion parallel to the NH bond vector or on time scales slower than molecular tumbling.[11]

^{15}N longitudinal (R_1) and transverse (R_2) auto-relaxation rates, transverse cross-relaxation rates, and heteronuclear {^{1}H}−^{15}N nuclear Overhauser enhancements were measured for the same 53 backbone NH groups in each mutant at 30 °C;[12,13] residues Y33 and E42 were omitted due to spectral overlap. Experimental conditions and data analysis methods were as described for the wild-type B1 domain,[13] unless noted. The rotational diffusion of all three mutants was best represented[14] by a prolate axially symmetric diffusion tensor with identical shape (within error) for the three mutants. Model-free calculations were performed using the weighted average values of $2D_{zz}/(D_{xx} + D_{yy}) = 1.28$, $\theta = 84.5°$, and $\phi = 182.2°$; angles are defined relative to the crystal structure coordinates of the wild-type B1 domain (PDB code: 2GB1).[15] Effective molecular correlation times were chosen to be optimally consistent with these values and the R_2/R_1 ratios for each mutant;[13] $\tau_m[= (2D_{\|} + 4D_{\perp})^{-1}]$= 3.57, 3.44, and 3.47 ns, for A53, M53, and T53 mutants, respectively. Initial dynamics model selection[16] was performed independently for each residue in each mutant. Final calculations were then performed for each residue using the simplest model consistent with the data for that residue in all three mutants.

The backbone NH groups of the three mutants show a similar variation of order parameters across the sequence (Figure 1a). Thus, S^2 values correlate well between mutants; r^2 = 0.88 for A53 versus M53, 0.75 for A53 versus T53, and 0.77 for M53 versus T53. Nevertheless, there is an offset between the order parameter curves. The 10% trimmed weighted average order parameters are lowest for the T53 mutant and highest for the A53 mutant (Table 1). Qualitatively, these data indicate that the most stable mutant (T53) has the greatest backbone flexibility (entropy) whereas the least stable mutant (A53) has the most rigid backbone. Thus, these results provide experimental support for the proposal

† Department of Molecular Biophysics and Biochemistry and Department of Chemistry, Yale University, P.O. Box 208114, 266 Whitney Avenue, New Haven, CT 06520-8114.

(1) (a) Chou, P. Y.; Fasman, G. D. *Biochemistry* **1974**, *13*, 211−222. (b) Creamer, T. P.; Rose, G. D. *Proc. Natl. Acad. Sci. U.S.A* **1992**, *89*, 5937−5941. (c) Chakrabartty, A.; Baldwin, R.L. *Adv. Protein Chem.* **1995**, *46*, 141−176. (d) Kim, C. A.; Berg, J. M. *Nature* **1993**, *362*, 267−270. (e) Minor, D. L.; Kim, P. S. *Nature* **1994**, *371*, 264−267. (f) Munoz, V.; Serrano, L. *Proteins* **1994**, *20*, 301−311. (g) Smith, C. K.; Regan, L. *Science* **1995**, *270*, 980−982. (h) Smith, C. K.; Regan, L. *Acc. Chem. Res.* **1997**, *30*, 153−161. (i) Merkel, J. S.; Sturtevant, J. M.; Regan, L. *Struct. Folding Des.* **1999**, *7*, 1333−1343.

(2) Smith, C. K.; Withka, J. M.; Regan, L. *Biochemistry* **1994**, *33*, 5510−5517.

(3) Minor, D. L.; Kim, P. S. *Nature* **1994**, *367*, 660−663.

(4) Bai, Y.; Englander, S. W. *Proteins* **1994**, *18*, 262−266.

(5) Street, A. G.; Mayo, S. L. *Proc. Natl. Acad. Sci. U.S.A* **1999**, *96*, 9074−9076.

(6) Boyle, M. *Bacterial Immunoglobulin-Binding Proteins*; Academic Press: San Diego, CA, 1990.

(7) Abbreviations and symbols used: ΔG_{conf}, NH conformational free energy; NMR, nuclear magnetic resonance; R_1, longitudinal relaxation rate constant; R_2, transverse relaxation rate constant; S^2, NH order parameter; S_{conf}, NH conformational entropy.

(8) Yang, D.; Kay, L. E. *J. Mol. Biol.* **1996**, *263*, 369−382.

(9) (a) Akke, M.; Bruschweiler, R.; Palmer, A. G., III. *J. Am. Chem. Soc.* **1993**, *115*, 9832−9833. (b) Li, Z. G.; Raychaudhuri, S.; Wand, A. J. *Protein Sci.* **1996**, *5*, 2647−2650.

(10) Lipari, G.; Szabo, A. *J. Am. Chem. Soc*. **1982**, *104*, 4546−4559.

(11) (a) Daragan, V. A.; Mayo, K. H. *Prog. Nucl. Magn. Reson. Spectrosc.* **1997**, *31*, 63−105. (b) Fischer, M. W. F.; Zeng, L.; Majumdar, A.; Zuiderweg, E. R. P. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8016−8019. (c) Forman-Kay, J. D. *Nat. Struct. Biol.* **1999**, *6*, 1086−1087. (d) Cavanagh, J.; Akke, M. *Nat. Struct. Biol*. **2000**, *7*, 11−13.

(12) Farrow, N. A.; Muhandiram, R.; Singer, A. U.; Pascal, S. M.; Kay, C. M.; Gish, G.; Shoelson, S. E.; Pawson, T.; Forman-Kay, J. D.; Kay, L. E. *Biochemistry* **1994**, *33*, 5984−6003.

(13) Seewald, M. J.; Pichumani, K.; Stowell, C.; Tibbals, B. V.; Regan, L.; Stone, M. J. *Protein Sci.* **2000**, *9*, 1177−1193.

(14) (a) Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 12562−12566. (b) Blackledge, M.; Cordier, F.; Dosset, P.; Marion, D. *J. Am. Chem. Soc*. **1998**, *120*, 4538−4539.

(15) Gronenborn, A. M.; Filpula, D. R.; Essig, N. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. *Science* **1991**, *253*, 657−661.

(16) Mandel, A. M.; Akke, M.; Palmer, A. G., III. *J. Mol. Biol*. **1995**, *246*, 144−163.

10.1021/ja003094l CCC: $20.00 © 2001 American Chemical Society
Published on Web 12/08/2000

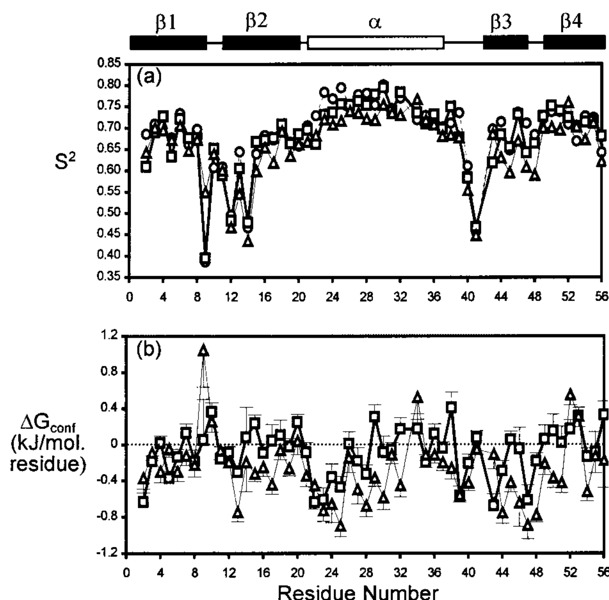

Figure 1. (a) NH order parameters for each residue in the A53 (circles, broken lines), M53 (squares, thick lines), and T53 (triangles, thin lines) mutants; the average standard error in S^2 values is 0.011. (b) NH group ΔG_{conf} ($= -T\Delta S_{conf}$) values for each residue of the M53 (squares, thick lines) and T53 (triangles, thin lines) mutants relative to the corresponding residue in the A53 mutant; error bars are shown in only one direction for clarity. The positions of the β-strands and the α-helix are indicated at the top.

Table 1. Thermodynamic and Dynamics Parameters for B1 Domain Mutants at 30 °C

	A53	M53	T53
wt. av S^2 values[a]	0.70 (0.05)	0.69 (0.06)	0.67 (0.05)
wt. av ΔG_{conf} (kJ/mol.residue)[a,b,c]	0	−0.07 (0.23)	−0.29 (0.24)
total ΔG_{conf} (kJ/mol)[b,c]	0	−4.3	−14.2
$\Delta\Delta G_{folding}$ (kJ/mol)[b,d]	0	−4.2	−8.9
$\Delta\Delta A$ (kJ/mol)[b,e]	0	−2.5	−6.3
S^2 (residue 53)	0.67 ± 0.01	0.70 ± 0.01	0.70 ± 0.01

[a] 10% trimmed weighted average; standard deviations in parentheses. According to the two-tailed z-statistic test (ref 17), the S^2 distributions for A53 and T53 differ significantly ($p = 0.07$), whereas the differences for A53 versus M53 and M53 versus T53 are less significant ($p = 0.62$ and $p = 0.65$, respectively). [b] Relative to the A53 mutant. [c] Determined from order parameters (eq 1). [d] From ref 2, calculated using ΔC_p of wild type (2.6 kJ/mol·K; ref 3). [e] From ref 5; these values incorporate an empirical scaling factor.

that the stability of these mutants is influenced to some extent by backbone entropy.

The backbone conformational free energy changes ($\Delta G_{conf} = -T\Delta S_{conf}$) calculated from eq 1 for each residue of the M53 and T53 mutants relative to the A53 mutant are presented in Figure 1b. Table 1 lists the trimmed weighted average ΔG_{conf} values and the total ΔG_{conf} estimated from the sum of ΔG_{conf} values across the whole domain. The total ΔG_{conf} values are in the same order as the total free energy differences between the mutants ($\Delta\Delta G_{folding}$)[2] and the calculated Helmholtz free energy differences ($\Delta\Delta A$)[5] (Table 1). To ensure that the total ΔG_{conf} values were not biased by the above choices of data analysis methods, several alternative methods were also investigated (see Supporting Information). In all cases, the conclusion that backbone conformational entropy makes a significant contribution to the stability differences between these three mutants is supported by the calculated conformational free energies.

Despite the striking agreement between the estimated total ΔG_{conf} and $\Delta\Delta G_{folding}$ values, rigorous quantitative comparison of these values is complicated by several factors. First, the current analysis relies on the assumption of independent NH group motions.[8] Any correlated motions will tend to decrease the magnitude of the ΔG_{conf} values making the backbone entropy effect less dramatic. Second, the $\Delta\Delta G_{folding}$ values depend on the properties of the unfolded state which are not observed in the current experiments. This concern is somewhat reduced by the likelihood that the backbone dynamics of the unfolded peptide chains are similar for the three mutants. It is important to emphasize that the current study does not provide any information about differences in side chain flexibility, solvent entropy, or any enthalpic contributions between these three mutants.

Although, the total ΔG_{conf} values calculated from order parameters correspond well to the Helmholtz free energy differences of Street and Mayo,[5] the latter calculations were based solely on the conformational distribution of a single amino acid (Xaa) within an Ala-Xaa-Ala tripeptide, whereas in the current study, it is the entropy differences calculated over the whole domain that appear to give rise to the stability differences between the three mutants. The mechanism by which a mutation at position 53 causes changes in flexibility throughout the domain is not evident from the current data. One speculation is that substitution of Ala-53 by Thr causes a slight local structural rearrangement leading to a reduction in the tertiary packing efficiency of the domain and thus an increase in global motions. Such a mechanism might reflect cooperative interactions stabilizing the structure, further highlighting the possibility of correlated motions between spatially proximal groups.

In summary, the current analysis suggests that backbone entropy contributes to the variation in stability of the three B1 domain mutants and that the effect is spread throughout the domain rather than being localized at the site of mutation. Whether this effect can provide a general explanation for β-sheet propensities awaits the analysis of additional B1 domain mutants and of other proteins, in which the β-sheet residues are often more rigid than those in the B1 domain.[18]

Acknowledgment. We thank Drs. A. G. Palmer, III (Columbia University), M. Rance (University of Cincinnati), and L. E. Kay (University of Toronto) for software and Drs. C. K. Smith and J. S. Merkel for discussions and reagents. This work was supported by grants from the National Science Foundation (MCB-9600968 to M.J.S and MCB-9316863 to L.R.), from the National Institutes of Health (GM-55055 to M.J.S. and GM-57265 to L.R.), and from the American Chemical Society Petroleum Research Fund (to M.J.S.).

Supporting Information Available: Tables of relaxation and dynamics parameters for each mutant. A summary of ΔG_{conf} results using alternative data analysis methods (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA003094L

(17) Devore, J. L. Probability and Statistics for Engineering and the Sciences; Duxbury, Pacific Grove, CA, 2000, pp 360−361.
(18) Goodman, J. L.; Pagel, M. D.; Stone, M. J. *J. Mol. Biol.* **2000**, *295*, 963−978.