

# Critical role of $\beta$ -hairpin formation in protein G folding

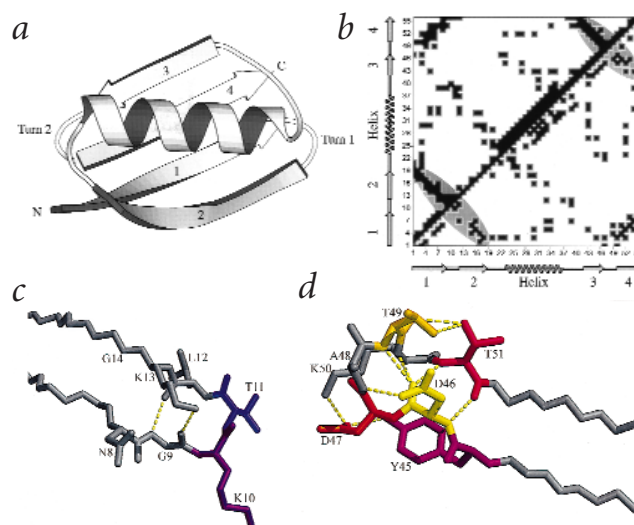
Erika L. McCallister, Eric Alm and David Baker

Department of Biochemistry, University of Washington, Seattle, Washington 98195, USA.

**Comparison of the folding mechanisms of proteins with similar structures but very different sequences can provide fundamental insights into the determinants of protein folding mechanisms. Despite very little sequence similarity, the ~60 residue IgG binding domains of protein G and protein L both consist of a single helix packed against a four-stranded sheet formed by two symmetrically disposed  $\beta$ -hairpins. We demonstrate that, as in the case of protein L, one of the two  $\beta$ -turns of protein G is formed and the other disrupted in the folding transition state. Unlike protein L, however, in protein G it is the second  $\beta$ -turn that is formed in the folding transition state ensemble. Substitution of an Asp residue by Ala in protein G that eliminates an  $i,i+2$  side chain-main chain hydrogen bond in the second  $\beta$ -turn slows the folding rate ~20-fold but has virtually no effect on the unfolding rate. Taken together with previous results, these findings suggest that the presence of an intact  $\beta$ -turn in the folding transition state is a consequence of the overall topology of protein L and protein G, but the particular hairpin that is formed is determined by the detailed interatomic interactions that determine the free energies of formation of the isolated  $\beta$ -hairpins.**

Native state topology is an important determinant of protein folding mechanisms<sup>1</sup>. Proteins with similar structures but very different sequences often<sup>2–4</sup>, but not always<sup>5</sup>, have structurally similar folding transition state ensembles, and folding rates are correlated with the contact order, the average sequence separation between residues that contact each other in the native state<sup>6</sup>. Recently developed models have shown considerable promise in predicting folding rates and the distribution of structure in the folding transition state ensemble from the structure of the native state<sup>7–9</sup>. For proteins with considerable symmetry, however, the choice between topologically equivalent folding mechanisms must be determined by other factors. For example, the folding of protein L must be determined by factors beyond chain topology because one of two symmetrically disposed  $\beta$ -hairpins is formed and the other disrupted in the folding transition state<sup>10,11</sup>. To gain insight into the factors that contribute to this asymmetry in folding, we have studied the folding mechanism of the 57 residue IgG binding domain of protein G<sup>12,13</sup> (referred to as protein G in this paper), a protein with a very similar overall topology but only 15% sequence identity to protein L. As in the case of protein L, both the overall structure of protein G (Fig. 1a) and the distribution of core contacts (Fig. 1b) are highly symmetric.

Previous work has provided a rich store of information on the folding of protein G. Folding of protein G at high pH occurs without detectable intermediates<sup>14–16</sup>, but at pH 4.0 in the presence of 0.4 M sodium sulfate double exponential folding kinetics suggest the presence of a collapsed intermediate<sup>17</sup>. The second  $\beta$ -hairpin is stable in isolation<sup>18</sup>, and hydrogen/deuterium (H/D) exchange studies<sup>19</sup> indicate that this portion of the structure may be formed in the denatured state. The isolated second  $\beta$ -hairpin



**Fig. 1** Protein G. **a**, Backbone ribbon diagram of the crystal structure of protein G with the strands,  $\beta$ -hairpin turns, and N-terminus and C-terminus labeled<sup>13</sup>. The image was created using Molscript<sup>32</sup>. **b**, Contact map of protein G with backbone hydrogen bonds indicated above the diagonal and side chain contacts below the diagonal. Shaded areas indicate side chain contacts in each of the  $\beta$ -hairpins. **c**, Interactions within the first  $\beta$ -turn of protein G. **d**, Interactions within the second  $\beta$ -turn of protein G. Residues are colored by  $\Phi$  from 0.0 (blue) to 0.5 (red) to 1.0 (yellow). Hydrogen bonds are shown in yellow. The point of view is rotated 180° about the y axis from the view in (a). The images were created using Midas Plus<sup>33,34</sup>.

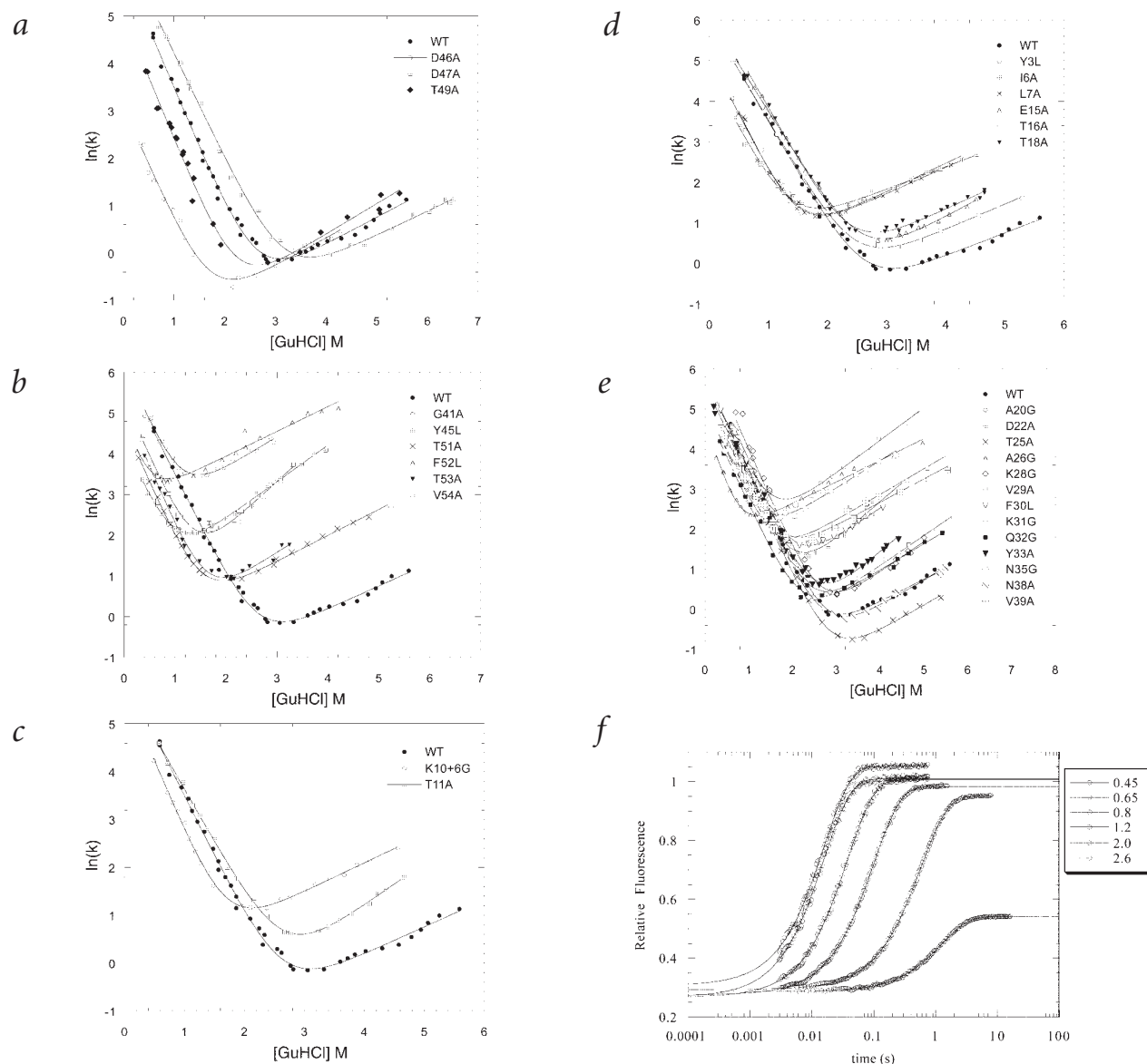
folds on the 10  $\mu$ s time scale, two orders of magnitude faster than the intact protein folds<sup>20</sup>.

## Effect of mutations on folding kinetics

To probe the role of the second  $\beta$ -turn in protein G folding, we examined the effects of mutations in this region on the folding and unfolding rates (Fig. 2a). Truncation of Asp 46 to Ala, which removes a hydrogen bond between one of the side chain oxygen atoms and the amide proton of residue 48 (Fig. 1d), slowed the folding rate ~20-fold (from 412 s<sup>-1</sup> to 21.8 s<sup>-1</sup>) but had virtually no effect on the unfolding rate. Truncation of Thr 49 to Ala, which removes side chain-side chain hydrogen bonds between Thr 49 and Thr 51 (Fig. 1d), also slowed the folding rate with no effect on unfolding. The  $\Phi$  values for both D46A and T49A mutations are very close to 1 (Table 1), suggesting that the hydrogen bonds made by these residues are largely intact in the folding transition state ensemble. These hydrogen bonds may play critical roles in establishing the register of the  $\beta$ -hairpin and the conformation around the  $\beta$ -turn. While the hydrogen bonds made by Asp 46 and Thr 49 stabilize both the transition state and the native state, the D47A mutation, which removes a buried salt bridge with Lys 50 across the turn, and a partially buried hydrogen bond between Asp 47 and Tyr 45 (Fig. 1d), stabilizes the protein and increases the rate of folding (Fig. 2a). Interestingly, the D47A mutation also stabilizes the isolated  $\beta$ -hairpin<sup>21</sup>.

Additional mutations suggest that the second  $\beta$ -hairpin may be more completely formed near the  $\beta$ -turn than at the base in the folding transition state. The Y45L, T51A, F52L and T53A mutations, which remove interactions in the portion of the hairpin closest to the turn, have intermediate  $\Phi$  values, while the V54A and G41A mutations at the base of the hairpin destabilize the protein considerably but have quite low  $\Phi$  values (Table 1).

## letters



Mutations in the first  $\beta$ -turn, a six-Gly insertion following Lys 10 (K10 + 6G) and T11A, destabilize the protein but have little effect on the folding rate (Fig. 2c). The low  $\Phi$  values of these mutations suggest that the first turn is formed after the rate-limiting step in folding (Table 1). Mutations in the second  $\beta$ -strand (E15A, T16A and T18A) also have low  $\Phi$  values. In contrast, mutations in the first  $\beta$ -strand (Y3L, I6A and L7A), which remove interactions with the second  $\beta$ -hairpin, slow the folding rate, increase the unfolding rate (Fig. 2d) and have intermediate  $\Phi$  values. A plausible interpretation of these data is that the first  $\beta$ -hairpin is not formed at the folding transition state, but the hydrophobic residues in the first strand do make some interactions with the relatively ordered second  $\beta$ -hairpin.

The kinetic consequences of mutations in the helix are complex and suggest that the helix may sample a variety of conformations in the folding transition state ensemble. Of four Gly mutations of solvent exposed residues, three in the C-terminal portion of the helix (K31G, Q32G and N35G) slow the folding rate  $\sim 2.5$ -fold, while K28G near the center of the helix increases the folding rate  $\sim 2$ -fold, suggesting that the helix is disrupted

**Fig. 2** Folding kinetics of protein G mutants. Dependence of the folding and unfolding rate constants of the mutant proteins on the denaturant concentration. **a**, Second  $\beta$ -turn. **b**, Second  $\beta$ -hairpin. **c**, First  $\beta$ -turn. **d**, First  $\beta$ -hairpin. **e**, Helix. The data for the wild type protein G (filled circles) is shown in all panels. **f**, Lack of burst phase in refolding of the D46A mutant in the presence of 0.4 M  $\text{Na}_2\text{SO}_4$ .

near its center and better formed towards the C-terminus in the folding transition state ensemble. Mutations that remove hydrogen bonds at the beginning (A20G, D22A) and end (N35G) of the helix have large effects on the unfolding rate but little effect on the folding rate (Fig. 2e), suggesting these interactions are made late in folding. Mutations that remove buried interactions between the helix and the first  $\beta$ -hairpin (T25A, A26G, V39A, F30L, Y33A, A34G, V39A) have relatively low  $\Phi$  values (Table 1).

In summary, the second  $\beta$ -turn appears to be largely formed in the folding transition state, while the first  $\beta$ -turn and the helix appear to be relatively disordered (Fig. 3a). Mutations in the first, third, and fourth  $\beta$ -strands have intermediate  $\Phi$  values,

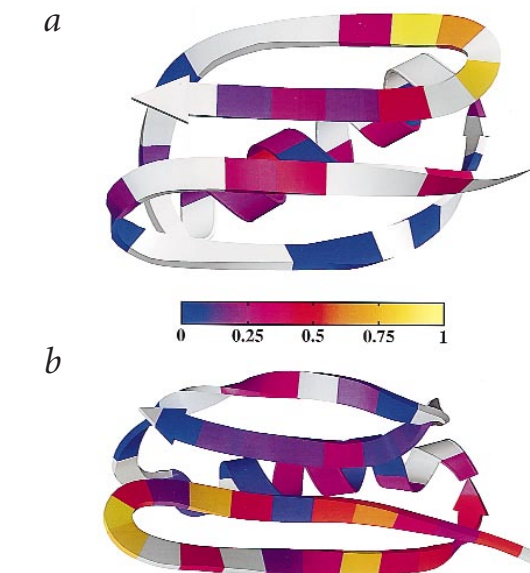
**Fig. 3** Comparison of  $\Phi$  value distributions. **a**, Protein G<sup>13</sup>. **b**, Protein L<sup>11</sup>. Residues are colored by  $\Phi$  from 0.0 (blue) to 0.5 (red) to 1.0 (yellow).  $\Phi_F$  values were calculated from  $k_f$  at 0.5 M GuHCl and  $k_u$  at 4 M GuHCl in order to avoid extrapolations<sup>11</sup>.  $\Phi_F = -RT \ln(k_{0.5 M}^{mut} / k_{0.5 M}^{wild\ type}) / \Delta\Delta G$ ;  $\Delta\Delta G = -RT[(\ln(k_{0.5 M}^{mut} / k_{0.5 M}^{wild\ type})) + (\ln(k_{4 M}^{wild\ type} / k_{4 M}^{mut}))]$ . The images were created using Molscript<sup>32</sup> and Raster3d<sup>35,36</sup>.

suggesting partial formation of the three-stranded  $\beta$ -sheet involving these strands in the folding transition state.

Simultaneous characterization using both continuous-flow and stopped-flow methods of the folding kinetics of wild type protein G between 0.4 and 0.8 M guanidine hydrochloride in the presence of 0.4 M Na<sub>2</sub>SO<sub>4</sub> suggested the presence of two distinct kinetic phases<sup>22</sup>. Because the stopped-flow instrument used in our experiments has a dead time of 2.9 ms, we are only able to accurately monitor folding kinetics for reactions with rates  $<200\text{ s}^{-1}$ . Under such conditions (up to 0.8 M guanidine for the wild type protein and the more rapidly folding mutants), the folding reactions of all proteins were well described by a single exponential without significant missing amplitude. In particular, even under highly stabilizing conditions (0.4 M Na<sub>2</sub>SO<sub>4</sub>), there was no change in fluorescence in the first milliseconds after initiation of refolding of the D46A mutant (Fig. 2f). The mutation thus appears to abolish the rapid conformational transition observed by Roder and coworkers<sup>17, 22</sup> for wild type protein G, suggesting that formation of the second hairpin may be involved in the fast kinetic phase. This conclusion, however, must be regarded as somewhat preliminary since we were unable to detect biphasic folding kinetics for our wild type protein G construct because of the very rapid folding rate. Further analysis of the refolding of the mutants under highly stabilizing conditions using a combination of continuous-flow<sup>22</sup> and stopped-flow methods should shed considerable light on the early events in protein G refolding.

### Protein G folding free energy landscape

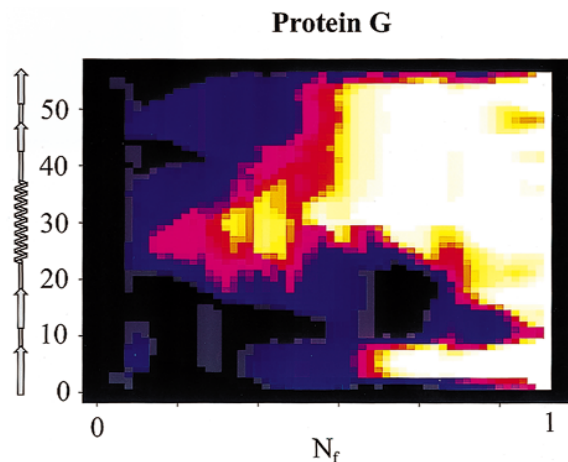
We have recently described a simple model for folding free energy landscapes that attempts to capture the tradeoff between the formation of attractive native interactions and the loss of chain configurational entropy during folding<sup>7</sup>. The folding landscape for protein G projected onto the reaction coordinate  $N_f$ , the total number of residues ordered, is shown in Fig. 4. Interestingly, late in the folding reaction (high  $N_f$ ), the first, third, and fourth  $\beta$ -strands are predicted to be largely ordered, while the second  $\beta$ -strand appears to be ordered last. The helix is (incorrectly) predicted to form early in the folding reaction, which may in part



be a consequence of the limitation that the only interactions considered are between segments that are ordered as in the native state. More consistent with our data on both protein G and protein L is a model in which the hydrophobic residues in the helix make contacts with the second (first in the case of protein L) hairpin at the rate limiting step in folding, independent of the extent of formation of the helix. Interestingly, full atom molecular dynamics (MD) simulations of protein G refolding<sup>23</sup> also suggest the early formation of the second hairpin and the helix. In any event, the differences between the two hairpins appear to be at least in part captured by the surface area based representation of the attractive native interactions used in the simple model.

### Role of second hairpin in protein G folding

Our finding that the second  $\beta$ -turn is the most intact portion of protein G in the folding transition state complements and extends previous studies of protein G folding. The second hairpin was found to be stable in isolation, protected from H/D exchange early in folding<sup>19</sup>, and relatively well ordered in unfolded configurations sampled in MD simulations<sup>24</sup>. The 10  $\mu\text{s}$  time scale found for formation of the isolated second hairpin by Munoz and Eaton<sup>20</sup> is interesting in light of the finding that there is no detectable rapid folding event in the D46A mutant, which has been shown to significantly destabilize the isolated  $\beta$ -hairpin<sup>21</sup>.



**Fig. 4** Hierarchy of protein G folding. In the configurations considered in this model each residue is either ordered as in the native state or completely disordered. Ordered residues are required to occur within one or two contiguous segments in the linear sequence. The interaction energy between ordered residues is proportional to the surface area buried between them in the native state; disordered residues make no interactions. Folding is favored by formation of these attractive native interactions and disfavored by the entropic costs of ordering residues and closing loops between ordered segments. The reaction coordinate,  $N_f$ , is the fraction of the ordered residues ( $N_f = 0$  is the fully unfolded state and  $N_f = 1$  is the fully folded state). The y-axis indicates position along the sequence. All configurations of the system were enumerated, and the Boltzmann averaged frequency of ordering of each residue (y-axis), as a function of  $N_f$  (x-axis), is indicated by the color: black–blue, 0–0.25; blue–magenta, 0.25–0.50; magenta–red, 0.50–0.75; red–yellow, 0.75–0.88; and yellow–white, 0.88–1.0 (ref. 7).

## letters

Table 1 Kinetic parameters for protein G folding<sup>1</sup>

	$k_f^{0.5M}$ (s <sup>-1</sup> )	$m_f$ (kcal mol <sup>-1</sup> M <sup>-1</sup> )	$k_u^{4M}$ (s <sup>-1</sup> )	$m_u$ (kcal mol <sup>-1</sup> M <sup>-1</sup> )	$\Delta\Delta G_u$ (kcal mol <sup>-1</sup> )	$\Phi_f$
Wild type	116.68	1.48	1.22	0.33		
Y3L	39.83	1.71	6.87	0.33	1.62	0.38
I6A	29.23	1.63	11.18	0.33	2.09	0.38
L7A	42.02	1.79	10.72	0.38	1.85	0.32
K10 + 6G	69.61	1.75	8.46	0.36	1.42	0.21
T11A	114.82	1.30	3.40	0.52	0.60	0.02
E15A	135.74	1.36	3.21	0.46	0.47	-0.19
T16A	116.31	1.45	2.34	0.36	0.38	0.00
T18A	166.88	1.50	3.84	0.39	0.46	-0.45
A20G	106.16	1.42	68.02	0.48	2.39	0.02
D22A	58.99	1.40	12.69	0.49	1.75	0.23
T25A	85.46	1.40	0.61	0.37	-0.22	-0.81
A26G	23.77	2.14	41.24	0.32	2.96	0.31
K28G	225.79	1.50	2.57	0.45	0.05	- <sup>2</sup>
V29A	85.59	1.39	2.99	0.44	0.70	0.26
F30L	103.32	1.40	12.53	0.43	1.42	0.05
K31G	51.73	1.45	17.62	0.37	2.02	0.23
Q32G	45.37	1.39	2.67	0.38	1.00	0.55
Y33A	84.61	1.29	4.30	0.54	0.92	0.20
A34G	47.72	1.55	36.11	0.42	2.48	0.21
N35G	52.20	1.54	40.99	0.40	2.50	0.19
N37A	145.03	1.45	1.13	0.37	-0.17	- <sup>2</sup>
V39A	71.62	1.46	14.48	0.34	1.72	0.16
G41A	126.90	1.71	176.71	0.44	2.84	-0.02
Y45L	20.36	1.64	67.52	0.53	3.34	0.30
D46A	6.61	1.43	1.38	0.35	1.74	0.96
D47A	204.70	1.30	0.92	0.30	-0.49	0.67
T49A	41.13	1.52	1.50	0.38	0.72	0.84
T51A	27.84	1.66	7.35	0.37	1.87	0.44
F52L	36.93	-0.36	173.85	-2.31	3.54	0.19
T53A	48.47	1.70	10.82	0.48	1.91	0.27
V54A	52.74	1.65	86.06	0.67	2.93	0.16

<sup>1</sup>Rate of folding ( $k_f$ ) is reported at 0.5 M GuHCl; rate of unfolding ( $k_u$ ) is reported at 4 M GuHCl to avoid extrapolation. The  $m$ -values,  $m_f$  and  $m_u$ , indicate the denaturant dependencies of the folding and unfolding rates, respectively. Experimental conditions and calculations are as described in the Methods section.

<sup>2</sup>Mutation has little or no effect on stability ( $\Delta\Delta G_u < 0.3$  kcal mol<sup>-1</sup>).

A resemblance between the distribution of structure in the denatured state and the folding transition state may be a general feature of the folding of small proteins. In protein L, for example, the first  $\beta$ -turn appears to be more ordered than the second  $\beta$ -turn in the denatured state<sup>25,26</sup> and in the folding transition state<sup>11</sup> (the SH3 domain is another example<sup>27</sup>). While certainly not necessary, such similarity between the denatured state and the transition state ensembles is not surprising if the barrier to folding is primarily entropic; structural elements with relatively low free energies of formation are more likely to be formed in these ensembles than less stable structural elements.

Despite the partial formation of the second  $\beta$ -hairpin in the denatured state, the large effect of the D46A mutation on both the folding rate and overall stability (a decrease of  $\sim 2$  kcal mol<sup>-1</sup>) suggests that the mutation increases the free energy of the denatured state much less than that of the transition state and native state, perhaps because of large numbers of nearly isoenergetic conformers lacking the  $\beta$ -turn (the loss of the energetically favorable interaction may be compensated by a large increase in entropy).

## Comparison of protein G and protein L

Protein L has an overall structure very similar to that of protein G, but a very different sequence (15% of the residues are identical in a structure-based alignment)<sup>28</sup>. There are interesting similarities and differences in the folding mechanisms of the two proteins (Fig. 3). An overall similarity in the folding mechanism is suggested by the fact that the folding rates of the two proteins differ by less than an order of magnitude and, in both proteins, one of the two  $\beta$ -turns is largely formed and the other largely disrupted in the folding transition state. However, in protein L it is the first  $\beta$ -turn<sup>10</sup>, and in protein G the second  $\beta$ -turn, that is formed in the transition state ensemble.

The differences between the order of formation of the two  $\beta$ -turns in protein L and protein G are likely to reflect differences in the free energies of formation of the isolated  $\beta$ -hairpins in the two proteins. In protein G, only the second  $\beta$ -hairpin is stable in isolation, suggesting that local interactions within the second hairpin are responsible for the predominance of the second turn in the protein G folding transition state. Indeed, there is a dense network of interatomic interactions (including many hydrogen bonds) in the second turn (Fig. 1d) in the native structure, but relatively few interactions between side chains in the first  $\beta$ -turn (Fig. 1c). All side chain-side chain contacts expected within a  $\beta$ -hairpin are formed in the case of the second (Fig. 1b, top right shaded area) but not the first (Fig. 1b, bottom right shaded area)  $\beta$ -hairpin; the presence of two Gly residues in the first  $\beta$ -hairpin significantly reduces the side chain contact density near the  $\beta$ -turn. These differences in the contact density are responsible for the favoring of formation of the second hairpin in the simple model calculations (Fig. 4) and, we speculate, in the actual folding process. In protein L we believe the distinction between the two turns during folding results from high backbone torsional strain in the second  $\beta$ -turn, as three consecutive residues have positive  $\Phi$  angles.

In summary, the lowest free energy path to the native state for the protein G/protein L fold appears to involve formation of one of the two  $\beta$ -hairpin. The hairpin that is formed appears to be the one with the lowest free energy (largest number of favorable interactions and least backbone torsional strain) and hence it is also more intact than the other  $\beta$ -hairpin in the denatured state ensemble. For proteins with symmetric structures, several alternative routes to the native state may be equally consistent with the native state topology, and the route that is selected is likely to involve formation of the lowest free energy substructures.

## Methods

**Materials.** All reagents, solutions, and enzymes used for molecular biology procedures were as described<sup>29,30</sup>.

**Cloning, expression and purification.** GB1-57 was generously provided by K.T. O'Neil (Dupont). For ease of expression and purification in our studies, GB1-57 was cloned into the *NdeI* and *BamHI* sites of the pET 15b expression vector (Novagen) containing a His tag. Point mutants were made using the QuickChange site-directed mutagenesis kit (Stratagene). For protein production, transformed

BL21(DE3) pLysS cells were grown in LB broth containing 100 mg ml<sup>-1</sup> carbenicillin to OD<sub>600</sub> = 0.6 and induced with 100 mM IPTG for 2.5–3.0 h (ref. 31). All proteins were soluble and were purified as described without removal of the His tag; mass spectrometry was used to confirm the identity of each mutant protein<sup>29</sup>.

**Biophysical analysis.** In all experiments, protein solutions were made in 50 mM sodium phosphate, pH 6.0, and experiments were carried out at 295 K. Equilibrium circular dichroism to determine protein stability and stopped-flow kinetic experiments were carried out as described<sup>30</sup>.

#### Acknowledgments

We thank K. O'Neil, L. Regan, and P. Alexander for providing constructs containing the protein G B1 domain. We also thank J. Tsai for Fig. 1b and members of the Baker lab for their helpful comments on the manuscript. This work was funded by a grant from the NIH.

Correspondence should be addressed to D.B. *email: dabaker@u.washington.edu*

Received 7 March, 2000; accepted 1 June, 2000.

- Alm, E. & Baker, D. *Curr. Opin. Struct. Biol.* **9**, 189–196 (1999).
- Chiti, F. *et al. Nature Struct. Biol.* **6**, 1005–1009 (1999).
- Martinez, J.C. & Serrano, L. *Nature Struct. Biol.* **6**, 1010–1016 (1999).
- Riddle, D.S. *et al. Nature Struct. Biol.* **6**, 1016–1024 (1999).
- Ternstrom, T., Mayor, U., Akke, M. & Oliveberg, M. *Proc. Natl Acad. Sci. USA* **96**, 14854–14859 (1999).
- Plaxco, K.W., Simons, K.T. & Baker, D. *J. Mol. Biol.* **277**, 985–994 (1998).
- Alm, E. & Baker, D. *Proc. Natl Acad. Sci. USA* **96**, 11305–11310 (1999).
- Munoz, V. & Eaton, W.A. *Proc. Natl Acad. Sci. USA* **96**, 11311–11316 (1999).
- Galzitskaya, O.V. & Finkelstein, A.V. *Proc. Natl Acad. Sci. USA* **96**, 11299–11304 (1999).
- Gu, H., Kim, D. & Baker, D. *J. Mol. Biol.* **274**, 588–596 (1997).
- Kim, D.E., Fisher, C. & Baker, D. *J. Mol. Biol.* **298**, 971–984 (2000).
- Gronenborn, A.M. *et al. Science* **253**, 657–661 (1991).
- Gallagher, T., Alexander, P., Bryan, P. & Gilliland, G.L. *Biochemistry* **33**, 4721–4729 (1994).
- Alexander, P., Orban, J. & Bryan, P. *Biochemistry* **31**, 7243–7248 (1992).
- Alexander, P., Fahnestock, S., Lee, T., Orban, J. & Bryan, P. *Biochemistry* **31**, 3597–3603 (1992).
- Orban, J., Alexander, P., Bryan, P. & Khare, D. *Biochemistry* **34**, 15291–15300 (1995).
- Park, S.H., O'Neil, K.T. & Roder, H. *Biochemistry* **36**, 14277–14283 (1997).
- Blanco, F.J., Rivas, G. & Serrano, L. *Nature Struct. Biol.* **1**, 584–590 (1994).
- Kuszewski, J., Clore, G.M. & Gronenborn, A.M. *Protein Sci.* **3**, 1945–1952 (1994).
- Munoz, V., Thompson, P.A., Hofrichter, J. & Eaton, W.A. *Nature* **390**, 196–199 (1997).
- Kobayashi, N., Honda, S., Yoshii, H. & Muneakata, E. *Biochemistry* **39**, 6564–6571 (2000).
- Park, S.H., Shastry, M.C. & Roder, H. *Nature Struct. Biol.* **6**, 943–947 (1999).
- Sheinerman, F.B. & Brooks, C.L., 3rd. *Proteins* **29**, 193–202 (1997).
- Sheinerman, F.B. & Brooks, C.L., 3rd. *J. Mol. Biol.* **278**, 439–456 (1998).
- Scalley, M.L., Nauli, S., Gladwin, S.T. & Baker, D. *Biochemistry* **38**, 15927–15935 (1999).
- Yi, Q., Michelle L. Scalley, Eric J. Alm & David Baker. *J. Mol. Biol. in the press* (2000).
- Kortemme, T., Kelly, M.J., Kay, L.E., Forman-Kay, J. & Serrano, L. *J. Mol. Biol.* **297**, 1217–1229 (2000).
- Wikstrom, M., Drakenberg, T., Forsen, S., Sjobring, U. & Bjorck, L. *Biochemistry* **33**, 14011–14017 (1994).
- Gu, H. *et al. Protein Sci.* **4**, 1108–1117 (1995).
- Scalley, M.L. *et al. Biochemistry* **36**, 3373–3382 (1997).
- O'Neil, K.T., Hoess, R.H., Raleigh, D.P. & DeGrado, W.F. *Proteins* **21**, 11–21 (1995).
- Kraulis, P.J. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
- Ferrin, T.E., Huang, C.C., Jarvis, L.E. & Langridge, R. *J. Mol. Graph.* **6**, 13–27 (1988).
- Huang, C.C., Pettersen, E.F., Klein, T.E., Ferrin, T.E. & Langridge, R. *J. Mol. Graph* **9**, 230–236, 242 (1991).
- Merritt, E.A. & Murphy, M.E.P. *Acta Crystallogr. D* **50**, 869–873 (1994).
- Bacon, D.J.A. & Anderson, W.F. *Mol. Graph.* **6**, 219–220 (1988).