

changes associated with ATP and GroES binding. GroEL provides binding surfaces for substrate proteins on a ring of highly mobile domains. □

Received 20 May; accepted 24 August 1994.

- Hendrick, J. P. & Hartl, F. U. *A. Rev. Biochem.* **62**, 349–384 (1993).
- Martin, J., Mayhew, M., Langer, T. & Hartl, F. U. *Nature* **366**, 228–233 (1993).
- Fisher, M. J. *biol. Chem.* **269**, 13629–13636 (1994).
- Schmidt, M., Buchner, J., Todd, M. J., Lorimer, G. H. & Viitanen, P. V. *J. biol. Chem.* **267**, 10304–10311 (1994).
- Kubota, H., Hynes, G., Carne, A., Ashworth, A. & Willison, K. *Curr. Biol.* **4**, 89–99 (1994).
- Hendrix, R. W. *J. molec. Biol.* **129**, 375–392 (1979).
- Hutchinson, E. G., Tichelaar, W., Hofhaus, G., Weiss, H. & Leonard, K. *EMBO J.* **8**, 1485–1490 (1989).
- Saibil, H. R. *et al. Curr. Biol.* **3**, 265–273 (1993).
- Langer, T., Pfeifer, G., Martin, J., Baumeister, W. & Hartl, F. U. *EMBO J.* **11**, 4757–4765 (1992).
- Braig, K., Simon, M., Furuya, F., Hainfeld, J. F. & Horwich, A. L. *Proc. natn. Acad. Sci. U.S.A.* **90**, 3978–3982 (1993).
- Ellis, R. J. & Hemmingsen, S. M. *Trends biochem. Sci.* **14**, 339–342 (1989).
- Goloubinoff, P., Christeller, J. T., Gatenby, A. A. & Lorimer, G. H. *Nature* **342**, 884–889 (1989).
- Nilsson, B. & Anderson, S. A. *Rev. Microbiol.* **45**, 607–635 (1991).
- Jackson, G. S. *et al. Biochemistry* **32**, 2554–2563 (1993).
- Zahn, R., Spitzfaden, C., Ottiger, M., Wüthrich, K. & Plückthun, A. *Nature* **368**, 261–265 (1994).
- Peralta, D., Hartman, D. J., Hoogenraad, N. J. & Høj, P. B. *FEBS Lett.* **339**, 40–45 (1994).
- Martin, J. *et al. Nature* **352**, 36–42 (1991).
- Mendoza, J. A., Lorimer, G. H. & Horowitz, P. M. *J. biol. Chem.* **266**, 16973–16976 (1991).
- Badcoe, I. G. *et al. Biochemistry* **30**, 9195–9200 (1991).
- Todd, M. J., Viitanen, P. V. & Lorimer, G. H. *Science* **265**, 659–666 (1994).
- Ishii, N., Taguchi, H., Sasabe, H. & Yoshida, M. *J. molec. Biol.* **236**, 691–696 (1994).
- Bochkareva, E. S. & Girshovich, A. S. *J. biol. Chem.* **267**, 25672–25675 (1992).
- Harris, J. R., Plückthun, A. & Zahn, R. *J. struct. Biol.* (in the press).
- Llorca, O., Marco, S., Carrascosa, J. L. & Valpuesta, J. M. *FEBS Lett.* **345**, 181–186 (1994).
- Schmidt, R. *et al. Science* **265**, 656–659 (1994).
- Azem, A., Kessel, M. & Goloubinoff, P. *Science* **265**, 653–656 (1994).
- Staniforth, R. A. *et al. FEBS Lett.* **344**, 129–135 (1994).

ACKNOWLEDGEMENTS. We thank J. Munn for EM support, R. Westlake and J. Bouquiere for computing support, R. J. Ellis and G. Laxer for discussion, the Birkbeck Photo Unit for photographic work and the Wellcome Trust and AFRC for project grants. A.R.C. is a Lister Institute Fellow.

## Context is a major determinant of $\beta$ -sheet propensity

Daniel L. Minor Jr\* & Peter S. Kim†

Departments of \* Chemistry and † Biology, Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA

**RESIDUES in  $\beta$ -sheets occur in two distinct tertiary contexts: central strands, bordered on both sides by other  $\beta$ -strands, and edge strands, bordered on only a single side by another  $\beta$ -strand<sup>1</sup>. The  $\Delta\Delta G$  values for  $\beta$ -sheet formation measured at an edge  $\beta$ -strand of the IgG-binding domain of protein G (GB1) are quite different from those obtained previously<sup>2,3</sup> at a central position in the same protein. In particular, there is no correlation at the edge position with statistically determined  $\beta$ -sheet-forming preferences<sup>4</sup>. The differences between  $\beta$ -sheet propensities measured at central and edge  $\beta$ -strands,  $\Delta\Delta\Delta G$  values, correlate with the values of water/octanol transfer free energies<sup>5</sup> and side-chain non-polar surface area for the amino acids<sup>6</sup>. These results strongly suggest that, unlike  $\alpha$ -helix formation,  $\beta$ -sheet formation is determined in large part by tertiary context, even at solvent-accessible sites, and not by intrinsic secondary structure preferences.**

The 20 naturally occurring amino acids were substituted at a solvent-exposed edge  $\beta$ -strand position, residue 44, by site-directed mutagenesis in a host molecule in which local inter-

TABLE 1 Parameters of  $\beta$ -sheet formation

Amino acid	$\Delta\Delta G$ (kcal mol <sup>-1</sup> )	$T_m$ (°C)	$K_a/K_a^{\text{AAA-44Thr}}$	$\Delta\Delta\Delta G$ (kcal mol <sup>-1</sup> )
Thr	0.83	60.2	1.0	-0.27
Ser	0.63	59.4	3.1	-0.07
Glu	0.31	57.0	1.5	0.30
Val	0.17	56.1	1.5	-0.65
Phe	0.16	56.1	1.1	-0.70
Tyr	0.11	55.9	1.1	-0.75
Cys	0.08	55.1	0.7	-0.44
Gln	0.04	54.7	1.9	-0.19
Ile	0.02	54.8	1.6	-0.98
Ala	0	54.7	2.9	0
His	-0.01	54.6	1.8	0.01
Met	-0.02	54.2	1.8	-0.74
Asp	-0.10	53.7	1.2	0.84
Trp	-0.17	53.3	3.0	-0.77
Asn	-0.24	52.6	1.0	-0.16
Leu	-0.24	52.8	1.0	-0.75
Lys	-0.40	51.4	0.9	-0.67
Arg	-0.43	51.2	1.0	-0.88
Gly	-0.85	47.6	1.4	0.35
Pro	< -4	< 0	0	—

Listed are  $\Delta\Delta G$  values for  $\beta$ -sheet formation at an edge position relative to alanine, thermal melting temperatures ( $T_m$ ) for the AAA-44Xaa proteins, relative binding constants ( $K_a$ ) to human Fc for the AAA-44Xaa proteins, and  $\Delta\Delta\Delta G$  values for  $\beta$ -sheet formation, comparing  $\Delta\Delta G_{\text{edge}} - \Delta\Delta G_{\text{centre}}$  values (see also ref. 2). The  $K_a$  for wild-type GB1 is  $1.4 \times 10^8 \text{ M}^{-1}$  (ref. 30). The  $K_a/K_a^{\text{AAA-44Thr}}$  value for wild-type GB1 is 7.1.  $\Delta G$  values for unfolding at 321 K were calculated as described previously<sup>2</sup> using data obtained from CD thermal unfolding measurements and the Gibbs-Helmholz equation. A positive  $\Delta\Delta G$  value indicates an increase in stability relative to alanine. Estimated errors in determination of  $T_m$  and  $\Delta G$  are  $\pm 0.5^\circ\text{C}$  and  $\pm 0.06 \text{ kcal mol}^{-1}$  respectively. Fc binding was measured as described previously<sup>2</sup> by adding variable amounts of competitor protein to a fixed quantity of protein G-alkaline phosphatase at  $5^\circ\text{C}$  in 96-well plates. The ratio of affinity constants for the mutants was normalized to AAA-44Thr. GB1-Thr1 (see Fig. 1 legend) was included as an internal standard on each plate.

actions to the guest site had been minimized by replacing the nearest neighbours with alanine (see Fig. 1 legend). This edge  $\beta$ -strand is bordered on one side by another  $\beta$ -strand and on the other side by solvent (Fig. 1a). The stability of each protein (denoted AAA-44Xaa) was measured by thermal unfolding as monitored by circular dichroism (CD) at 218 nm (Fig. 1b).  $\Delta\Delta G$  values for  $\beta$ -sheet formation, referenced to alanine, were obtained by assuming that changes in global stability result entirely from changes in the ability of the residue at the guest site to adopt a  $\beta$ -sheet conformation. In support of this assumption, molecules representative of the entire  $\Delta\Delta G$  range were found to have  $\Delta H_{\text{van't Hoff}}/\Delta H_{\text{cal}}$  ratios near unity, indicating that the two-state nature of the equilibrium unfolding transition observed for wild-type GB1 (ref. 7) remains intact (see Fig. 1 legend). The relative free energy differences for  $\beta$ -sheet formation at the edge position are listed in Table 1.

All the proteins tested, with the exception of unfolded AAA-44Pro, bind Fc with around sevenfold reduced affinity relative to wild-type GB1 (Table 1). There is some variation in binding constants but this is not correlated with protein stability. As residues 42–46 have been identified as participants in the Fc-binding interface<sup>8,9</sup> it seems likely that the observed affinity differences reflect direct effects on binding by substitutions at residue 44.

As a more detailed check on the conformation at the guest site of each molecule, the chemical shifts of the aromatic ring protons of Trp 43 were measured in each of the 20 variants. Trp 43 is expected to be sensitive to changes in structure as it immediately precedes the guest site and is part of the hydrophobic core of the molecule. The chemical shifts of the Trp 43 ring protons are similar in all the variants, with the exception

of unfolded AAA-44Pro, and are significantly different from the chemical shifts for free tryptophan (see Fig. 2 legend). Additionally, unambiguous cross-strand NOEs between the guest strand and its neighbouring strand can be found between upfield shifted  $H_{54}^{\zeta 1}$  and  $H_{54}^{\zeta 2}$  protons of Val 54 and the  $H_{43}^{\delta}$ ,  $H_{43}^{\zeta 2}$ ,  $H_{43}^{\zeta 3}$ ,  $H_{43}^{\epsilon 3}$ ,  $H_{43}^{\eta 2}$ ,  $H_{43}^{\eta 3}$  ring protons of Trp 43 in each folded protein.

Two molecules with substantially different stability—AAA-44Thr and AAA-44Ala—which also have different Fc-binding affinities, were characterized further by NMR. The  $^1H$ - $^{15}N$  correlation spectra of these two molecules are very similar except for resonances from residues in the immediate vicinity of the guest site (Fig. 2a, b). The NMR spectra of both proteins also contain cross-strand nuclear Overhauser effect (NOE) patterns near the guest site that are present in wild-type GB1 (Fig. 2c). Taken together with the Fc-binding and Trp 43 chemical shift results, these data indicate that any structural changes between the variants are small and do not involve significant disruption of the backbone  $\beta$ -sheet structure.

The dramatic effect that context can have on  $\beta$ -sheet propensities is seen in the comparison of the thermodynamic preferences for  $\beta$ -sheet formation measured at the central and edge  $\beta$ -sheet positions (Fig. 3a). Although the overall magnitude of the scale for  $\beta$ -sheet formation at the edge position ( $\sim 2$  kcal mol $^{-1}$ ,

excluding proline) is similar to that measured at the central  $\beta$ -sheet position $^{2,3}$ , there is no apparent correlation with the statistically measured  $\beta$ -sheet frequencies of Chou and Fasman $^4$  (Fig. 3b). In addition, unlike the results obtained at the central position, for which the  $\Delta\Delta G$  values are distributed fairly evenly over a wide range, more than half (13/20) of the  $\Delta\Delta G$  values measured at the edge position fall within a small range ( $\sim 0.4$  kcal mol $^{-1}$ ).

The context dependence of  $\beta$ -sheet formation suggests that two components contribute to  $\beta$ -sheet propensity: (1) intrinsic ability to form a local extended  $\beta$ -strand structure; and (2) ability to interact with the surrounding tertiary  $\beta$ -sheet structure. The small range of  $\beta$ -sheet propensities measured at the edge position, together with the relative lack of preference for particular side chain rotamers in  $\beta$ -strand residues $^{10}$ , suggests that (1) has only a minor role in determining  $\beta$ -sheet propensity. This situation contrasts with  $\alpha$ -helix formation where a significantly biased side chain rotamer distribution $^{10}$  and a well-distributed range of  $\alpha$ -helix propensity values are found $^{11-18}$ .

The difference in  $\beta$ -sheet propensity between the edge and centre sites, designated as  $\Delta\Delta G = \Delta\Delta G_{edge}^{Xaa} - \Delta\Delta G_{centre}^{Xaa}$ , correlates with both the free energy of transfer for the amino acids from octanol to water $^5$  and with the non-polar accessible

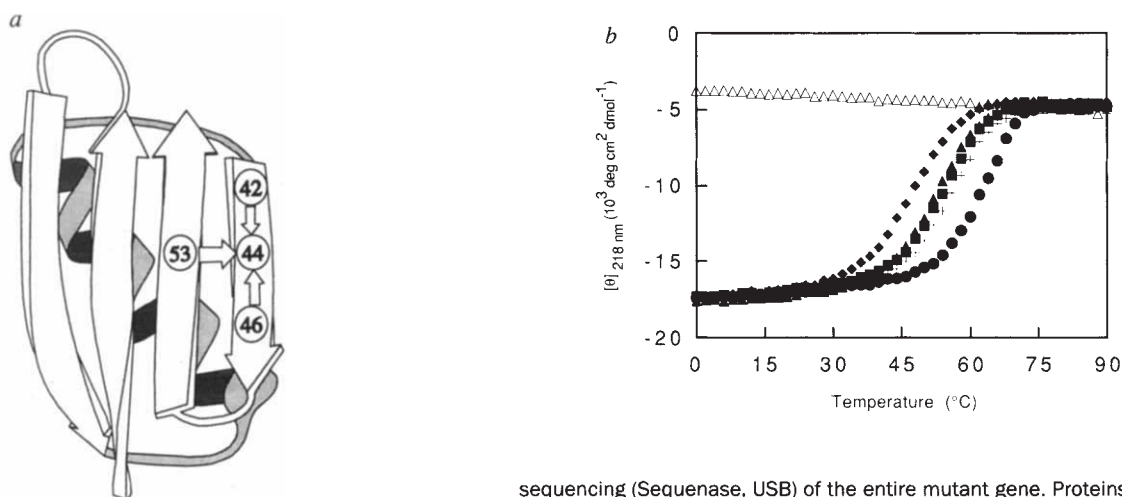


FIG. 1 a, Ribbon drawing $^{21}$  of GB1 based on the coordinate set 2GB1 $^{22}$ . The positions of the guest site and the surrounding residues are indicated. The arrows represent the inter-residue contacts to the guest site made by the surrounding residues. b, Temperature dependence of the CD signal from representative GB1 variants: ●, AAA-44Thr; +, AAA-44Val; ■, AAA-44Ala; ▲, AAA-44Asn; ◆, AAA-44Gly; and △, AAA-44Pro, in 150 mM NaCl, 50 mM Na-acetate, pH 5.4.

METHODS. As described previously $^2$ , major inter-residue contacts were identified $^{23}$  from the coordinate set 2GB1 and were defined as any atom pair having a centre-to-centre distance  $\leq 150\%$  of the sum of the van der Waals radii. Major contacts to residue 44 arise from residue 53 on the adjacent strand with fewer, more distant contacts being made by residues at positions  $i+2$  and  $i-2$  from the guest site. These residues were changed to alanine to create the host molecule E42A/D46A/T53A (denoted AAA). To test for possible effects from the  $i+2$  and  $i-2$  residues, which had been serine in our central position study $^2$ , we also created the host molecule E42S/D46S/T53A (denoted SSA). Substitution at the guest position with threonine, an amino acid expected to be a very good  $\beta$ -sheet former, and glycine, an amino acid expected to be a very poor  $\beta$ -sheet former, yields identical  $\Delta\Delta G$  values relative to alanine for  $\beta$ -sheet formation in both the AAA and SSA backgrounds. The host molecule bearing the smaller amino-acid substitutions, AAA, was chosen for further study. Recombinant GB1 mutants were expressed from a synthetic gene bearing the mutation Met 1 $\rightarrow$ Thr (GB1-Thr1) described previously $^2$ . Mutations were generated by single-strand mutagenesis $^{24}$  and were verified by dideoxyribonucleotide

sequencing (Sequenase, USB) of the entire mutant gene. Proteins were expressed in *Escherichia coli* (BL21 (DE3) pLysS) and were induced at an  $A_{600}$  of  $\sim 0.5$ – $0.8$  with a final concentration of isopropylthiogalactoside of 0.4 mM for 2–4 h. All folded proteins were purified by affinity chromatography $^7$  with IgG 6 Fast Flow Sepharose (Pharmacia) followed by reverse phase high pressure liquid chromatography (HPLC) purification on a Vydac preparative C18 column using a linear H $_2$ O-acetonitrile gradient in the presence of 0.1% TFA. The proteins are expressed as mixtures of N-terminally processed protein beginning with threonine at position 1 (56 residues) and non-processed protein (57 residues) beginning with methionine at position 0. The ratio of processed to unprocessed protein seems to depend on the stability of the molecule (data not shown). The HPLC purification step separates these two species and in all cases the 56-residue protein was used. For the unfolded molecule AAA-44Pro, the IgG affinity column step was replaced by G75 Sephadex chromatography in a buffer of 10 mM phosphate, 150 mM NaCl, pH 7.3. The identity of each HPLC purified protein was confirmed by laser desorption mass spectrometry (Finnigan Mat Laser-mat). All measured molecular weights were within 3 atomic mass units of the expected mass. CD measurements were made and thermal unfolding curves were fitted as described previously $^2$ .  $\Delta C_p$  values for unfolding (data not shown) were similar ( $\pm 15\%$ ) to those measured for the centre site AASS-Xaa molecules $^2$ . Protein concentration was determined by measuring absorbance of the unfolded protein $^{25}$ . Differential scanning calorimetry was performed with a Microcal MC-2 scanning calorimeter (Northampton, MA) as described previously $^2$ . AAA-44Thr, AAA-44Ser, AAA-44Gln, AAA-44Ala and AAA-44Gly were found to have  $\Delta H_{van't Hoff}/\Delta H_{cal}$  ratios of 0.98, 0.94, 1.03, 0.98 and 0.99, respectively.





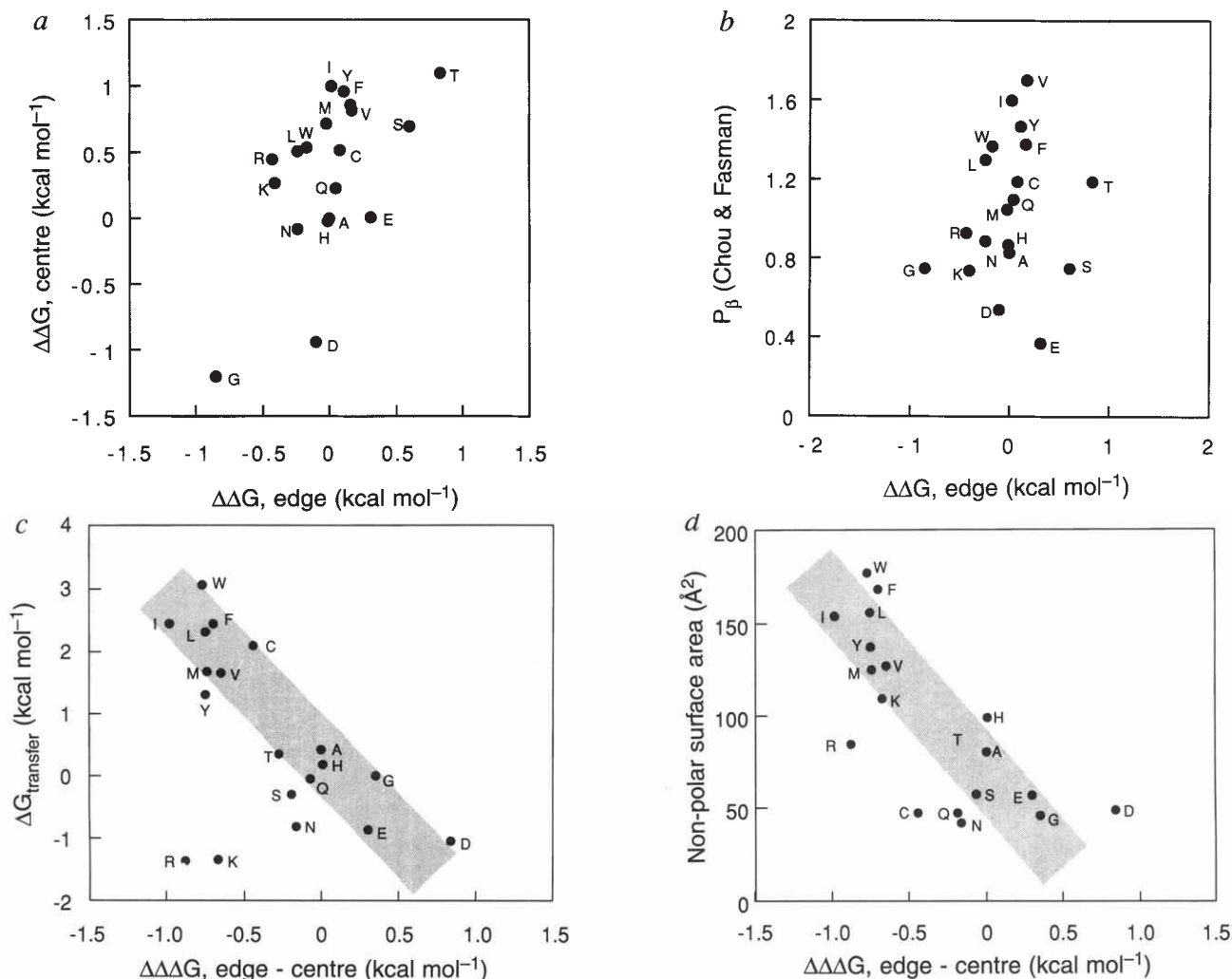


FIG. 3 *a*, Comparison of  $\beta$ -sheet forming propensities at central<sup>2</sup> and edge  $\beta$ -sheet positions. *b*, Comparison of the edge site  $\beta$ -sheet propensities with the  $\beta$ -sheet forming frequencies of Chou and Fasman<sup>4</sup>. The data are not correlated ( $r=0.15$ ,  $P>0.25$ ). *c*, Correlation of the difference in  $\beta$ -sheet forming propensity measured at edge and centre  $\beta$ -sheet positions,  $\Delta\Delta G$ , with  $\Delta G$  of transfer for the amino acid side chains from octanol to water<sup>5</sup> ( $r=0.57$ ,  $P<0.005$ ;  $r=0.80$ ,  $P<0.0005$ , excluding values of R and K, see methods) and *d*, side chain non-polar surface area<sup>6</sup> ( $r=0.72$ ,  $P<0.0005$ ). Amino acids are identified by the labels. Shaded areas are meant only to emphasize the overall trends in the data.

apparent  $\beta$ -sheet propensity measured in the zinc finger<sup>19</sup> and GB1 (refs 2, 3) model systems. More generally, our results emphasize that  $\beta$ -sheets are elements of both secondary and tertiary structure<sup>20</sup>. □

Received 13 June; accepted 2 August 1994.

- Chothia, C. *J. molec. Biol.* **105**, 1–14 (1976).
- Minor, D. L. Jr & Kim, P. S. *Nature* **367**, 660–663 (1994).
- Smith, C. K., Withka, J. M. & Regan, L. *Biochemistry* **33**, 5510–5517 (1994).
- Chou, P. Y. & Fasman, G. D. *Biochemistry* **13**, 211–222 (1973).
- Fauchere, J.-L. & Pliska, V. *Eur. J. med. Chem.-Chim. Ther.* **18**, 369–375 (1983).
- Livingstone, J. R., Spolar, R. S. & Record, M. T. Jr *Biochemistry* **30**, 4237–4244 (1991).
- Alexander, P., Fahnestock, S., Lee, T., Orban, J. & Bryan, P. *Biochemistry* **31**, 3597–3603 (1992).
- Frick, I.-M. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 8532–8536 (1992).
- Gronenborn, A. M. & Clore, G. M. *J. molec. Biol.* **223**, 331–335 (1993).
- McGregor, M. J., Islam, S. A. & Sternberg, M. J. E. *J. molec. Biol.* **198**, 295–310 (1987).
- Lyu, P. C., Liff, M. I., Marky, L. A. & Kallenbach, N. R. *Science* **250**, 669–673 (1990).
- O'Neil, K. T. & DeGrado, W. F. *Science* **250**, 646–651 (1990).
- Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T. M. & Baldwin, R. L. *Nature* **344**, 268–270 (1990).

METHODS. Values for amino acid side chain non-polar surface areas are for 'set 1' from ref. 14. The values from 'set 2' show a similar trend. With respect to the correlation between  $\Delta\Delta G$  and  $\Delta G$  of transfer there are two clear outlying points, arginine and lysine. The apparent anomalous behaviour of these two residues probably reflects the large distance between the side chain charge and the backbone  $\beta$ -sheet structure. Correlation coefficients were calculated for the  $(x_i, y_i)$  pairs using the formula<sup>29</sup>:

$$r^2 = \left[ \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\left[ \sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2 \right]^{1/2}} \right]^2$$

- Wojcik, J., Altman, K.-H. & Scheraga, H. A. *Biopolymers* **30**, 121–134 (1990).
- Chakrabarty, A., Schellman, J. A. & Baldwin, R. L. *Nature* **351**, 586–588 (1991).
- Horowitz, A., Matthews, J. M. & Fersht, A. R. *J. molec. Biol.* **227**, 560–568 (1992).
- Serrano, I., Neira, J.-L., Sancho, J. & Fersht, A. R. *Nature* **356**, 453–455 (1992).
- Blaber, M., Zhang, X. & Matthews, B. W. *Science* **260**, 1637–1640 (1993).
- Kim, C. A. & Berg, J. M. *Nature* **362**, 267–270 (1993).
- Lifson, S. & Sander, C. *J. molec. Biol.* **139**, 627–639 (1980).
- Priestle, J. P. *J. appl. Crystallogr.* **21**, 572–576 (1988).
- Gronenborn, A. M. et al. *Science* **253**, 657–661 (1991).
- Oas, T. G. & Kim, P. S. *Nature* **336**, 42–48 (1988).
- Kunkel, T. A., Roberts, J. D. & Zakour, R. A. *Meth. Enzym.* **154**, 367–382 (1987).
- Edelhoc, H. *Biochemistry* **6**, 1948–1954 (1967).
- Grathwohl, C. & Wüthrich, K. *Biopolymers* **20**, 2623–2633 (1981).
- McIntosh, L. P., Wand, A. J., Lowry, D. F., Redfield, A. G. & Dahlquist, F. W. *Biochemistry* **29**, 6341–6362 (1990).
- Wüthrich, K. *NMR of Proteins and Nucleic Acids* (John Wiley, New York, 1986).
- Rosner, B. *Fundamentals of Biostatistics* (Duxbury, Boston, 1982).
- Fahnestock, S. R., Alexander, P., Filipula, D. & Nagle, J. in *Bacterial Immunoglobulin-Binding Proteins* (ed. Boyle, M. D. P.) (Academic, San Diego, 1990).

ACKNOWLEDGEMENTS. We thank Z.-Y. Peng for helpful advice and suggestions and L. M. Mayr, M. G. Oakley and P. A. Pettilo critical reading of the manuscript. This work was supported by the Howard Hughes Medical Institute. D.L.M. is supported by a National Institutes of Health Biophysical Training Grant.