FEBS Letters 347 (1994) 304–309

FEBS 14188

Stability and solvation of Thr/Ser to Ala and Gly mutations at the N-cap of α-helices

Yu Wai Chen, Alan R. Fersht*

Centre for Protein Engineering, Medical Research Council Centre, Hills Road, Cambridge CB2 2QH, UK

Received 27 May 1994

Abstract

The solvation of polar groups at the N-terminal end of α-helices was studied by comparing the crystal structures of T4 lysozyme, barley chymotrypsin inhibitor 2 (CI2), barnase and their respective N-cap mutants. Whether or not the N3 residue is solvated on mutating the N-cap Thr/Ser to Ala or Gly appears to be related to the identities and the side-chain conformations of the N2 and N3 residues. When these two residues are alanines, as is in the pseudo-wild-type CI2 (E33A/E34A), the main-chain NH at the N3 position is exposed to the solvent and can be solvated. If the N2 residue is an Asp or a Glu, it is more likely that the side-chain of these residues will form a surrogate N-cap with the amide NH at N3 to compensate for the lost –OH group. In this case, no additional solvation will be observed. In general, Gly can be more stable than Ala at the N-cap because its small side-chain allows nearby polar groups to form hydrogen bonds with optimal geometry with solvent molecules or other polar groups.

Key words: α-Helix; N-cap; Solvation; Solvent accessibility; Hydrogen bond; Crystal structure

1. Introduction

The first and last four residues of α-helices differ from the rest by not being able to make the intra-helical hydrogen bonds between the backbone CO groups of one turn and the NH groups of the next [1]. There are statistical preferences at the N-cap position of an α-helix for the residues Asn, Ser, Asp, Gly and Thr. Conversely, Ala, Leu, Val, Ile, Trp, Arg, Gln and Glu are not preferred [2,3]. Apart from Gly, all the preferred residues have side-chain atoms that can form a hydrogen bond to the main-chain NH at the N3 position, thereby offering additional stability [2]. It has been postulated that these residues function as initiation and stop signals for secondary structure formation during the early stage in protein folding [1].

Recent advances in protein engineering have made possible the experimental verification of α-helix-forming propensities of various amino acids [4–7], including the relative stabilities of Ala versus Gly in various locations of an α-helix [8,9]. It is now generally established that Ala is a good residue for stabilizing an α-helix [10], except as capping residues at the ends [8,11]. The statistical preference for Gly over Ala at the N-cap position has been investigated by several groups using protein engineering experiments. Investigations on barnase, mutating the wild-type Thr residues to Ala (T6A and T26A) and to Gly (T6G and T26G), found that the Gly mutants are more stable than the Ala mutants by 0.5 to 1 kcal·mol^{-1} [12]. However, Gly and Ala have similar effects on stability at the N-caps of T4 lysozyme and the barley chymotrypsin inhibitor 2 (CI2) [13,14].

It has been proposed that the small side-chain of Gly provides extra stability by allowing water molecules to form optimal hydrogen bonds with polar groups at the N-terminus of the helix [9]. This hypothesis is supported by the findings in the crystal structures of the Ser → Gly31 (S31G/E33A/E34A) and Ser → Ala31 (S31A/E33A/E34A) mutants of CI2 [15]. A new water molecule is observed, compensating for the removed –OH of Ser at the wild-type N-cap, and solvating the main chain NH at the N3 position in both mutants. Further, the geometry of this hydrogen bond in the Ser → Gly mutant is indeed better than that in the Ser → Ala mutant. Crystallographic work on mutants of T4 lysozyme, however, showed contradicting results. No additional water molecule was found in both the Thr→Gly59 (T59G) and Thr→Ala59 (T59A) mutants near the N-terminus of the long α-helix [13].

In this article, we examine why the effect of stabilization by increased solvation is not consistently observed in the α-helices mentioned above. By studying the respective crystal structures and by using molecular modelling, we also attempt to discover the structural basis for the various stability measurements obtained from protein engineering.

2. Experimental

The structure of the barnase T26A mutant has been solved recently [16]. A pseudo-T26G mutant structure of barnase was generated by deleting the C^{β} atom from the co-ordinates of the T26A structure. Barnase wild-type structure is obtained from Cameron [17].

The other structures analyzed here were taken from the Protein Data

*Corresponding author. Fax: (44) (223) 402 140.

SSDI 0014-5793(94)00574-F

Bank, Brookhaven National Laboratory [18,19]. The entry identification codes for the CI2 protein structures [15] are: 1YPA (S31A mutant, also with E33A/E34A mutations); 1YPB (S31G mutant, also with E33A/E34A); and 1YPC (E33A/E34A pseudo-wild-type). The entry identification codes for the T4 lysozyme structures [13,20] are: 1LYH (T59G mutant, also with C54T/C97A); 1LYJ (T59A mutant, also with C54T/C97A) and 1L63 (C54T/C97A pseudo-wild-type). The backbone conformations of the α-helices of interest among all these structures are very similar to that of α-helices 2 of barnase. The r.m.s. derivation in main-chain atomic positions for the first 5 residues (N to N4) for any two of these α-helices does not exceed 0.25 Å. Therefore, the solvation pattern of all these proteins can be compared directly.

All structural analysis and molecular modelling were performed with the program, O [21]. Solvent contact surface areas were calculated with the program MS [22], using a probe radius of 1.4 Å. The solvent contact surface of an atom is the van der Waals surface of that atom that is in contact with the solvent probe; i.e. no re-entrant surface is included. Accessible surface areas of residues were calculated with the program DSSP [23].

3. Results

3.1. Solvent structure at the mutation site of barnase T26A mutant

There are no overall or local structural changes in protein atoms observed in the barnase T26A mutant structure, within the limit of error, compared to the wild-type structure [16]. The solvent structure at the N-terminal end of α-helix 2, of which residue 26 is the N-cap residue, is very well conserved and no new solvent molecule is observed in the vicinity of the mutation site. Specifically, there is no change in solvation for the main-chain N atoms of the N (N-cap), N1, N2 and N3 residues. This is consistent with the findings of Bell et al. for T4 lysozyme [13].

3.2. Solvent contact surface area calculations

We attempted to model a water molecule into the barnase T26A structure so that it occupies the equivalent position of the extra water in the S31A mutant of CI2 [15] and solvating the main-chain NH at N3 (Glu29). It was found that it is impossible for this water molecule to be positioned more closely than 3.5 Å from the N atom of the N3 residue, because of steric hindrance of the side-chain O^{γ} of Ser28 (N2 residue) and the C^{γ} of Glu29 (N3 residue). It follows that the main-chain NH at the N3 position of α-helix 2 in barnase T26A is simply not solvent-accessible.

The solvent accessibilities of the N atom at the N3 position for the wild-type and the N-cap mutant structures of CI2 and T4 lysozyme as well as for barnase were calculated (Table 1). It is found that the solvent contact surface areas of this atom for all the wild-type and N-cap mutants of both barnase and T4 lysozyme are near or equal to zero (Fig. 1A and B). On the other hand, the solvent contact surface area of the N atom at N3 of the two CI2 mutants increased from 0.6 $Å^2$ to more than 1.2 $Å^2$ on mutating the N-cap Ser31 to Ala and Gly (Table 1). The results of these calculations are consistent with the observations of solvation and lack of solvation at the N3 residue for these three proteins.

3.3. Geometry of the solvation hydrogen bond at the main-chain NH of the N3 residue

The geometry of the main-chain NH to water hydrogen bond at the N3 residue of the α-helix in CI2 was examined in detail. The location of the water molecule relative to the main-chain N atom was specified by the spherical polar co-ordinate system adopted by Thanki et al. [24] (Fig. 1A). The N-terminal region of the α-helix of CI2 S31A mutant is shown in Fig. 2A and B. The increase in solvent accessible molecular surface area in the S31A mutant (Table 1) is manifested by a larger range of θ ($\theta = 15°$ to 135°) available for solvation as illustrated in Fig. 2B. This larger range of solvent accessibility is made possible by the lack of side-chain atoms at the γ-position for both the N2 and the N3 residues. The C^{β} of Ala31 (N-cap) prevents the water from making a good linear hydrogen bond with the NH at the N3 position (Fig. 2A and B) [15]. On the other hand, the S31G mutant allows a hydrogen bond of optimal geom-

Table 1
Solvent accessible molecular contact surface area of the amide N-atom at the N3 position for barnase, T3 lysozyme, chymotrypsin inhibitor 2, and their respective N-cap mutants

Protein	Mutation at N-cap	Residue at N2	Residue at N3	Solvent contact area of NH at N3, $Å^2$
Barnase	Thr26 (WT)	Ser28	Glu29	0.0
	Thr → Ala26			0.0
	Thr → Gly26*			0.0
T7 Lysozyme	Thr59 (WT)	Asp61	Glu62	0.1
	Thr → Ala59			0.0
	Thr → Gly59			0.1
CI2	Ser31 (WT)	Ala33	Ala34	0.6
	Ser → Ala31			1.2
	Ser → Gly31			1.5

*Pseudo barnase T26G mutant structure was generated by deleting C^{β} of residue 26 from T26A mutant structure

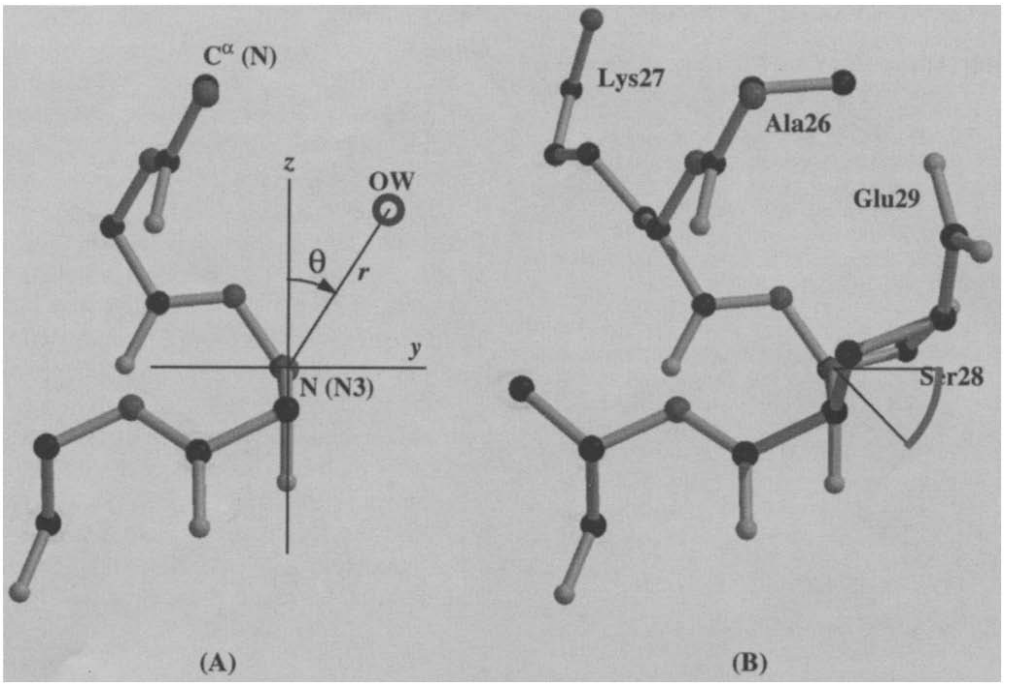

Fig. 1. Analysis of solvent accessibility near the N-cap of barnase α-helix 2. Shown here are the structures of the T26A mutant, oriented so that the peptide plane between the N2 and N3 residues lies vertical and perpendicular to the plane of the paper; i.e. it lies on the x–z plane. In this orientation, the C^{α} at N3 coincides with the C at N2; and the N at N3 coincides with the C^{α} of N2. The spherical polar co-ordinate system used is the one adopted by Thanki et al. [24]. (A) The backbone atoms are shown with the convention and symbols used to specify the location of a solvent molecule (OW), relative to the N atom of the N3 residue as origin. The y–z plane represents $\phi = \pm 90°$. (B) The solvent accessible molecule surface (grey arc) of the N atom at N3 for barnase wild-type and T26A mutant. The range of spherical angles describing this arc is $\theta = 90°$ to 135° and $\phi = +90°$. This geometry is not suitable for forming a hydrogen bond. The low solvent accessibility at this location is due to the steric hindrance of the side-chains of the N2 and N3 residues (Ser28 and Glu29). Similar results are obtained in the analysis of T4 lysozyme. The ranges of θ angles describing solvent accessibility of the N atom at N3 in the wild-type, T59A and T59G mutants of T4 lysozyme are similar to that illustrated here for barnase. This figure and Fig. 2 were generated with MOLSCRIPT [26] and RASTER3D [27, 28] and rendered in IRIS Showcase (Silicon Graphics, Inc., Mountain View, CA, USA) on a Silicon Graphics Indigo^2 workstation. Carbon atoms are represented by black spheres; nitrogen atoms are dark grey; oxygen atoms and water molecules are light grey.

etry to be formed (Fig. 2A). The absence of a C^{β} atom at the N position apparently offers the largest range of θ ($\theta = 0°$ to 135°, $\phi = +90°$; and $\theta = 0°$ to 15°, $\phi = -90°$; Fig. 2A and B).

3.4. *The effect of side-chains at the N2 and N3 residues on solvation of the amide NH at N3*

The CI2 pseudo-wild-type protein was designed such that both the N2 and N3 residues were truncated from a Glu to an Ala (E33A/E34A) to simplify the analysis, because Glu33 is disordered in the original wild-type structure [15,25]. We have reconstructed these two Glu side-chains on the S31A and S31G mutant structures by modelling, in order to study the effects of the side-chains at these positions on the solvation of the N atom at N3. We found that with both Glu33 and Glu34 reconstructed, certain side-chain conformations will obstruct the N atom at N3 from forming a reasonable solvation hydrogen bond. If the side-chain of the N2 residue has a torsion angle χ_1 in the range of −30° to +150° (180° coverage), then the γ-atom of this residue will block the hydration. The side-chain of N3 itself can also block the solvation of its own N atom, if the torsion angle χ_1 falls in the range of −75° to +15° (90° coverage). It is apparent that the identity of the amino acid at the N2 position is most important in determining the solvent accessibility of the main-chain N atom at N3 because the range of χ_1 of this residue that is unfavourable for hydration is much larger than that for the N3 residue (180° compared to 90°).

The amide NH site at N3 in barnase α-helix 2 is sterically hindered by the side-chain O^{γ} of the N2 residue (Ser28) and the side-chain, especially the C^{γ} atom, of the N3 residue (Glu29). The side-chain torsion angles (Table 2) of the N2 and N3 residues of both the wild-type and the T26A mutant fall into the range that blocks the N atom from solvation. Likewise, the N2 residue (Asp61) and the N3 residue (Glu62) in the two T4 lysozyme mutants have side-chain conformations that inhibit solvation of amide NH at N3. When compared with wild-type lysozyme, the side-chains of Asp61 of both T59A and T59G mutant rotate along the C^{α}–C^{β} bond (χ_1) so that the side-chain carboxylic oxygen atoms can form a surrogate N-cap with the amide NH of N3 [9,13].

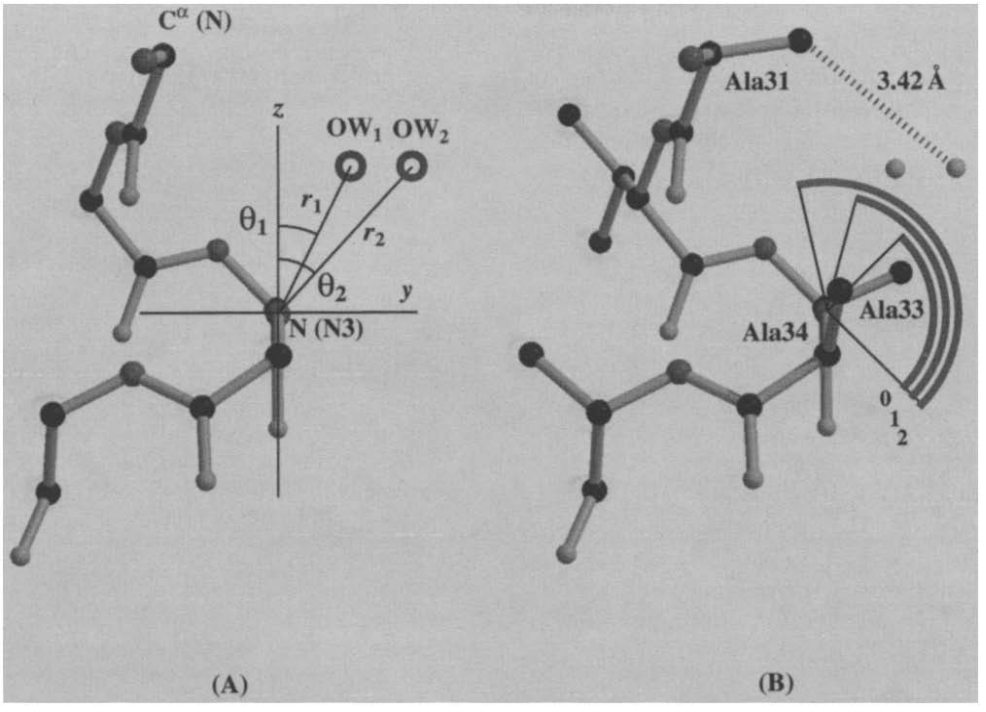

Fig. 2. Analysis of solvent accessibility near the N-cap of chymotrypsin inhibitor 2 (CI2). Shown here are the structures of the S31A mutant, oriented in the same way as in Fig. 1 so that the geometry of these two systems can be compared. The structure of wild-type CI2 and the S31G mutant are similar and superimposable to that of the S31A mutant. (A) The backbone atoms are shown with the locations of the water molecules found in the S31A and S31G mutants (Wat116 in the respective co-ordinate files), relative to the N atom of the N3 residue. OW_1 is the water molecule of the S31G mutant and OW_2 is that of S31A. The spherical polar co-ordinates of OW_1 are $r_1 = 2.74$ Å, $\theta_1 = 22.8°$, $\phi_1 = 84.6°$ and that for OW_2 are $r_2 = 3.28$ Å, $\theta_2 = 42.0°$, $\phi_2 = 89.6°$. (B) The solvent accessible surfaces (grey arcs) of the wild-type (denoted by 0), S31A (denoted by 1) and S31G (denoted by 2) compared. The ranges of spherical angles describing these arcs are: wild-type (arc 0): $\theta = 45°$ to 135°, $\phi = +90°$; S31A (arc 1): $\theta = 15°$ to 135°, $\phi = +90°$; and S31G (arc 2): $\theta = 0°$ to 135°, $\phi = +90°$and $\theta = 0°$ to 15°, $\phi = -90°$. The water molecule in the S31A structure cannot form a linear hydrogen bond because of steric hindrance of the C^{β} of N-cap Ala31. The higher solvent accessibility of the N atom at N3 position of the S31A and S31G mutants is due to the N2 and N3 residues both being an Ala (Ala33 and Ala34). This is to be compared with the N2 and N3 residues in Fig. 1.

4. Discussion

4.1. *Structural correlation of stability measurements at the N-terminus of α-helices*

As discussed in earlier studies [9,12,14], the effect on stability of Ala versus Gly at N-caps is a balance of several events. First, Gly is more destabilizing because on folding it loses the greatest amount of entropy. Secondly, Ala is not suitable as an N-cap residue because, instead of burying hydrophobic surface area in internal helical positions, it actually exposes more hydrophobic area than Gly does at the caps. Thirdly, polar groups that are hindered from solvation by the C^{β} may be exposed to the solvent if the N-cap is a Gly instead of an Ala. In stability studies of all these helices mentioned here, the delicate balance among these three factors is manifested by a range of stability measurements spanning from 0 to 0.5 kcal·mol^{-1}, with Gly being more stable than Ala.

We would like to discuss the structural basis for the relative preference of Ala versus Gly as the N-cap residue in the three helices for which stability data of N-cap mutations are available. The results of CI2 and lysozyme can be compared directly, because the N-cap of these two helices are exposed to solvent. The accessible surface areas of these residues are: Ser31 (CI2), 40 Å^2; and Thr59 (lysozyme), 67 Å^2, respectively. Barnase α-helix 2 is treated separately, since the N-cap residue (Thr26) is partially buried by packing against the β-sheet and it has an accessible surface area of 31 Å^2.

In a recent study of CI2 mutants, it was implicated that the N2 and N3 residues are not important in determining the stability of the helix [14], because the mutations of the N-cap Ser31 into either Gly or Ala give similar amount of destabilization, both with reference to the wild-type and to the pseudo-wild-type (E33A/E34A). However, our structural analysis here showed that the identities as well as the side-chain conformations of these two residues are important. We suggest that the wild-type and the pseudo-wild-type sets of mutants have different mechanisms of stabilizing the 'free' amide NH at N3, but making similar contributions. In the pseudo-wild-type set, both mutants are stabilized by an additional solvent hydrogen bond. In the wild-type set, the side-chain of Glu33 is most likely to rotate along χ_1 and forms a surrogate N-cap with the N atom of N3, in a

similar way to that which is observed in the lysozyme T59A and T59G mutant structures [13]. The long and flexible Glu side-chain allows this hydrogen bond to form with good geometry comparable to those observed in the pseudo-wild-type sets. Further, the conformation that Glu33 takes up to form a surrogate N-cap allows the negative charge of this residue to have better alignment with the helix dipole and thereby offering additional stability. Although the geometry of the hydrogen bond in the S31G/E33A/E34A mutant is better than that in the S31A/E33A/E34A mutant, the difference in stability is not obvious (0.05 kcal·mol^{-1}) [14]. Similarly, the stabilities of the S31A mutant and the S31G mutants are almost equal [14]. It follows from this argument that, the stability in the two sets being similar implies that the stabilization energy provided by a solvent molecule is similar in magnitude to that provided by a surrogate N-cap hydrogen bond offered by Glu33.

In the lysozyme mutants, a surrogate hydrogen bond is formed by the N2 residue (Asp61) with the N atom of N3 and has been discussed previously [9,13]. Although the side-chain atoms of the N2 residue have high *B*-factors (over 50 Å^2) [9], this surrogate N-cap does block the amide NH at N3 from solvation. The surrogate N-cap confers similar stabilization to both T59A and T59G mutants. Similar surrogate N-caps are frequently observed when Gly serves as the N-cap residue [3]. The similarity in stabilities of the two mutants is consistent with the findings in the CI2 mutant studies.

For barnase, the T26G mutant in α-helix 2 is more stable than the T26A counterpart, by 0.5 kcal·mol^{-1}. Assuming that there are no structural changes caused by the T26G mutation, we suggest that the differences in stabilities are caused by Gly allowing better hydrogen bonding geometry in the neighbourhood of the site of mutation. At the N-cap of α-helix 2, no solvation of amide NH at N3 is observed in the T26A mutant. We do not expect to see solvation for the T26G mutant neither. We observed that the C^β of Ala at the N-cap restricts some hydrogen bonds in its vicinity to form with optimal geometry. On mutating the N-cap residue 26 into a Gly, the nearby amide NH of Gly52 can make a hydrogen bond with much better geometry with Wat16 than those observed in the wild-type and the T26A structures.

Table 2
Side-chain conformation at the N2 and N3 positions for barnase, T4 lysozyme, and their respective N-cap mutants

Protein	Mutation at N-cap	χ_1 angle of residue at N2	χ_1 angle of residue at N3
Barnase	Thr26 (WT)	+75°	−45°
Barnase	Thr → Ala26	+70°	−65°
T7 Lysozyme	Thr59 (WT)	−80°	−80°
T7 Lysozyme	Thr → Ala59	+55°	−75°
T7 Lysozyme	Thr → Gly59	+45°	−80°

The ranges of χ_1 angle of that inhibit hydration of amide NH at N3 are: −30° to +150° for N2 residue and −75° to +15° for N3 residue. There are no χ_1 angles for the CI2 crystal structures because it has Ala at both N2 and N3.

4.2. Conclusions

We found that contradicting solvation observations of N-cap mutations can be satisfactorily accounted for by the solvent accessibility of the amide NH atom of the N3 residue. We also suggested how stability measurements from different protein engineering experiments can be correlated with the structural changes. The three helices we examined in this study apparently represent slight variations of a single theme: that relative stability of Gly versus Ala at the N-caps can depend on whether the removal of the side-chain C^β of Ala leads to nearby polar groups forming extra hydrogen bonds with the solvent, or having existing hydrogen bond geometry improved. Detailed study of structures also revealed that improved hydrogen bond formation of polar groups is not limited to the free amide NH's on the helix: it can be the neighbouring atoms if the N-cap is buried.

Acknowledgements: Y.W.C. is supported by a Croucher Foundation Fellowship. The atomic co-ordinates and the structure factors of the barnase T26A mutant have been deposited in the Brookhaven Protein Data Bank.

References

[1] Presta, L.G. and Rose, G.D. (1988) Science 240, 1632–1641.
[2] Richardson, J.S. and Richardson, D.C. (1988) Science 240, 1648–1652.
[3] Dasgupta, S. and Bell, J.A. (1993) Int. J. Peptide Protein Res. 41, 499–511.
[4] Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T.M. and Baldwin, R.L. (1990) Nature 344, 268–270.
[5] Horovitz, A., Matthews, J.M. and Fersht, A.R. (1992) J. Mol. Biol. 227, 560–568.
[6] Blaber, M., Zhang, X.-J. and Matthews, B.W. (1993) Science 260, 1637–1640.
[7] Blaber, M., Zhang, X.-J., Lindstrom, J.D., Pepiot, S.D., Baase, W.A. and Matthews, B.W. (1994) J. Mol. Biol. 235, 600–624.
[8] Strehlow, K.G. and Baldwin, R.L. (1989) Biochemistry 28, 2130–2133.
[9] Serrano, L., Sancho, J., Hirshberg, M. and Fersht, A.R. (1992) J. Mol. Biol. 227, 544–559.
[10] Marqusee, S., Robbins, V.H. and Baldwin, R.L. (1989) Proc. Natl. Acad. Sci. USA 86, 5286–5290.
[11] Serrano, L., Neira, J.-L., Sancho, J. and Fersht, A.R. (1992) Nature 356, 453–455.
[12] Serrano, L. and Fersht, A.R. (1989) Nature 342, 296–299.
[13] Bell, J.A., Becktel, W.J., Sauer, U., Baase, W.A. and Matthews, B.W. (1992) Biochemistry 31, 3590–3596.
[14] elMasry, N.F. and Fersht, A.R. (1994) Protein Eng. 7, in press.
[15] Harpaz, Y., elMasry, N., Fersht, A.R. and Henrick, K. (1994) Proc. Natl. Acad. Sci. USA 91, 311–315.
[16] Chen, Y.W. (1994) Structural Studies of Barnase Mutants, Ph.D. Thesis, University of Cambridge, UK.
[17] Cameron, A.D. (1992) Structural Studies on Myoglobin and Barnase, Ph.D. Thesis, University of York, UK.

[18] Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, J. E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) J. Mol. Biol. 112, 535–542.
[19] Abola, E.E., Bernstein, F.C., Bryant, S.H., Koetzle, T.F. and Weng, J. (1987) in: Crystallographic Databases – Information Content, Software Systems, Scientific Applications (Allen, F.H., Bergerhoff, G. and Sievers, R., Eds.) pp. 107–132, Data Commission of the International Union of Crystallography, Bonn, Cambridge, Chester.
[20] Nicholson, H., Anderson, D.E., Dao, P.S. and Matthews, B.W. (1991) Biochemistry 30, 9816–9828.
[21] Jones, T.A. and Kjeldgaard, M. (1993) O. Uppsala University, Sweden.
[22] Connolly, M.L. (1983) Science 221, 709–713.
[23] Kabsch, W. and Sander, C. (1983) Biopolymers 22, 2577–2637.
[24] Thanki, N., Umrania, Y., Thornton, J.M. and Goodfellow, J.M. (1991) J. Mol. Biol. 221, 669–691.
[25] McPhalen, C.A. and James, M.N.G. (1987) Biochemistry 26, 261–269.
[26] Kraulis, P. (1991) J. Appl. Crystallogr. 24, 946–950.
[27] Bacon, D.J. and Anderson, W.F. (1988) J. Mol. Graphics 6, 219–220.
[28] Merritt, E.A. and Murphy, M. (1994) in preparation.