

## Discrepancies between the NMR and X-ray Structures of Uncomplexed Barstar: Analysis Suggests That Packing Densities of Protein Structures Determined by NMR Are Unreliable<sup>†,‡</sup>

Girish S. Ratnaparkhi,<sup>§</sup> S. Ramachandran,<sup>||</sup> Jayant B. Udgaonkar,<sup>||</sup> and R. Varadarajan<sup>\*,§,⊥</sup>

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India, Chemical Biology Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur P.O., Bangalore 560 004, India, and National Centre for Biological Sciences, TIFR Centre, Indian Institute of Science Campus, Bangalore 560 012, India

Received November 21, 1997; Revised Manuscript Received December 30, 1997

**ABSTRACT:** The crystal structure of the C82A mutant of barstar, the intracellular inhibitor of the *Bacillus amyloliquefaciens* ribonuclease barnase, has been solved to a resolution of 2.8 Å. The molecule crystallizes in the space group  $I4_1$  with a dimer in the asymmetric unit. An identical barstar dimer is also found in the crystal structure of the barnase–barstar complex. This structure of uncomplexed barstar is compared to the structure of barstar bound to barnase and also to the structure of barstar solved using NMR. The free structure is similar to the bound state, and there are no significant main-chain differences in the 27–44 region involved in barstar binding to barnase. The C82A structure shows significant differences from the average NMR structure, both overall and in the binding region. In contrast to the crystal structure, the NMR structure shows an unusually high packing value based on the occluded surface algorithm, indicating errors in the packing of the structure. We show that the NMR structures of homologous proteins generally show large differences in packing value, while the crystal structures of such proteins have very similar packing values, suggesting that protein packing density is not well determined by NMR.

Barstar is a small protein of 89 residues existing in the bacterium *Bacillus amyloliquefaciens*. It functions as the intracellular inhibitor of barnase, a ribonuclease found in the same organism (1, 2). It has been shown in vitro that barnase and barstar form a 1:1 complex. The dissociation constant for the formation of the complex is  $10^{-14}$  M (3). This system has served as a typical model to understand protein–protein interactions.

Barstar is one of the two proteins for which information is available on ribonuclease inhibitor structure, the other being the mammalian ribonuclease inhibitor (4). This inhibitor binds mammalian ribonucleases tightly ( $10^{-14}$ – $10^{-16}$  M<sup>-1</sup>), and the structure of free as well as complexed inhibitor is known (5, 6). In the case of barstar, the free structure is known from NMR methods (7), whereas the complexed structure is known from crystallographic studies (8, 9). The interaction between barnase and barstar has been studied by using protein engineering (10), NMR (11), X-ray crystallography (8, 9), and thermodynamic techniques (12).

Barstar has also been used as a model protein to understand the effects of structure on folding and stability. An analysis of the refolding of barstar in 1 M guanidine hydrochloride

(GuHCl)<sup>1</sup> has shown that an initial hydrophobic collapse precedes the formation of secondary structure (13, 14). Multiple intermediates and transition states have been detected on the unfolding pathway of the protein (15, 16). Several mutagenesis as well as chemical modification studies have been carried out on this protein (14, 17, 18). A proper interpretation of these results requires the availability of the structure of the protein in its native state.

Although crystallization reports are available in the literature (19, 20), both for uncomplexed wild type as well as the double mutant (C40A:C82A), a crystal structure of barstar in the absence of barnase has yet to be reported. An unpublished structure of *wt* barstar is available from Mauguen's group (2). Here, we report the crystal structure of the C82A mutant of barstar in its free form, that is, uncomplexed to barnase. The C82A mutant of barstar has activity (3) and thermal stability similar to those of *wt* barstar (21) with a  $\Delta G^\circ$  (25 °C) of 25 kJ mol<sup>-1</sup> and is marginally destabilized to GuHCl denaturation (18) compared to *wt* barstar. The C82A structure shows significant differences from the solution structure of the protein reported previously (7, 22).

<sup>†</sup> Financial support was from DST Grant SP/SO/D-21/93 and CSIR Grant 37(913)96-EMR-II. G.S.R. is a KS Krishnan Fellow, DAE.

<sup>‡</sup> The coordinates as well as the structure factors have been deposited in the Protein Data Bank with the file names 1a19 and r1a19sf, respectively.

<sup>\*</sup> Corresponding author: Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012. E-mail: varadar@mbu.iisc.ernet.in. Fax: 91-80-3348535 or 91-80-3341683.

<sup>§</sup> Molecular Biophysics Unit, Indian Institute of Science.

<sup>||</sup> National Centre for Biological Sciences.

<sup>⊥</sup> Jawaharlal Centre for Advanced Scientific Research.

<sup>1</sup> Abbreviations: C82A, crystal structure of the Cys82Ala mutant of barstar at 2.8 Å solved in this study; C40A:C82A, barstar double mutant; *wt*, wild-type barstar; rmsd, root-mean-square deviation; *B*-factor, crystallographic temperature factor; MC, main chain; SC, side chain; SA, simulated annealing; 1BTA, the uncomplexed NMR structure (7); 1BGS, 2.6 Å structure of barstar C40A:C82A complexed to barnase (9); 1BRS, identical to 1BGS except at a higher resolution of 2.0 Å (8); OS, occluded surface;  $\Delta C_p$ , change in excess heat capacity upon unfolding; DTNB, 5,5'-dithiobis(2-nitrobenzoic acid); TNB, 5-thio-2-nitrobenzoic acid; GuHCl, guanidine hydrochloride.

## MATERIALS AND METHODS

**Crystallization and Data Collection.** The expression, purification, and chemical modification of barstar have been carried out as previously reported (18) except that the protein was further purified by FPLC on a Superose-6 gel filtration column. We attempted to crystallize the DTNB-labeled barstar (18) along with the C40A and C82A mutants, as the chemical modification of barstar stabilizes these proteins in solution and reduces aggregation due to the absence of free thiol groups. However, the C40A and DTNB-labeled crystals did not diffract well and were not stable in the X-ray beam. Crystals of barstar were grown at 20 °C using the hanging drop method (23) under conditions similar to those used previously for *wt* barstar (20). To crystallize barstar, 2  $\mu$ L of a stock solution of protein [45 mg/mL based on  $\epsilon = 23\,000\text{ M}^{-1}\text{ cm}^{-1}$  in 50 mM phosphate, pH 6.5 (24)] was mixed with an equal volume of the precipitant (40–60% ammonium sulfate and 50 mM phosphate, pH 6.5). Crystals were obtained in 24–48 h and reached their maximum size in 4–7 days. The protein was dissolved to its maximum concentration in 50 mM Tris, pH 8.0, and then the Tris was replaced with 50 mM phosphate, pH 6.5, in a Centricon concentrating unit. The crystals were sealed in a glass capillary with a small amount of mother liquor. X-ray diffraction data was collected on these crystals at 20 °C using a MAR imaging plate detector system attached to a Rigaku RU200 rotating anode equipped with a 0.2-mm focusing cup. The X-ray source was Cu K $\alpha$  radiation, and the crystal to detector distance was 100 mm. Data sets were collected by recording 1.0° oscillation images. The anode was operated at 50 mA and 40 kV, and each frame was exposed for 210 s. Diffraction data were recorded over an oscillation range of 250°. The data set was processed using the XDS (25) suite of programs.

A single crystal (1.0  $\times$  0.5  $\times$  0.4 mm<sup>3</sup>) was used to collect a complete data set for the barstar C82A mutant to 2.76 Å resolution at 20 °C. The crystal indexed in the space group  $I4_1$  with cell parameters  $a = b = 104.18\text{ Å}$ ,  $c = 36.00\text{ Å}$ . The space group obtained has not been reported in the crystallization reports published previously (19, 20). Assuming a solvent content of the unit cell of approximately 50%, we expected 2 molecules in the asymmetric unit. We have been able to obtain a solution for the orientation and position of the C82A mutant in the unit cell using the molecular replacement program AMoRe (26) with the X-ray structure of the C40A:C82A barstar mutant complexed to barnase, 1BGS (9), as a starting model. The best solution had a correlation coefficient of 65% and an  $R$ -factor of 40% (after rigid body refinement; resolution range, 15–3.5 Å). This was significantly better than the next solution. A similar solution was obtained using 1BRS (8) as a starting model. Initially, the NMR structure of free barstar, 1BTA (10), was also used as a starting model for AMoRe calculations, but it did not give a clear solution (correlation, 24–28%;  $R$ -factor, 48–50%) despite testing various resolution ranges and using only rigid units such as  $\beta$ -sheets or  $\alpha$ -helices in the calculations. Table 1 summarizes the data collection, reduction, molecular replacement, and refinement of the C82A mutant.

**Refinement.** Refinement was carried out by simulated annealing (SA) torsion angle refinement (27, 28) using X-PLOR (29). Torsion angle refinement is more powerful

Table 1: Summary of X-ray Data Reduction and Refinement

|  |  |
|--|--|
| data collection                          |  |
| space group                              | $I4_1$   |
| unit cell                                | $a = b = 104.182\text{ Å}$ ;<br>$c = 36.00\text{ Å}$ |
| molecules per asymmetric unit            | 2(dimer)   |
| resolution                               | 2.76 Å   |
| unique reflections ( $I/\sigma(I) > 0$ ) | 4796   |
| unique reflections ( $I/\sigma(I) > 2$ ) | 3706   |
| multiplicity of data set                 | 7.0  |
| completeness                             | 99.0% (98.9%)  |
| $R_{\text{merge}}^a$                     | 7.6%   |
| molecular replacement                    |  |
| model, 1BGS (15–3.5 Å range)             | correlation = 65%;<br>$R$ -factor = 30%              |
| refinement                               |  |
| crystallographic $R$ -factor             | 20.0%  |
| $R_{\text{free}}^b$                      | 29.2%  |
| range used for refinement                | 10.0–2.76 Å  |
| completeness <sup>c</sup>                | 2.9 Å  |
| rmsd bonds                               | 0.008 Å  |
| rmsd angles                              | 1.12°  |
| rmsd dihedrals                           | 26.39°   |
| rmsd impropers                           | 0.5°   |
| av $B$ -factor <sup>d</sup>              |  |
| main-chain atoms                         | 46   |
| side-chain atoms                         | 48   |

<sup>a</sup>  $R_{\text{merge}}$  is defined as  $(\sum_{hkl}|I_i - \langle I \rangle|/\sum_{hkl} I_i) \times 100.0$ , where  $I_{hkl}$  is the intensity of reflection  $hkl$  and  $\langle I \rangle$  is the average of all measurements for reflection  $hkl$ . <sup>b</sup>  $R_{\text{free}} = \sum_{hkl}|F_o| - |F_c|/\sum_{hkl}|F_o|$  for a test set of 15% of all reflections. <sup>c</sup> Completion = 50% in the 2.9–3.0 Å shell for  $3\sigma$  reflections. <sup>d</sup> The average  $B$ -factors refer to chain A of the two molecules in the asymmetric unit.

than conventional simulated annealing. It has a larger radius of convergence and reduces the ratio of free parameters to observed data (30). Fifteen percent of all reflections were set aside in order to calculate  $R_{\text{free}}$  (30–32). However, in the final stage of refinement, all reflections were used. The starting  $R_{\text{free}}$  and  $R$  values were 45% and 42%, respectively. Refinement was started with a model assuming restrained NCS symmetry (33) in the initial stages of refinement. Both molecules in the asymmetric unit were considered similar, and a high NCS restraint weight of 300 kcal mol<sup>-1</sup> Å<sup>-1</sup> was applied. This weight was decreased by steps of 50 during the course of refinement and was set at 100 kcal mol<sup>-1</sup> Å<sup>-1</sup> at the end of SA refinement. Various refinement protocols were tried, and it was found that SA torsional angle refinement reduced the  $R_{\text{free}}$  value more than the conventional SA protocol. Refinement steps also included Powell minimization,  $B$ -factor refinement, and constant temperature torsion angle refinement (500 K) followed by manual rebuilding.  $2F_oF_c$  as well as  $F_oF_c$  difference Fourier maps were calculated with X-PLOR using  $\sigma_A$ -weighting (34), and manual rebuilding was done using the package FRODO (35). Water addition (10 waters) did not improve the  $R_{\text{free}}$  significantly, and waters were not included in the final structure. Both grouped  $B$ -factor refinement and individual restrained isotropic  $B$ -factor refinement with the default restraints (29) for atoms forming pairs (bonds) and triplets (angles) were used. The  $B$ -factors between atoms related by NCS-symmetry were restrained. The individual restrained  $B$ -factor protocol gave the lowest  $R_{\text{free}}$  values. Model bias in the structure was also minimized by starting refinement with the NMR structure (the starting model was the NMR structure, superimposed on the molecular replacement solution found using 1BGS) and confirming that this structure

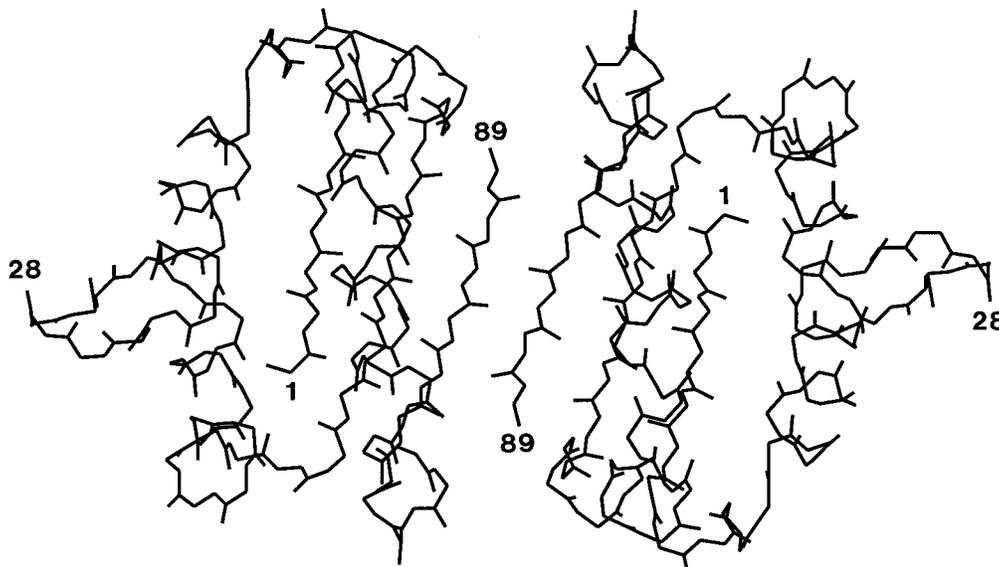


FIGURE 1: Main-chain trace of the dimer in the asymmetric unit.

after complete refinement was similar to the final refined structure obtained using the X-ray structure of barstar (1BGS) as a starting model. Model bias in the binding region (27–44) was minimized by rebuilding the region after SA  $\sigma_A$ -weighted omit map calculations in X-PLOR. The last stage of refinement used all reflections and included constant temperature torsion angle refinement (500 K), positional refinement, and *B*-factor refinement with an NCS weight of 100 kcal mol<sup>-1</sup> Å<sup>-1</sup>. Removing all NCS restraints did not decrease  $R_{\text{free}}$  significantly, and therefore the monomers in the asymmetric unit were restrained in the final refined structure.

**Analysis.** All calculations were done in X-PLOR (29). The structures were analyzed using rmsd and B-Factor plots for main-chain as well as side-chain residues. Parameters such as accessibility (36) and packing value using the occluded surface (OS) algorithm (37) were examined. The mean protein packing value was reported on the basis of the normalized packing value for each residue (38). Interactions between barnase and barstar as well as between individual monomers of the dimer in the asymmetric unit were analyzed by using the INTCHOS program (a part of the OS suite of programs) to calculate interchain occluded surfaces as well as by using the CONTACT algorithm, which is part of the CCP4 (39) suite of programs. All of the analyses were carried out on chain A, as it was similar to chain B due to NCS restraints. All structural results were confirmed by visual inspection of  $2F_oF_c$  and  $F_oF_c$  maps. Omit maps ( $\sigma_A$ -weighted; omit maps 29, 34) were calculated to confirm important structural results.

## RESULTS AND DISCUSSION

**Description of the Final Structure.** The final structure has a crystallographic *R*-factor of 20% and an  $R_{\text{free}}$  value of 29.2% and contains all 89 residues without any deletions. The mean coordinate error calculated by a Luzatti plot is 0.30 Å for the experimental data set and 0.40 Å for the 15% test reflections. The cysteine at position 40 shows clear density and does not show any additional density that would indicate a mixed disulfide. The presence of the Cys82 to Ala mutation is confirmed by  $2F_oF_c$  density as well as the

lack of  $F_oF_c$  density. Omit SA maps were also used to confirm the C82A mutation in the structure. Although mass spectrometry (21) shows that an N-terminal methionine is attached to the polypeptide chain, it has not been included in the structure, since its inclusion is not confirmed by electron density and omit map calculations. The asymmetric unit contains two molecules of barstar which form an antiparallel dimeric unit with the beta strands of the C-termini of both molecules aligned with each other. The two monomers in the dimeric unit are related by a 2-fold axis of symmetry (Figure 1). Both chains in the asymmetric unit are similar due to the NCS restraints imposed on them (the main-chain rmsd of chain B with A is 0.21 Å). Hence, only chain A is used in further analysis. The electron density map for both molecules in the asymmetric unit is of good quality for a map at this resolution due to the presence of NCS. The  $2F_oF_c$  map has breaks in the main chain in the stretch of residues from 58 to 62, but no deletions have been made from the final model (residues 64 and 65 have been deleted in 1BRS). The stretch of residues from 6 to 9 shows weak electron density, especially around glycine 7. Despite these breaks and the fact that these regions show small deviations between molecules in the asymmetric unit, these regions have been treated with restraints similar to those used with the rest of the structure. Treatment of the loop and flexible regions with smaller (e.g., 5–50 kcal mol<sup>-1</sup> Å<sup>-1</sup>) NCS restraints during SA refinement does not significantly decrease  $R_{\text{free}}$  or improve the map. The final structure has excellent stereochemistry, with the deviations from ideal bond lengths and bond angles being 0.008 Å and 1.3°, respectively. There are no violations for bonds, angles, dihedrals, or impropers on the basis of the default threshold values in X-PLOR. The structure has been analyzed using PROCHECK (40) and shows that 80% of the residues have backbone  $\phi, \psi$  angles in the core regions of the Ramachandran plot and 20% lie in the allowed regions. None of the residues have generously allowed or disallowed  $\phi, \psi$  angles. The main-chain and side-chain stereochemistries are better than that derived in the database of high-resolution structures used by PROCHECK, and the structure has an overall *G*-factor of 0.1.

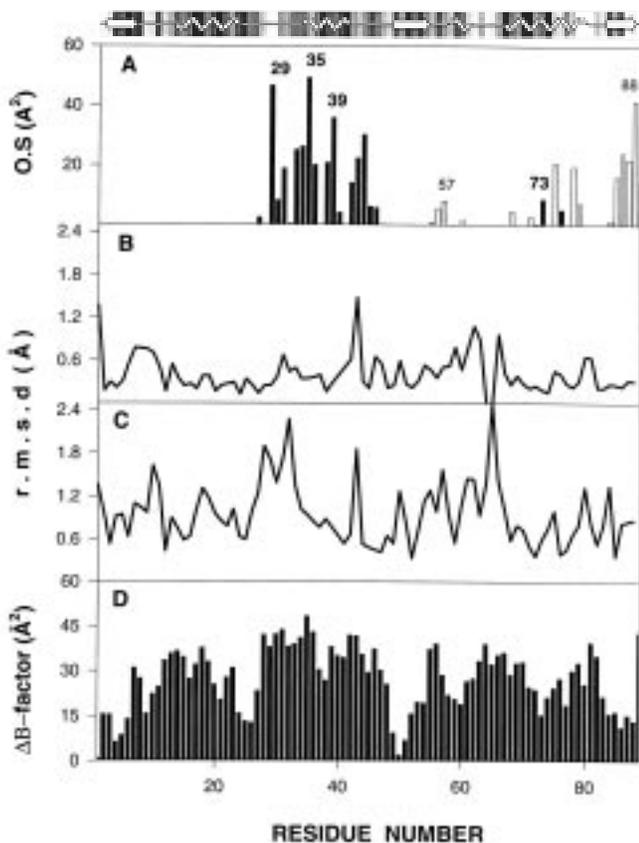


FIGURE 2: (A) Residues which occlude each other in either the dimer interface (empty bars) or the barnase–barstar (1BRS) interface (filled bars). The occluded surface was calculated by the OS algorithm (37). (B, C) Main-chain rmsd plots relative to the C82A crystal structure of 1BRS (crystal structure of complexed barstar) (B) and 1BTA (NMR structure) (C). In each case, the structure was superimposed on the C82A structure before calculating the rmsd. (D)  $\Delta B$ -Factor plot of C82A relative to the 1BGS structure. The secondary structure and accessibility representation at the top of the figure is an output of PROCHECK (40).

Table 2: List of Hydrogen Bonds between the Monomers of the Dimer in the Asymmetric Unit<sup>a</sup>

| chain A | chain B | distance (Å) |
|---------|---------|--------------|
| E68 OE1 | R75 NH2 | 3.03         |
| R75 NH2 | E68 OE1 | 2.88         |
| R75 NH2 | R75 NH2 | 3.21         |
| I86 N   | L88 O   | 3.03         |
| L88 N   | I86 O   | 2.73         |
| L88 O   | I86 N   | 3.00         |
| S89 OT1 | K78 NZ  | 2.66         |

<sup>a</sup> The list was generated using the program CONTACT, from the CCP4 (39) suite of programs.

**The Dimer in the Asymmetric Unit.** The asymmetric unit contains two monomers which are in a dimeric association. The two monomers are in contact via the  $\beta$ -strand at the C-terminus of barstar (Figure 1). Figure 2A shows all the residues in contact (empty bars) between the two monomers, based on interchain OS. The major contribution for the dimer is from the C-terminal  $\beta$ -sheet and includes three main-chain hydrogen bonds (Table 2). In addition, there are contributions from helix-4 and the region around Glu 57. Table 2 lists the major hydrogen bonds which are part of the dimeric interface. Figure 2A also shows the residues occluded in the barnase–barstar interface (filled bars). The

residues involved in the barnase–barstar interface and the dimer interface are different sets of residues and do not overlap. This was confirmed by the absence of short contacts on superposition of two barnase–barstar molecules on the dimer. Thus, if a dimeric unit exists in solution, in principle, it should be capable of simultaneously inhibiting two barstar molecules. We generated the symmetry-related molecules in 1BRS and, surprisingly, found that the dimer in the C82A asymmetric unit is present in the 1BRS (and therefore the 1BGS) barnase–barstar complex structure also. The 1BRS dimer superimposed well on the C82A dimer and gave a  $C_{\alpha}$  rmsd of 0.72 Å. Despite the dimer being found in the crystal structures of both free and bound barstar, there is no evidence for barstar forming dimers in solution. Under oxidizing conditions, a small amount of aggregation is observed (20), but, this has been shown to be caused by the formation of intermolecular disulfide bonds. The presence of the dimer in the crystal structures of free and complexed barstar may be an artifact of the extremely high protein concentration in the crystals.

**Comparison of C82A Crystal Structure with Barstar Bound to Barnase.** The overall fold of uncomplexed barstar closely resembles that of barstar in the complex with barnase (1BRS, 1BGS). 1BRS (8) and 1BGS (9) are independently solved structures of the barnase–barstar complex from the laboratories of Fersht and Mauguen, respectively. Both of the structures are in the same space group with similar cell parameters, but 1BRS is at a higher resolution (2.0 Å) than 1BGS (2.6 Å). Both structures have three pairs of molecules in the asymmetric unit and have a barstar dimer in the crystal, though the dimer has not been explicitly mentioned previously (8, 9). The NMR structure of the uncomplexed protein (1BTA) also resembles our crystal structure, but with significant differences. A superimposition of the  $C_{\alpha}$  traces of the 1BRS crystal structure and the C82A crystal structure, (Figure 3A), as well as main chain rmsd plots (Figure 2B), shows that the differences in structures are restricted to residues 58–62, 6–9, and Gly 43. Thus, the C82A structure is very similar to that of barstar in complex with barnase, with the average rmsd of the main chain being 0.55 and 0.48 Å when compared to 1BGS and 1BRS, respectively. The average side-chain rmsd for a similar comparison yields values of 1.06 and 1.11 Å, respectively. The main-chain residues of barstar (27–44) that lie at the binding interface with barnase do not show significant deviations in comparison to the structure in the complex, except for Gly 43 (Figures 2B, 3A, and 4A). Taking into consideration the resolution of this structure as well as the quality of the electron density map, the side-chain residues in the binding site which show significant changes in C82A when compared to 1BRS are Y29, D35, and D39 (Figure 4A). These differences as well as the difference in the MC of Gly 43 were confirmed by SA omit maps. The hydrogen bonding groups in these residues (Y29OH, D35OD1, and G43O) are unsatisfied and may be interacting with water molecules not seen at this resolution. The hydrogen bonding groups in the binding site which are satisfied internally include W38NE1, T42OG1, D39OD2, and D39N. The residues listed above form important hydrogen bonds directly (Table 2 in ref 8) or indirectly via water (Table 4 in ref 8) with barnase. Thus, apart from minor differences, the conformations of residues involved in the binding of barstar to barnase are similar in

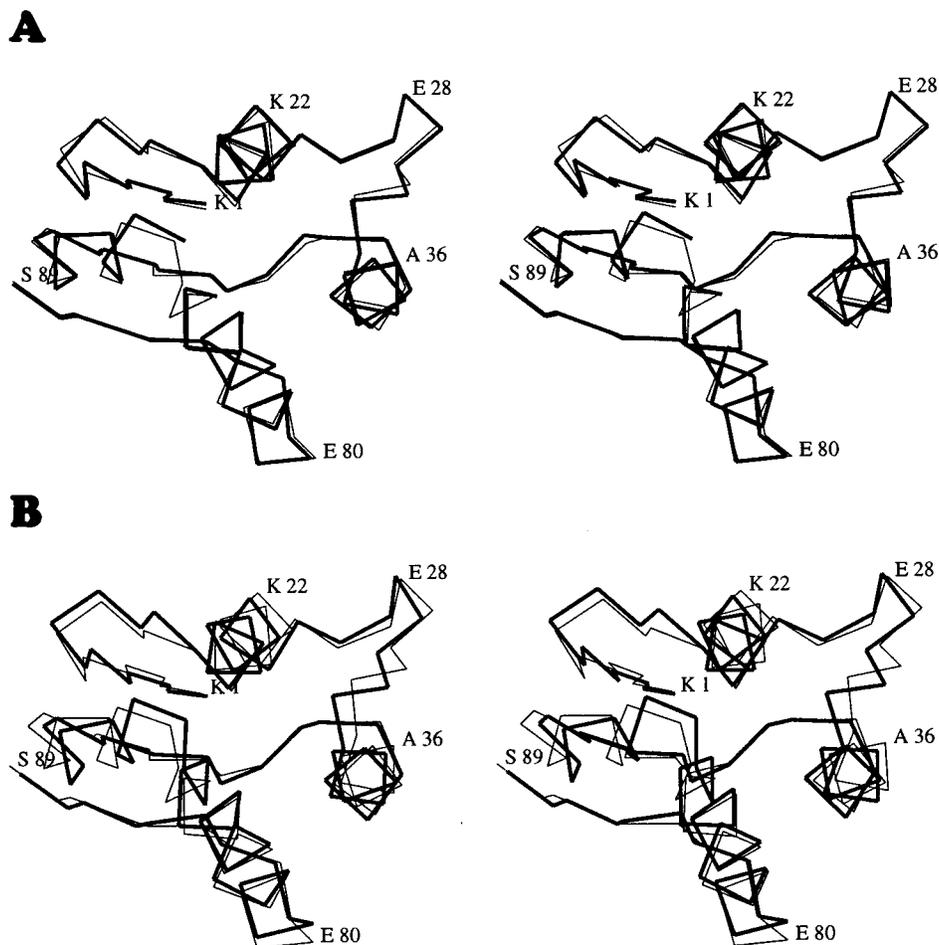


FIGURE 3: (A) C $\alpha$  trace of the crystal structure of complexed barstar (1BRS, E-chain) superimposed on the C82A structure. (B) C $\alpha$  trace of the barstar NMR structure (1BTA) superimposed on the C82A structure. In both cases, the C82A structure is represented by the thin line.

the free and bound states. To check for model bias in the structure, especially in the binding region, the refinement was repeated against the C82A reflection data, using the NMR structure (which shows 1 Å MC deviation from the final C82A structure; Figure 2C) as a starting model. The model was refined to a  $R$ -factor of 22% ( $R_{\text{free}} = 32\%$ ), and the refined structure was similar to the final C82A structure with a main-chain rmsd of 0.37 Å. In addition, to confirm the lack of change in the 27–44 region, the region was omitted and  $\sigma_A$ -weighed omit maps were calculated after SA torsion angle refinement. The region was rebuilt using FRODO (35), and the rebuilt structure was found to be similar to the final structure (rmsd of MC = 0.15 Å). We have examined the trends in  $B$ -factor variation along the polypeptide chain in the bound (1BGS) and free (C82A) structures. The 1BGS structure was chosen because it has a resolution (2.6 Å) similar to that of C82A (2.8 Å). In the 1BGS structure the 27–44 region is characterized by the lowest  $B$ -factors in the molecule. In our crystal structure this barnase-binding region has  $B$ -factors comparable to those of residues with weak electron density (residues 6–9). Figure 2D shows that the binding region exhibits the largest  $\Delta B$ -factor value. An analysis of  $B$ -factor trends is not straightforward because  $B$ -factors contain contributions from crystal mosaicity as well as static and dynamic disorder. We merely note that the  $B$ -factor trends in the free and bound structures are consistent with the expected decrease in flexibility of the 27–44 binding region upon complexation

with barnase. We compared the  $B$ -factor profiles in our structure with the mean rmsd's of the NMR structure set from the average NMR structure (22). The regions of large rmsd's for the NMR structures coincide with the three major peaks in the  $\Delta B$ -factor profile, except for the C-terminus of the binding helix, which does not show large rmsd's in the NMR structure. The largest rmsd's in the NMR structure set (22) are in the binding region. The above NMR data indicate flexibility in the binding region and are consistent with the large  $B$ -factors in our structure.

The NMR structure of free barstar had indicated that there were significant differences in the structures of the bound and free states (7, 8, 22). These differences were of three types: local changes in the conformations of residues in the binding region, an overall movement of the binding loop away from barnase, and global inward movements of the four helices in going from the bound to the free state. We do not observe these differences when comparing the crystal structure of free C82A to the barnase–barstar complex. At the resolution of the C82A structure, we cannot address the differences in the water structure changes between the free and the bound state. These changes are important considering the dominant role of charged residues and water in the stabilization of the interface. A complete analysis will have to wait for a high-resolution crystal structure of free barstar.

*Correlation of C82A Crystal Structure with Previous Binding and Unfolding Studies.* At 298 K barstar binds barnase with  $K_D$ ,  $\Delta H^\circ$ , and  $\Delta S$  values of  $10^{-14} \text{ M}^{-1}$ ,  $-45 \text{ kJ}$

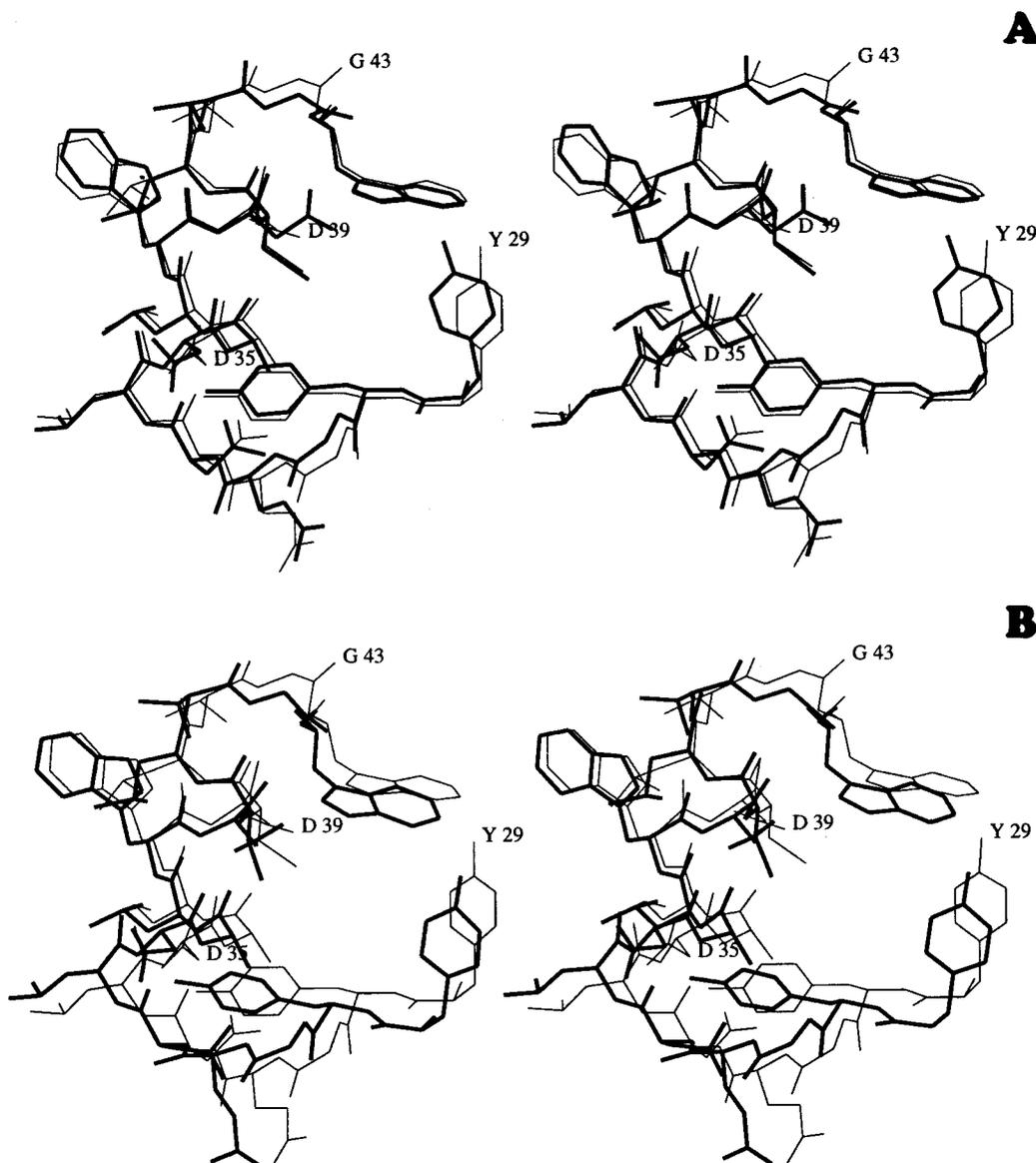


FIGURE 4: Binding regions of (A) 1BRS and (B) 1BTA superimposed on the C82A structure. The C82A structure is represented by the thin line.

$\text{mol}^{-1}$ , and  $51 \text{ J mol}^{-1} \text{ K}^{-1}$ , respectively. The positive entropic contribution has previously been suggested to be due to both an increase in flexibility of the molecule and a release of water molecules upon complex formation (12). Earlier measurements of hydrogen exchange rates by NMR (22, 42, 43) have suggested that free barstar is a flexible molecule. The C82A structure does not show any evidence for increased flexibility of barstar upon binding to barnase. In fact, the data are consistent with a decrease in flexibility of at least the binding residues 27–44. Indeed, ligand binding or complexation typically leads to a decrease rather than an increase in molecular flexibility. Hence, the positive entropy of binding is likely to be due solely to release of bound water molecules from both free barnase and free barstar upon complexation. The C82A structure shows that barstar appears to undergo very little conformational rearrangement upon binding. This is consistent with the very high association rate of  $6 \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$  (10). As noted previously (44), the observed  $\Delta C_p$  of binding is quite different from the  $\Delta C_p$  estimated from the amount of surface area buried upon binding ( $-384$  versus  $-85 \text{ cal mol}^{-1} \text{ K}^{-1}$ ).

This discrepancy may be due to a decrease in flexibility of barstar upon binding. A resultant decrease in vibrational contributions to  $C_p$  upon binding (45) may partly explain why the observed  $\Delta C_p$  of binding is considerably more negative than the  $\Delta C_p$  calculated from burial of surface area alone. Previous unfolding studies (46, 47) have found that the extrapolated values of  $\Delta H^\circ$  and  $\Delta S$  of unfolding at 383 K are lower than the corresponding values observed for most proteins at these convergence temperatures (48). These observations have suggested that barstar is more loosely packed than typical globular proteins. The crystal structure of free barstar does not show any evidence of such loose packing. Instead, the lower than average value of  $\Delta S$  is probably due to the high flexibility of the native state of free barstar.

*Comparison to the NMR Structure.* The NMR structure shows average rmsd's of  $1.1 \text{ \AA}$  for main-chain atoms and  $1.7 \text{ \AA}$  for side-chain atoms from the C82A free crystal structure (Figures 2C and 3B) and a similar deviation from the 1BRS-bound crystal structure. This deviation is in the range of deviations found between independently determined

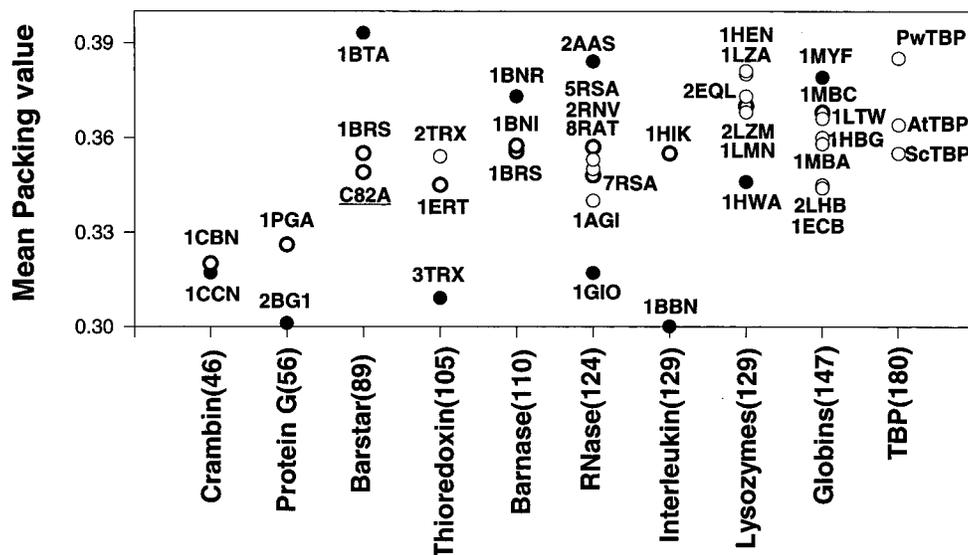


FIGURE 5: Mean normalized packing values of proteins for some representative crystal (empty circles) and NMR (filled circles) structures, calculated using the OS algorithm. Crystal structures which are the same protein as the NMR structure are represented by open circles with a thicker edge. The TBP packing values are from ref 38.

high-resolution NMR and X-ray structures (49, 50). The deviations are largest in the binding region and smallest in the  $\beta$ -sheet region as well as the helix-4 region. Since the C82A crystal structure is similar to the bound state of barstar, and since the differences between the bound state and the free NMR structure have already been described in detail (7, 8), we shall only briefly summarize the main differences between the NMR structure and the crystal bound (1BRS, 1BGS) and free (C82A) structures. The positions of the binding loop are significantly different in the crystal and NMR structures (Figure 4). In addition, the NMR structure has a smaller radius of gyration (11.39 Å) than 1BGS (11.79 Å), 1BRS (11.67 Å), and C82A (11.83 Å). The radius of gyration was calculated by considering only the C $\alpha$  atoms. The differences in the radius of gyration between C82A and the two complexed structures are minor and are due to the variation in the disordered helix-3 region. The NMR structure is thus a very compact structure (8, 9), and we have quantified and compared the packing in the NMR structure with that in our structure.

**Packing Analysis of NMR and Crystal Structures.** To quantify the increase in compactness of the NMR structure, we analyzed the atomic packing using the occluded surface (OS) algorithm (37). The overall packing in a protein structure is characterized by a parameter called the mean normalized protein packing value (38). For the NMR structure 1BTA, the value of this parameter was 0.393, which is higher than for all other protein crystal structures previously analyzed (38) as well as the structures analyzed in our study, suggesting that barstar is an exceptionally tightly packed protein. This high packing value was seen both for the averaged NMR structure and for the individual structures comprising the NMR structure set. In contrast, the C82A crystal structure has a packing value of 0.347. This indicates that the uncomplexed crystal structure is in fact marginally loosely packed compared to the complexed structures 1BRS and 1BGS (0.355).

To examine whether the apparently high packing of 1BTA is a general feature of all NMR structures, we calculated

the packing values for 70 proteins ranging in size from 29 to 180 residues from the PDB (41). We chose proteins for which both the crystal and NMR structures were available, either for the same structure or within the same protein family. Our analysis showed that all the protein structures solved by using crystallographic procedures had packing values lying within the narrow range of 0.34–0.37 (Figure 5), as pointed out by DeDecker et al. (38). The only exceptions were the well-packed hyperthermophilic TBP [packing value = 0.375 (38)] and a metallothionein (packing value = 0.25, 4MT2). The crystal structures of proteins with different sequences but belonging to the same family showed very similar packing values (e.g., myoglobin: average, 0.35; standard deviation, 0.008). In contrast, all of the structures determined by using NMR techniques showed a much larger scatter in packing value, ranging from 0.24 (1MHU) to 0.39–(1BTA). Except for one example (Crambin), no NMR structure showed a packing value similar to its crystal counterpart. Shown in Figure 5 are mean packing values of a few representative examples of the 70 proteins. Figure 5 clearly shows that the scatter in the mean packing values of the NMR structures is much larger than that in crystal structures. There are two possible explanations for the observed difference in packing values between NMR and crystal structures. It is possible that the packing value of a protein in solution is significantly different from that in the crystalline state. Alternatively, the packing value may not be a parameter that can be accurately determined by NMR data. To test the latter possibility, we examined the packing values of NMR structures of the same protein determined by different investigators. We also compared packing values of NMR structures of site-directed mutants of a protein with the wild-type packing value. In both situations, significant differences in packing values were observed. For example, in the case of human interleukin-4, the four NMR structures determined (1BBN, 1ITL, 1ITM, and 1CYK) show an average and a standard deviation of the mean packing values of 0.33 and 0.031, respectively. In contrast, the packing values for several different crystal structures of RNase A

determined in different solvents, pH's, and space groups show an average and a standard deviation of 0.353 and 0.005, respectively.

It has previously been stated (49) that the accuracy of good NMR structures for small proteins approaches that of high-resolution (2 Å) crystal structures. Our packing analysis shows that this is unlikely to be correct, since the variation in mean packing values appears to be much higher for NMR structures than for different crystal structures of the same protein. There is no generally accepted method for assessing the accuracy of an NMR structure. While a low rmsd of different members of an NMR structure set from the average structure is often taken to mean that the structure is well determined, this value is an estimate of the precision of the structure rather than its accuracy. A better estimate of accuracy would be the rmsd between NMR structures of the same protein determined independently by different investigators. Unlike the case of crystal structures, there are very few examples of such structure determinations using NMR. The barstar crystal structure has a mean packing value similar to that of mammalian ribonucleases, globins, and also barnase. The packing values show that the C82A structure is marginally loosely packed relative to the bound state (1BRS, 1BGS) and that free and bound barnase (1BNI, 1BRS) have similar packing values. Not only does the NMR structure of barstar (1BTA) appear to be a well-packed structure, it appears to be the best packed among the structures of the 70 proteins we have screened. This apparent tight packing is not consistent with the observation that barstar is a dynamically flexible and marginally stable protein with a stability of about 4.5 kcal/mol (18, 46). The tight packing in the free NMR structure may be an artifact of the NMR structure determination process, and the crystal structure of C82A is probably a better representation of the average structure adopted by uncomplexed barstar in solution.

## CONCLUSIONS

The free state of barstar is structurally very similar to the bound state, both in the overall conformation and at the binding site. The present structural data are useful in the interpretation of earlier thermodynamic studies of barstar unfolding and barnase-barstar complex formation. The packing analysis shows that protein packing values estimated by NMR are unreliable and that the accuracies of well-determined NMR structures are lower than those of high-resolution crystal structures. The present crystal structure is likely to be a better representation of the true structure of free barstar in solution than the NMR structure.

## ACKNOWLEDGMENT

The diffraction data was collected in the National Area Detector facility, IISc. We are grateful to the Supercomputer Education and Research Center, the Interactive Graphics Facility, and the Distributed Information Centre, IISc, for the use of their computational and computer graphics facilities. We thank Professor M. R. N. Murthy, B. Gopal, and R. Ravishankar for helpful discussions.

## REFERENCES

- Hartley, R. W. (1989) *Trends Biochem. Sci.* 14, 450–454.
- Hartley, R. W. (1997) in *Ribonucleases, Structure and Function* (D'Alessio, G., & Riordan, J. F., Eds.) pp 51–100, Academic Press, New York.
- Hartley, R. W. (1993) *Biochemistry* 32, 5978–5984.
- Blackburn, P., Wilson, G., and Moore, S. (1977) *J. Biol. Chem.* 252, 5904–5910.
- Kobe, B., and Deisenhofer, J. (1993) *Nature* 366, 751–756.
- Kobe, B., and Deisenhofer, J. (1995) *Nature* 374, 183–186.
- Lubienski, M. J., Bycroft, M., Freund, S. M. V., and Fersht, A. R. (1994) *Biochemistry* 33, 8866–8877.
- Buckle, A. M., Schreiber, G., and Fersht, A. R. (1994) *Biochemistry* 33, 8878–8889.
- Guillet, V., Laphorn, A., Hartley, R. W., and Mauguén, Y. (1993) *Structure* 1, 165–177.
- Schreiber, G., and Fersht, A. R. (1993) *Biochemistry* 32, 5145–5150.
- Jones, D. N., Bycroft, M., Lubienski, M. J., and Fersht, A. R. (1993) *FEBS Lett.* 331, 165–172.
- Martinez, J. C., Filimnov, V. V., Mateo, P. L., Schreiber, G., and Fersht, A. R. (1995) *Biochemistry* 34, 5224–5233.
- Agashe, V. R., Shastry, M. C. R., and Udgaonkar, J. B. (1995) *Nature* 377, 754–757.
- Nath, U., and Udgaonkar, J. B. (1997) *Biochemistry* 36, 8602–8610.
- Nath, U., Agashe, V. R., and Udgaonkar, J. B. (1996) *Nat. Struct. Biol.* 3, 920–923.
- Zaidi, F. N., Nath, U., and Udgaonkar, J. B. (1997) *Nat. Struct. Biol.* 12, 1016–1024.
- Nath, U., and Udgaonkar, J. B. (1995) *Biochemistry* 34, 1702–1713.
- Ramachandran, S., and Udgaonkar, J. B. (1996) *Biochemistry* 35, 8776–8785.
- Guillet, V., Laphorn, A., Fourniat, J., Benoit, J.-P., Hartley, R. W., and Mauguén, Y. (1993) *Proteins* 17, 325–328.
- Raghunathan, V., Khurana, S., Gupta, V., Khurana, R., Udgaonkar, J. B., and Salunke, D. M. (1994) *J. Mol. Biol.* 243, 533–536.
- Schoppe, A., Hinz, H.-J., Agashe, V. R., Ramachandran, S., and Udgaonkar, J. B. (1997) *Protein Sci.* 6, 2196–2202.
- Wong, K., Fersht, A. R., and Freund, S. M. V. (1997) *J. Mol. Biol.* 268, 494–511.
- McPherson, A. (1990) *Eur. J. Biochem.* 189, 1–23.
- Khurana, R., and Udgaonkar, J. B. (1994) *Biochemistry* 33, 106–115.
- Kabsch, W. (1993) *J. Appl. Crystallogr.* 26, 795–800.
- Navaza, J. (1994) *Acta Crystallogr.* A50, 157–163.
- Rice, L. M., and Brunger, A. T. (1994) *Proteins* 19, 277–290.
- Braig, K., Adams, P. D., and Brunger, A. T. (1995) *Nat. Struct. Biol.* 2, 1083–1094.
- Brunger, A. T. (1996) X-PLOR, version 3.851, Yale University, New Haven, CT.
- Kleywegt, G. J., and Jones, T. A. (1998) *Methods Enzymol.* (in press).
- Brunger, A. T. (1992) *Nature* 355, 472–474.
- Kleywegt, G. J., and Brunger, A. T. (1996) *Structure* 4, 897–904.
- Weis, W. I., Brunger, A. T., Skehel, J. J., and Wiley, D. C. (1990) *J. Mol. Biol.* 212, 737–761.
- Read, R. J. (1986) *Acta Crystallogr.* A42, 140–149.
- Jones, T. A. (1985) *Methods Enzymol.* 115, 157–171.
- Connolly, M. L. (1983) *Science* 221, 709–713.
- Pattabiraman, N., Ward, K. B., and Fleming, P. J. (1995) *J. Mol. Recognit.* 8, 334–344.
- DeDecker, B. S., O'Brien, R., Fleming, P., Geiger, J. H., Jackson, S. P., and Sigler, P. B. (1996) *J. Mol. Biol.* 264, 1072–1084.
- Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr.* D50, 760–763.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* 26, 283–291.
- Bernstein, F. C., Koetzle, K. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) *J. Mol. Biol.* 112, 535–542.
- Bhuyan, A. K., and Udgaonkar, J. U. (1998) *Proteins* 30, 295–308.

43. Lubienski, M. J., Bycroft, M., Jones, D. M., and Fersht, A. R. (1993) *FEBS Lett.* 332, 81–87.
44. Frisch, C., Schreiber, G., Johnson, C. M., and Fersht, A. R. (1997) *J. Mol. Biol.* 267, 696–706.
45. Sturtevant, J. M. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 2236–2240.
46. Wintrobe, P. L., Griko, Y. U., and Privalov, P. L. (1995) *Protein Sci.* 4, 1528–1534.
47. Agashe, V. R., and Udgaonkar, J. B. (1995) *Biochemistry* 34, 3286–3299.
48. Privalov, P. L. (1979) *Adv. Protein Chem.* 33, 167.
49. Wutrich, K. (1995) *Acta Crystallogr. D51*, 249–270.
50. Wagner, G., Hyberts, S. G., and Havel, F. H. (1992) *Annu. Rev. Biophys. Biomol. Struct.* 21, 167–198.

BI972857N