Structural and Energetic Responses to Cavity-Creating Mutations in Hydrophobic Cores: Observation of a Buried Water Molecule and the Hydrophilic Nature of Such Hydrophobic Cavities‡

Ashley M. Buckle, Patrick Cramer,§ and Alan R. Fersht*

Centre for Protein Engineering, Medical Research Council Centre, Hills Road, Cambridge CB2 2QH, U.K.

Received October 17, 1995; Revised Manuscript Received January 19, 1996⊗

ABSTRACT: We have solved the 2.0-Å resolution crystal structures of four cavity-creating Ile/Leu→Ala mutations in the hydrophobic core of barnase and compare and contrast the structural responses to mutation with those found for Leu→Ala mutations in T4 lysozyme. First, there are rearrangements of structure of barnase that cause the cavities to collapse partly, and there is an approximately linear relationship between the changes in stability and the volume of the cavity similar to that found for the mutants of T4 lysozyme. Second, although it is currently accepted that hydrophobic cavities formed on the mutation of large hydrophobic side chains to smaller ones are not occupied by water molecules, we found a buried water molecule in the crystal structure of the barnase mutant Ile76→Ala. A single hydrogen bond is formed between the water molecule and a polar atom, the carbonyl oxygen of Phe7, in the hydrophobic cavity that is formed on mutation. A survey of hydrophobic cavities produced by similar mutations in different proteins reveals that they all contain a proportion of polar atoms in their linings. The availability of such polar sites has implications for understanding folding pathways because a solvated core is presumed present in the transition state for folding and unfolding. Notably, the hydrogen bond between the cavity-water and the carbonyl group of Phe7 is also a marked early feature of very recent molecular dynamics simulations of barnase denaturation [Caflisch, A., & Karplus, M. (1995) *J. Mol. Biol. 252*, 672−708]. It is possible that cavities engineered into the hydrophobic cores of other proteins may contain water molecules, even though they cannot be detected by crystallographic analysis.

Globular proteins are characterized by the presence of well-packed hydrophobic cores, the formation of which is thought to be the major energetic driving force in protein folding (Kellis et al., 1988; Dill, 1990). The energetics of formation of the cores is amenable to study by making mutations in them, usually of large hydrophobic side chains to smaller ones, and analyzing the changes in energetics and structure (Kellis et al., 1988, 1989; Matsumura et al., 1988; Lim & Sauer, 1989; Shortle et al., 1990; Sandberg & Terwilliger, 1991a,b; Eriksson et al., 1992; Serrano et al., 1992; Buckle et al., 1993; Jackson et al., 1993a). The thermodynamic and structural consequences of mutations depend, however, on the exact location within the hydrophobic core and vary from virtually no structural movement, which is usually coupled to maximum destabilization, to almost complete collapse of surrounding atoms into the cavity along with some compensation of the loss of free energy of stabilization (Eriksson et al., 1992; Buckle et al., 1993; Jackson et al., 1993a). Since there is not a uniform response to mutation, it is necessary to study many mutations within a protein structure, as well as mutations in many different protein structures, in order to produce a database for benchmarking predictions. The linear relationship between the stability of the mutant and the size of the cavity resulting from cavity-creating mutations in T4 lysozyme was used to rationalize the variability in the structural response to mutation and offer an explanation as to why the strength of hydrophobic interactions as estimated from mutagenesis experiments is about two times larger than that estimated from model partition experiments (Eriksson et al., 1992).

One outstanding question is whether a water molecule could occupy a cavity created in the hydrophobic core on the mutation of a larger to a smaller residue. So far, no water molecules have been detected in any of the cavities created in mutants of lysozyme and barnase (Eriksson et al., 1992; Buckle et al., 1993). Further, the probability of a single water molecule being in a purely hydrophobic cavity is estimated to be approximately 1:20 000 (Wolfenden & Radzicka, 1994).

In this study, we describe the analysis of the crystal structures of four cavity-creating mutants of barnase, the small extracellular ribonuclease from *Bacillus amyloliquefaciens*, and compare the structural responses to mutation to those found for Leu→Ala mutants of T4 lysozyme (Eriksson et al., 1992). The barnase mutations are Leu14→Ala, Ile76→Ala, Ile88→Ala, and Ile96→Ala. The structure of barnase, indicating the location of the hydrophobic cores and mutated residues, is shown in Figure 1. The last three mutations are at the same positions as the barnase Ile→Val mutations previously characterized at the structural level (Buckle et al., 1993).

EXPERIMENTAL PROCEDURES

Mutagenesis, Protein Expression, and Purification. Mutagenesis of the barnase gene expression and purification of barnase mutants have been described (Serrano et al., 1990).

‡ Coordinates have been deposited in the Brookhaven Protein Data Bank (file names 1BRI, 1BRJ, 1BRK, and 1BRM).

§ Present address: European Molecular Biology Laboratory, Grenoble, c/o ILL BP 156, F-38042 Grenoble cedex, France.

⊗ Abstract published in *Advance ACS Abstracts,* March 1, 1996.

0006-2960/96/0435-4298$12.00/0

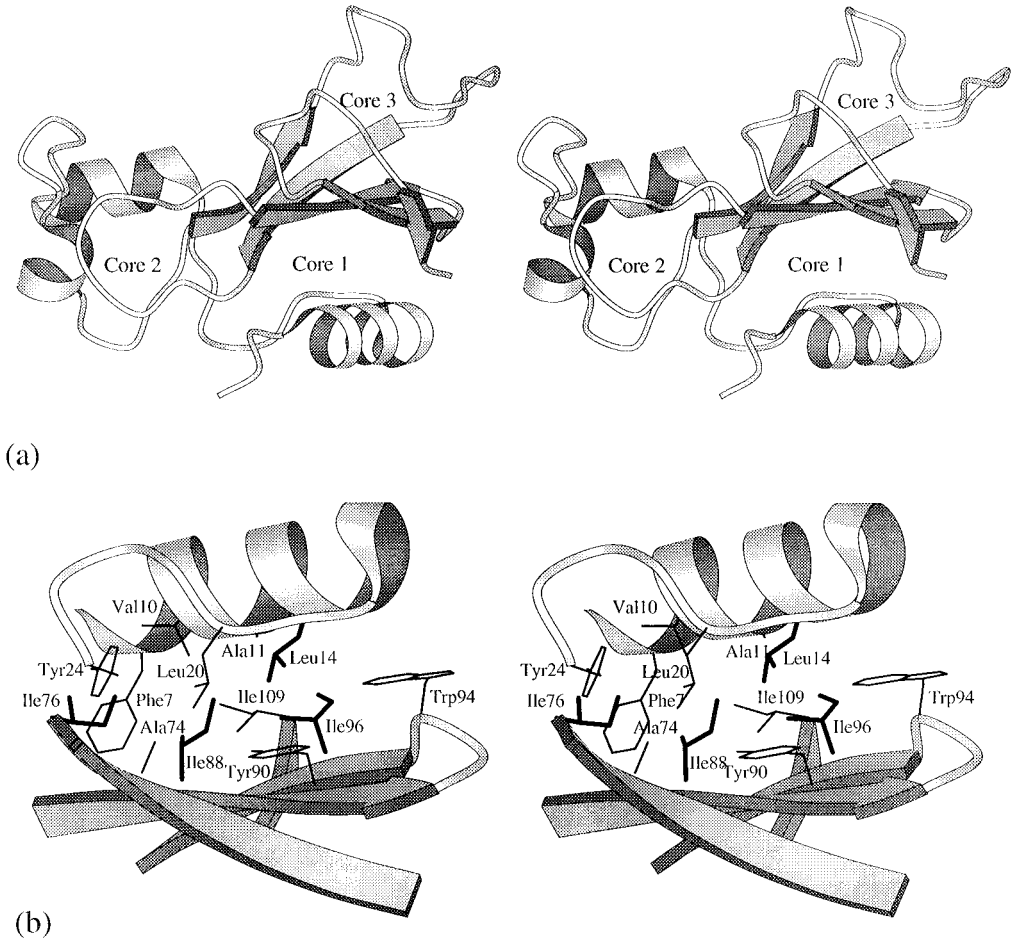

FIGURE 1: (a) Cartoon representation of barnase indicating secondary structure elements and the three hydrophobic cores. (b) View of $core_1$ showing the packing of side chains. Mutated residues are drawn in bold. The figure was produced with MolScript (Kraulis, 1991).

Structure Determination and Refinement. Crystallization, structure determination, and refinement of barnase mutants have been described in detail (Buckle et al., 1993). Data for the Ile88→Ala and the Ile96→Ala mutants were collected at 4 °C using a MAR research image plate and X-rays of wavelength 1.5418 Å generated by a rotating anode. Data for the Leu14→Ala and the Ile76→Ala mutants were collected at room temperature at the Synchrotron Radiation Source at Daresbury using a wavelength of 0.92 Å. Data collection and structure refinement statistics are summarized in Table 1.

Structural Analysis. All graphics work was carried out using the program O (Jones et al., 1991). Comparisons of barnase mutant and wild-type coordinates were done using chain C of the three chains in the unit cell and the wild-type structure solved at pH 7.5. Comparisons between wild-type and mutant structures were done after least-squares fitting of all Cα atoms. All least-squared fitting of coordinates was carried out with O and LSQKAB (CCP4, 1994). Atomic contact distances were calculated using O and CONTACT (CCP4, 1994). Cavity volumes were calculated with VOIDOO (Kleywegt & Jones, 1993), using the following parameters: initial grid spacing = 0.295 Å, VDW growth factor = 1.1, atom fattening factor = 1.1, and grid shrink factor = 0.9. Probe-occupied volumes were calculated. Cavity volumes were refined using successively finer grids until convergence was reached (convergence criteria = 0.1 $Å^3$ and 0.1%). The standard deviation calculated from this sample gives an estimate of precision. The accuracy of calculation of the volumes was estimated by repeating the calculations with nine randomly oriented copies of the molecule with respect to the grid. The cavities were produced in EZD format within VOIDOO (using a plot grid spacing of 0.15 Å) and converted with MAPPAGE (CCP4, 1994) for input into O, where they were inspected alongside space-filling models of the protein structures, to check whether the cavities generated by VOIDOO make physical sense. Different probe sizes were tried. Probe radii of 0.8 and 1.2 Å gave the most reliable results for barnase and lysozyme, respectively.

Coordinates of the T4 lysozyme structures were obtained from the Brookhaven database with the exception of Leu118→Ala and Leu121→Ala, which were made available by Prof. B. W. Matthews. Lysozyme mutant Leu46→Ala was not included in the analysis because Leu46 is in a different hydrophobic core from the other Leu→Ala mutants studied (Eriksson et al., 1992). Barnase mutant coordinates have been deposited in the Brookhaven database and have the following filenames: 1BRH (Leu14→Ala), 1BRI (Ile76→Ala), 1BRJ (Ile88→Ala), and 1BRK (Ile96→Ala).

RESULTS

There are no significant differences between the overall structures of barnase wild-type and mutant proteins. With the exception of the Ile76→Ala mutant, structural shifts at

Table 1: X-ray Data Collection and Refinement Statistics

mutant	Leu14→Ala	Ile76→Ala	Ile88→Ala	Ile96→Ala
		Data Collection		
no. of measured reflections	41880	35673	39885	32633
no. of unique reflections	20903	20394	19521	19313
unit cell dimensions[a] (Å)				
a, *b*	58.80	59.36	58.84	59.05
c	81.68	81.60	81.61	81.64
maximum resolution (Å)	2.0	1.9	2.0	2.0
data completeness (%)	97.5	80.5	91.6	89.8
R_{merge} (%)[b]	4.6	8.0	9.2	7.9
$\langle F/\sigma F \rangle$[c]	39.0	18.0	26.2	18.5
		Structure Refinement		
R_{cryst} (%), 7.0−2.0 Å, $F > 0$[d]	15.5	16.0	17.0	14.9
Δ_{bond} (Å)[e]	0.013	0.013	0.015	0.013
Δ_{angle} (deg)[e]	2.078	2.129	2.246	2.169
no. of atoms in refinement	2810	2801	2750	2875
no. of solvent molecules[f]	278	269	244	322
mean B_{mc} (Å^2)[g]	14.3	11.7	19.6	11.6
mean B_{sc} (Å^2)[g]	18.5	15.7	23.0	15.8

[a] All mutants are isomorphous with wild-type barnase (space group $P3_2$). Wild-type barnase (pH 7.5) unit cell dimensions are $a = b = 58.97$ Å and $c = 81.58$ Å. [b] R_{merge} gives the agreement between intensities of repeated measurements of the same reflections and can be defined as $\sum(I_{h,i} - \langle I_h \rangle)/\sum I_{h,i}$, where $I_{h,i}$ are individual values and $\langle I_h \rangle$ is the mean value of the intensity of reflection h. [c] Calculated by the CCP4 program TRUNCATE. [d] The crystallographic R factor, R_{cryst}, is defined as $\sum|F_o - F_c|)/\sum F_o$. [e] Average deviation of bond lengths and angles from ideal values. [f] Number of water molecules per asymmetric unit. [g] Average isotropic temperature (B) factor for main-chain (mc) and side-chain (sc) atoms (wild-type: $B_{mc} = 14.4$ Å^2; $B_{sc} = 16.8$ Å^2).

Table 2: Volumes of Internal Cavities, Percentage Cavity Collapse, and Atomic Shifts at the Site of Mutation

mutation	$\Delta\Delta G$ (kcal/mol)[a]	Eriksson V_{obs} (Å^3)[b]	V_{obs} (Å^3)[c]	SD range[d]	V_{mod} (Å^3)[e]	SD range	% collapse[f]	$\sum$(shifts) (Å)[g]
barnase								
Leu14→Ala	4.3		89.3 (1.5)	2.9−8.5	115.3 (1.3)	5.8−8.2	23	9.4
Ile76→Ala	1.9		31.1 (3.6)	1.8−11.0	63.0 (1.3)	3.0−3.8	51	22.1
Ile88→Ala	4.0		90.6 (0.9)	4.9−5.7	114.9 (0.2)	5.8−7.0	21	11.1
Ile96→Ala	3.2		83.0 (0.1)	3.9−4.7	103.0 (0.9)	5.2−8.0	19	23.6
lysozyme								
Leu99→Ala	5.0	122.5	126.3 (4.5)	2.2−8.2	109.9 (2.7)	1.6−8.4	−15	8.3
Leu118→Ala	3.5	66.5	86.8 (2.3)	1.7−7.5	83.1 (2.8)	1.6−12.2	0	12.6
Leu121→Ala	2.7	43.0	5.3 (9.7)	3.2−5.2	99.2 (3.2)	7.4−12.5	95	30.0
Leu133→Ala	3.6	77.8	80.7 (1.6)	10.9−14.5	99.9 (2.4)	10.9−15.5	19	10.5

[a] Free energies of unfolding for barnase and T4 lysozyme were taken from Serrano et al. (1992) and Eriksson et al. (1992), respectively. [b] Increase in cavity volume taken calculated by Eriksson et al. (1992). [c] The volume of the cavity in the mutant structure. This is the average of ten calculations using randomly rotated copies of the molecule, with respect to the grid. The standard deviation calculated from this sample, given in parentheses, gives an estimate of precision. Since there are two cavities in this hydrophobic core of wild-type T4 lysozyme, the volumes of these cavities (data not shown) are subtracted from the mutant cavity volumes to give the V_{mod} and V_{obs} values quoted in the table (which are equivalent to the "increase in cavity volume" used by Eriksson et al. (1992)). [d] Range of standard deviations for cavity volume calculations using successively finer grids. This gives an estimate of the accuracy of the cavity calculation. [e] The volume of the cavity remaining after deletion of the appropriate atoms from the coordinate file of the wild-type structure. [f] $[(V_{mod} - V_{obs})/V_{mod}] \times 100\%$. A value of 0% represents no movement of atoms into the created cavity whereas 100% indicates complete cavity closure. In the barnase mutant Leu14→Ala and the T4 lysozyme mutant Leu118→Ala, a tunnel connects the mutated side chain to the protein exterior, resulting in the probe escaping the cavity. This was avoided by placing a C atom in the center of the tunnel and 3.6 Å from the mutated side chain. [g] Sum of shifts of atoms within a cutoff distance of 6.0 Å from atoms deleted by the mutation. The lysozyme mutant Leu46→Ala was not included in the analysis because it is located in a separate hydrophobic core from the other four mutations.

the site of mutation are small (Table 2). There are no significant changes of temperature factors of atoms surrounding the mutation. In all cases, the volume of the engineered cavity is decreased by movements of surrounding atoms. Structural shifts are shown graphically in Figure 2 and are discussed separately below.

Mutants Leu14→Ala, Ile88→Ala, and Ile96→Ala. Mutations at residues 14 and 88 cause the smallest structural shifts, which are in the direction of the center of the created cavity (Table 2 and Figure 2a,b). The volumes of the engineered cavities decrease by 23% and 21%, respectively (Table 2). Mutation at position 96 causes larger structural shifts (Table 2 and Figure 2c). In particular, the side chain of Ile88 rotates by −22° (χ_1) and −41° (χ_2). Although some residues move toward the created cavity, thereby decreasing its volume, some actually move away from it (Arg110, Trp94, Ala11, and Gln15). Consequently, the cavity collapses by only 19% (Table 2). No solvent molecules were detected in any of the cavities.

Mutant Ile76→Ala. The structural response caused by this mutation is larger and more complicated than that in the other mutants (Figure 2d). Truncation of the side chain of Ile76 results in a small cavity. The cavity resulting from deletion of $C^{\gamma 1}$ and C^{δ} atoms is partly filled by the side-chain rotation of Ile88. The cavity created by this rotation has a volume of 31 Å^3 (Table 2) and contains a water molecule. This buried water molecule was clearly detected in both $2F_o - F_c$ and $F_o - F_c$ electron density maps (Figure 3). The spherical $F_o - F_c$ electron density peak surrounding the water molecule has a height of 4.7σ, where σ is the standard

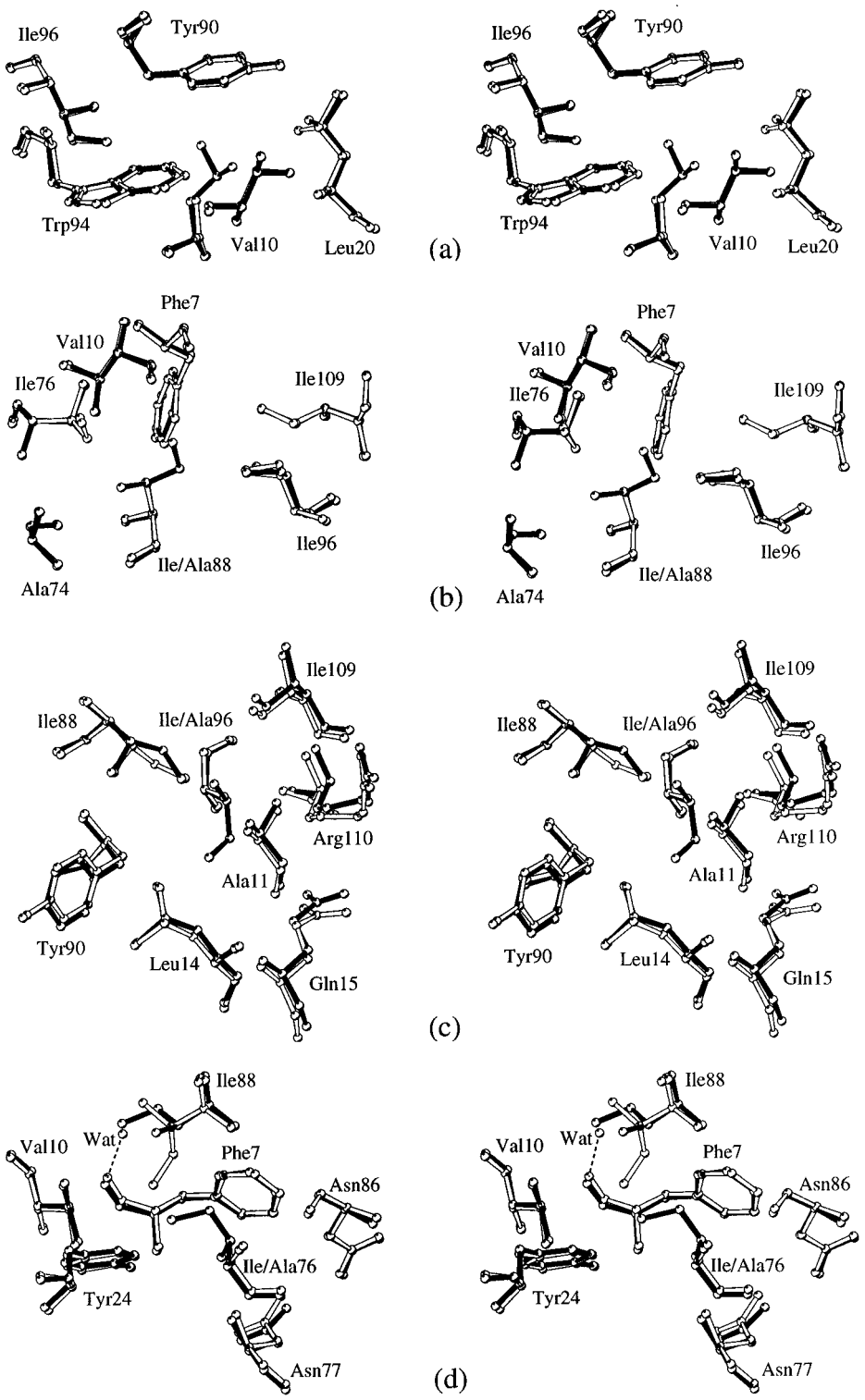

FIGURE 2: Stereo pictures showing the superposition of mutant structures (open bonds) on the structure of wild-type barnase (solid bonds). Residues within a sphere of radius 4.5 Å from the mutated residue are shown. (a) Leu14→Ala, (b) Ile88→Ala, (c) Ile96→Ala, and (d) Ile76→Ala. The hydrogen bond between the cavity-water molecule and protein backbone is drawn as a broken line. The figure was produced with MolScript (Kraulis, 1991).

deviation in the electron density map. Water molecules are usually modeled into $F_o - F_c$ electron density peaks of height greater than 3σ if "chemically sensible" water−protein interactions are made (weaker electron density is difficult to interpret). The asymmetric unit of the crystal consists of three protein molecules (A, B, and C). The electron density for the internal water molecule present in chain C is strong and well-defined, whereas density at the equivalent positions in chains A and B is weaker. The water molecule was modeled into the cavity in chain C only and has a *B*-factor of 36.5 Å^2 after full refinement (compared with an average of 27.0 Å^2 for all solvent molecules included).

The water molecule makes only one hydrogen bond with the surrounding protein (the backbone carbonyl oxygen of Phe7). With the exception of one NH group, the remaining 10 contacts within 4.0 Å are all nonpolar (Table 3). The protein−water hydrogen bond distance of

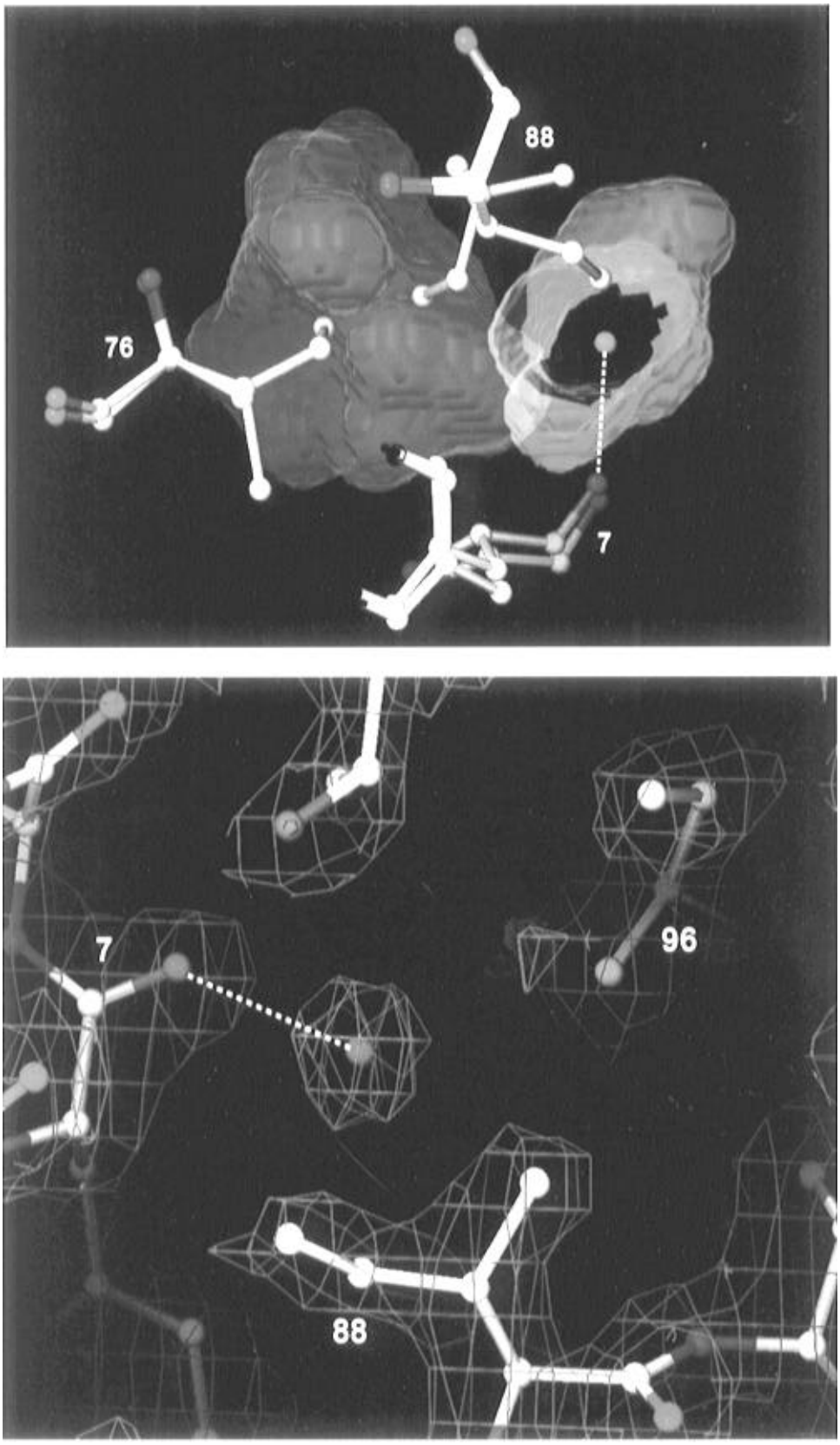

FIGURE 3: (a, top) The structural response to mutation in the barnase mutant Ile76→Ala (chain C) and the water molecule occupying the resulting cavity [wild-type bonds, white; Ile76→Ala bonds, yellow; "model" cavity (i.e., assuming no structural response to mutation), blue; observed cavity, red; water molecule, magenta]. (b, bottom) Ile76→Ala electron density in the region of the mutation ($2F_o - F_c$ density contoured at 1σ, red; $F_o - F_c$ density contoured at 3σ, green), calculated in the absence of the water molecule (chain C). The hydrogen bond between the cavity-water molecule and protein backbone is drawn as a broken line.

2.6 Å and bond angle of 135° fall within the ranges for hydrogen bond geometry of internal water molecules (Hubbard et al., 1994).

Cavities Engineered into Hydrophobic Cores of Proteins Contain Polar Groups in Their Lining. We have surveyed the cavities created by Ile/Leu→Ala mutations in barnase,

Table 3: Interactions between the Internal Water Molecule in the Barnase Mutant Ile76→Ala and Surrounding Protein Atoms

residue	atom	distance[a] (Å)	residue	atom	distance[a] (Å)
Phe7	C^{α}	3.8	Ile88	C^{β}	3.6
	C	3.6		$C^{\gamma 1}$	2.9
	O	2.6		$C^{\gamma 2}$	3.0
Ala11	C^{α}	3.6		C^{δ}	3.3
	C^{β}	3.4	Ile96	$C^{\gamma 2}$	3.6
	N	3.5	Ile109	C^{δ}	3.5

[a] Distance between the internal water molecule and surrounding protein atoms that are within 4.0 Å.

T4 lysozyme (Eriksson et al., 1992), staphylococcal nuclease (Shortle et al., 1988), and gene V protein (Sandberg & Terwilliger, 1991a,b). Although these mutations are in the so-called "hydrophobic core" of the structure, all of the resulting cavities are lined by a proportion of polar atoms (Table 4). We have included in our analysis a Cys→Ala mutant of human interleukin-1β because a positionally disordered water molecule has been found in its hydrophobic core (Ernst et al., 1995). This cavity is relatively large and completely nonpolar, in contrast to the other cavities analyzed.

Relationship of Structural Response to Stability: A Comparison with T4 Lysozyme. The structural response to mutation can be measured by the volumes of the resulting cavities and by the atomic shifts surrounding the mutation. In both cases, the barnase mutant structures are compared with Leu→Ala mutant structures of T4 lysozyme (Eriksson et al., 1992). The results of the cavity volume and atomic shifts calculations are summarized in Table 2. The magnitude of the atomic shifts around the site of mutation are similar for mutants of both proteins. The volumes of the cavities in barnase and lysozyme mutants cannot be directly compared because different probe sizes were used in each case. Also, the hydrophobic core of wild-type T4 lysozyme contains cavities, whereas hydrophobic $core_1$ of wild-type barnase is more tightly packed (no cavities were detected). There is an approximately linear relationship between cavity volume and stability for both lysozyme and barnase (Figure 4). That is, the most destabilized mutants have the largest cavities, whereas the least destabilized mutants have the smallest cavities. Lysozyme cavity volumes are very similar to those calculated by Eriksson et al. (1992), with the exception of the mutant Leu121→Ala, whose structure has been re-refined and shows larger atomic shifts at the site of mutation, including a significant side-chain rotation of a Phe residue (Prof. B. W. Matthews, personal communication).

There is an approximate linear relationship between stability and the sum of the shifts of atoms within a 6.0 Å radius of the deleted atoms, for mutants of both barnase and lysozyme (Figure 4). After trying several cutoff radii, we found that the best correlations resulted from radii of 6.0 Å for both proteins.

DISCUSSION

The Presence of a Water Molecule and the Hydrophilic Nature of Hydrophobic Cavities. The cavity created by the mutation Ile76→Ala in the hydrophobic core of barnase contains an ordered water molecule that makes one hydrogen bond to the protein backbone. This is a surprising observation because, so far, no water molecules have been detected in any of the cavities created in the hydrophobic cores of mutant proteins for which structures exist (Eriksson et al., 1992; Buckle et al., 1993; Jackson et al., 1993a). Moreover, the probability of a single water molecule being in a purely hydrophobic cavity is estimated to be approximately 1:20 000 (Wolfenden & Radzicka, 1994). However, cavities engineered into the hydrophobic cores of proteins barnase, T4 lysozyme, staphylococcal nuclease, and gene V are lined by a proportion of polar atoms (Table 4). Although no water molecules were detected in the cavities of those mutants whose X-ray structures have been determined (T4 lysozyme and barnase), it is clear from visual inspection of these cavities that a water molecule could find at least one hydrogen bond partner in the majority of cases. We suspect, therefore, that water molecules may also be present in the larger engineered cavities but are too mobile to be detected by crystallographic methods. A positionally disordered water molecule was detected by NMR methods in a cavity within the hydrophobic core of human interleukin-1β (Ernst et al., 1995). This cavity, which is lined purely by nonpolar atoms (Table 4), was found to be empty in three independently solved crystal structures of human interleukin-1β (Finzel, 1989; Preistle et al., 1989; Veerapandian, 1992). It is possible that we can detect the internal water molecule of the barnase mutant Ile76→Ala only because its cavity is the smallest of all cavities in barnase mutants and so the water molecule is more localized and because it hydrogen bonds with a surrounding polar group. Statistically, most buried water molecules make three hydrogen bonds with protein atoms, and the likelihood of a cavity being occupied by a water molecule increases with the number of available hydrogen bond partners (Hubbard et al., 1994; Williams et al., 1994). The electron density that suggests the presence of a water molecule in the cavity in chain C is much weaker in chains A and B. However, the environments of the three chains differ in the crystal, chain C being the most ordered of the three in the asymmetric unit.

The possibility of some degree of solvation of an engineered hydrophobic cavity has important implications for rationalizing the free energies of unfolding of such mutants. The Ile76→Ala mutant is the least destabilized of equivalent mutants of barnase. All of these mutations remove nonpolar interactions and create a cavity in the same hydrophobic core. Since a structural response to mutation always minimizes the destabilization, the presence of the water molecule in the Ile76→Ala mutant must have a stabilizing effect. The water−protein hydrogen bond and the van der Waals interactions (Table 1) must compensate for the unfavorable entropy cost of removing the water molecule from bulk solvent. Weak electrostatic interactions between the cavity water and surrounding nonpolar protein atoms should not be ruled out in the light of a recent analysis of the hydrogen bonding between water and small biological molecules, which finds a significant number of cases in which the only interaction present is of the type C−H···O (Steiner, 1995). Interestingly, a hypothetical cavity modeled by truncating the side chain Ile76 to Ala in the wild-type structure is more polar than the observed cavity which is occupied by a water molecule (50% of lining atoms in the modeled cavity are polar, compared with 22% in the observed cavity; Figure 3a).

The Role of Water in Protein Folding. Our observations have implications for protein folding pathways. It has been found that the rate-determining step in the unfolding of

Table 4: Survey of Hydrophobic Cavities in Protein Cores Arising from Ile/Leu→Ala Mutations[a]

protein	mutation	PDB code	cavity volume (Å^3)	total no. of lining atoms	no. of polar atoms in lining	% polar atoms in lining
barnase[b]	Ile76→Ala	1BRI	31 (63)[c]	32 (28)	7 (14)	22 (50)
	Ile88→Ala	1BRJ	91	52	13	25
	Ile96→Ala	1BRK	83	47	12	26
	Leu14→Ala	1BRH	89	52	11	19
T4 lysozyme[d]	Leu46→Ala	1L67	17	25	4	16
	Leu99→Ala	1L90	116	71	10	14
	Leu118→Ala		97	48	9	19
	Leu121→Ala		45	29	3	10
	Leu133→Ala	1L69	95	55	10	18
staph nuclease[e]	Ile72→Ala	1STN	16	19	3	16
	Leu103→Ala	1STN	28	30	6	20
	Leu125→Ala	1STN	59	43	12	28
gene V protein[e]	Ile47→Ala	1BGH	43	36	9	25
interleukin-1β	Cys71→Ser	31BI	41	31	0	0

[a] All calculations of probe-occupied volumes and cavity lining atoms were carried out with VOIDOO (Kleywegt & Jones, 1993). A probe radius of 1.4 Å was used, unless stated otherwise. This size probe was chosen since it is similar in size to a water molecule. [b] Cavities in barnase are generally smaller than cavities in the other proteins. Thus, not all parts of the cavities could be detected with a probe size of 1.4 Å and a probe size of 0.8 Å had to be used. As a result, the cavity volumes cannot be compared with those of the other proteins. [c] The values given in parentheses refer to the modeled cavity (i.e. assuming no structural response to mutation). [d] The volumes of the preexisting cavities in the wild-type structure have not been subtracted; therefore, these volumes are slightly larger than those quoted in Table 2. Coordinates of the lysozyme mutant structures Leu118→Ala and Leu121→Ala were kindly provided by Prof. B. W. Matthews. [e] Since X-ray structures for the mutants of staphylococcal nuclease and the gene V protein have not been determined, cavities were modeled by truncating the mutated residues in the wild-type structure.

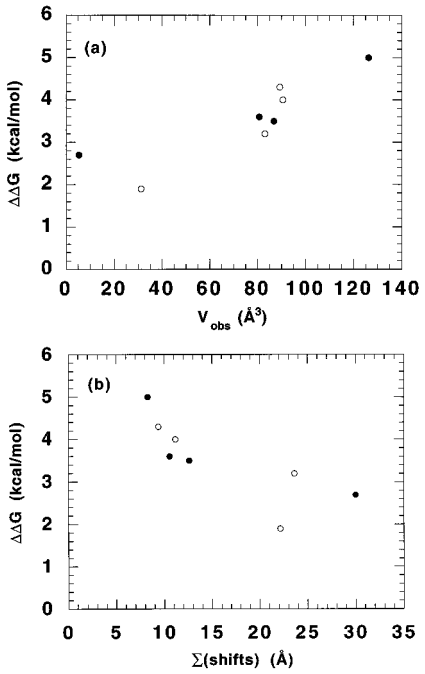

FIGURE 4: Change in the free energy of unfolding ($\Delta\Delta G$) plotted as a function of (a) observed cavity volumes and (b) total shifts of atoms surrounding the site of mutation for Ile/Leu→Ala mutants of barnase (open circles) and T4 lysozyme (solid circles). A cutoff distance of 6.0 Å from the deleted atoms was used in the atomic shifts calculation (see legend to Table 2).

barnase and chymotrypsin inhibitor 2 involves the penetration of water molecules into their cores (Fersht, 1993; Jackson et al., 1993b) and that the major transition state for folding involves the expulsion of solvent from the major hydrophobic core (Fersht et al., 1992). This is consistent with molecular dynamics simulations which suggest that the first step of unfolding involves the solvation of the major cores (Caflisch & Karplus, 1994, 1995; Li & Daggett, 1994). Recent molecular dynamics simulations of the acid and thermal denaturation of barnase show that, during the early stages of unfolding, the side chain of Phe7 is almost completely solvated and an invading water molecule hydrogen bonds with its carbonyl group, as in the barnase Ile76→Ala mutant structure (Caflisch & Karplus, 1995). Hydrogen-bonding partners are available to solvent molecules even in an environment of hydrophobic side chains. Further, the presence of a water molecule close to such side chains is less unfavorable than previously believed since one hydrogen bond is enough to immobilize a water molecule in a small hydrophobic cavity.

Structural Changes on Mutation and Their Relation to Changes in Energy of Stabilization and to the Hydrophobic Effect. The overall structural response to mutation for Ile/Leu→Ala mutations in barnase parallels that of the equivalent Ile→Val mutations and is similar to Leu→Ala mutations of T4 lysozyme; there are small and variable rearrangements and various degrees of collapse of the cavities formed on mutation. The structural response can be measured by the volumes of the resulting cavities and by the atomic shifts at the site of mutation. The volume of a cavity is not a very suitable parameter for analyzing a structural response to mutation since the calculation is nontrivial and its accuracy and precision depend on the size of the cavity and the probe (the small probes needed to measure small cavities will find the small "packing defects" in the protein core, whereas larger probes used in larger cavities may not find all of the cavity space). Thus, we cannot directly compare the cavity volumes of barnase mutants with those of lysozyme mutants, or make conclusions of a quantitative nature. The structure of the barnase mutant Ile96→Ala shows that the residues surrounding a mutation can adjust their positions with little change in the cavity volume, suggesting that some free energy can be gained by small shifts of atoms that improve packing, and not necessarily by a collapse of atoms into the cavity.

The linear relationship between cavity volume and free energy of unfolding for mutants of T4 lysozyme has been used to estimate the strength of the hydrophobic effect in proteins (Eriksson et al., 1992). That is, a free energy of approximately 2.0 kcal/mol that was calculated from an extrapolation of data to zero cavity volume was considered

to represent the difference in energy required to desolvate leucine relative to alanine, since it is in good agreement with values derived from model partition experiments. Eriksson et al. (1992) suggest that the slope of the linear relationship represents the context-dependent nature of the substitution, in terms of the extent of collapse of surrounding atoms into the cavity created by mutation. Although the qualitative nature of this relationship is of general importance, it is questionable whether values derived from such plots are of sufficient accuracy to be of predictive use. Also, complete collapse of surrounding atoms into the cavity involves a free energy of reorganization, which presumably accounts for a significant fraction of the free energy derived from the extrapolation. There is a fundamental difference between the hydrophobic effect as measured from simple model partition experiments and those involving proteins (Fersht et al., 1993). The transfer of a hydrophobic solute from water to organic solvent can be divided into the following notional steps: (i) transfer of the hydrophobic solute from water to vacuum; (ii) collapse of the cavity in water; (iii) opening a cavity in the hydrophobic phase; and (iv) transfer of solute to the cavity. The transfer of a hydrophobic solute to a preexisting cavity in a protein does not have step iii in the former process and is thus more favorable. If the cavity is empty, then the transfer will be more favourable for transfer to the protein by the energy expended in opening a cavity. If the cavity is filled with water, then it will be slightly more favorable still. Any degree of collapse of surrounding atoms into the cavity will be accompanied by a favorable free energy change, the reorganization energy. A macroscopic surface tension model predicts that, all things being equal, the transfer to a water-filled cavity is exactly twice as energetically favorable as transfer to an organic solvent (Fersht et al., 1993). A second difference between the results from models and with proteins results from the high packing density of atoms in the interior of proteins (Harpaz et al., 1994). This is illustrated experimentally from a correlation between the observed change in energy on deletion of methylene groups from hydrophobic side chains in a core and the number of $-CH_2-$ and CH_3- groups in a sphere of radius 6 Å surrounding the methylene groups that are deleted (Serrano et al., 1992; Otzen et al., 1995). This suggests that, since the van der Waals interactions are additive, the higher the packing density, the stronger the van der Waals interactions.

There is an approximately linear relationship between the changes in the free energy of unfolding on mutation and the local atomic shifts for mutants of both lysozyme and barnase (Figure 4). This measure of structural response is better than the use of cavity volumes because the calculation is simpler. The scatter in the data is too great to allow a meaningful extrapolation to zero atomic shift, but in theory this would allow the estimation of the destabilization caused by an Ile/Leu→Ala mutation when there is no rearrangement of the protein (free energies of reorganization are eliminated).

In conclusion, we have shown that the structural responses to cavity-creating mutations in the hydrophobic cores of T4 lysozyme and barnase are similar in terms of the magnitude of the local atomic shifts and the relationships between cavity volume and atomic shifts to the free energy of destabilization. The presence of polar atoms in the lining of cavities engineered into the hydrophobic cores of proteins, coupled with the observation of a water molecule inside such a cavity raises the possibility that cavities engineered into the hydrophobic cores of other proteins may contain solvent molecules, even though they cannot be detected by crystallographic analysis.

ACKNOWLEDGMENT

We are grateful to Prof. Brian Matthews for sending us the coordinates of the T4 lysozyme mutants Leu118→Ala and Leu121→Ala.

REFERENCES

Buckle, A. M., Henrick, K., & Fersht, A. R. (1993) *J. Mol. Biol. 234*, 847−860.
Caflisch, A., & Karplus, M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 1746−1750.
Caflisch, A., & Karplus, M. (1995) *J. Mol. Biol. 252*, 672−708.
CCP4 (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D50*, 760−763.
Dill, A. K. (1990) *Biochemistry 29*, 7133−7155.
Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P., & Matthews, B. W. (1992) *Science 255*, 178−183.
Ernst, J. A., Clubb, R. T., Zhou, H.-X., Groneborn, A., & Clore, G. M. (1995) *Science 267*, 1813−1817.
Fersht, A. R. (1993) *FEBS Lett. 325*, 5−16.
Fersht, A. R., Matouschek, A., Sancho, J., Serrano, L., & Vuilleumier, S. (1992) *Faraday Discuss. 93*, 183−193.
Fersht, A. R., Jackson, S. E., & Serrano, L. (1993) *Philos. Trans. R. Soc. London A 345*, 141−151.
Finzel, B. C. (1989) *Science 209*, 779.
Harpaz, Y., Gerstein, M., & Chothia, C. (1994) *Structure 2*, 641−649.
Hubbard, S. J., Gross, K. H., & Argos, P. (1994) *Protein Eng. 7*, 613−626.
Jackson, S. E., Moracci, M., elMasry, N., Johnson, C. M., & Fersht, A. R. (1993a) *Biochemistry 32*, 11259−11269.
Jackson, E. E., elMasry, N., & Fersht, A. R. (1993b) *Biochemistry 32*, 11270−11278.
Jones, T. A., Zou, J.-Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
Kellis, J. T., Jr., Nyberg, K., Sali, D., & Fersht, A. R. (1988) *Nature 333*, 784−786.
Kellis, J. T., Jr., Nyberg, K., & Fersht, A. R. (1989) *Biochemistry 28*, 4914−4922.
Kleywegt, G. J., & Jones, T. A. (1993) *Acta Crystallogr. D50*, 178−185.
Kraulis, P. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Li, A., & Daggett, V. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 10430−10434.
Lim, W. A., & Sauer, R. T. (1989) *Nature 339*, 31−36.
Matsumura, M., Becktel, W. J., & Matthews, B. W. (1988) *Nature 334*, 406−410.
Otzen, D. E., Rheinnecker, M., & Fersht, A. R. (1995) *Biochemistry 34*, 13051−13058.
Preistle, J. P., Schar, H. P., & Grutter, M. G. (1989) *Proc. Natl. Acad. Sci. U.S.A 86*, 9667.
Sandberg, W. S., & Terwilliger, T. C. (1991a) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1706−1710.
Sandberg, W. S., & Terwilliger, T. C. (1991b) *Trends Biotechnol. 9*, 59−63.
Serrano, L., Horovitz, A., Avron, B., Bycroft, M., & Fersht, A. R. (1990) *Biochemistry 29*, 9343−9352.
Serrano, L., Kellis, J. T., Jr., Cann, P., Matouschek, A., & Fersht, A. R. (1992) *J. Mol. Biol. 224*, 783−804.
Shortle, D., Meeker, A. K., & Freire, E. (1988) *Biochemistry 27*, 4761−768.
Shortle, D., Stites, W. E., & Meeker, A. K. (1990) *Biochemistry 29*, 8033−8041.
Steiner, T. (1995) *Acta Crystallogr. D51*, 93−97.
Veerapandian, B. (1992) *Proteins 12*, 10.
Williams, M. A., Goodfellow, J. M., & Thornton, J. M. (1994) *Protein Sci. 3*, 1224−1235.
Wolfenden, R., & Radzicka, A. (1994) *Science 265*, 936−937.

BI9524676