Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 8237–8241, November 1989
Biochemistry

Hydrophobic packing in T4 lysozyme probed by cavity-filling mutants

(protein structure/bacteriophage T4/hydrophobicity/protein folding)

MIHAIL KARPUSAS, WALTER A. BAASE, MASAZUMI MATSUMURA*, AND BRIAN W. MATTHEWS

Institute of Molecular Biology and Department of Physics, University of Oregon, Eugene, OR 97403-1229

Contributed by Brian W. Matthews, July 17, 1989

ABSTRACT **To probe the nature of the hydrophobic cores of proteins and to test potential ways of increasing protein thermostability, an attempt was made to improve the packing within T4 bacteriophage lysozyme by engineered amino acid replacements. Two mutations, Leu-133 → Phe and Ala-129 → Val, which were designed to fill the largest cavities that exist in the folded structure of the native protein, were constructed. The mutant proteins have normal activities and their thermal stabilities are marginally lower than that of wild-type lysozyme. Crystal structure analysis of the mutant proteins shows that the introduced amino acids are accommodated with very little perturbation of the three-dimensional structure. Incorporation of the more bulky hydrophobic residues within the core of the protein is expected to provide an increase in hydrophobic stabilization, but this is seen to be offset by the introduction of strain. Inspection of the mutant structures shows that in each case the introduced amino acid side chain is forced to adopt a non-optimal dihedral angle χ_1. Strain is also observed in the form of bond angle distortion and in unfavorable van der Waals contacts. The results illustrate how the observed core structures of proteins represent a compromise between the hydrophobic effect, which will tend to maximize the core packing density, and the strain energy that would be incurred in eliminating all packing defects. The results also suggest that mutations designed to increase protein stability by filling existing cavities may be effective in some cases but are unlikely to provide a general method for increasing protein stability.**

It has been anticipated for some time that the folding of proteins is probably dominated by the hydrophobic effect, i.e., by the sequestering of nonpolar side chains away from solvent. Recent studies suggest that alterations in many solvent-exposed residues on the surfaces of proteins have little if any effect on folding and stability (1–4). These results lead one to consider protein folding from a somewhat different perspective. Is it possible that folding is determined not by the whole amino acid sequence but by a small subset of key residues? If so, which are the critical residues and how do they determine three-dimensional structure? In an attempt to address questions such as these, increasing attention, both theoretical (5) and experimental (6–8), is being directed to the hydrophobic cores of proteins. Such cores have densities comparable to crystals of simple organic molecules (9, 10) but the packing is not perfect and some cavities remain (11). What determines the size of these cavities?

In the present study we have used site-directed mutagenesis to probe the nature of the core of T4 bacteriophage lysozyme. Specifically, we have asked whether the hydrophobic packing within the molecule can be improved by appropriate amino acid replacements. The hydrophobic surface area that is buried within a folded protein contributes directly to its free energy of stabilization (12, 13). Therefore, amino acid substitutions that increase the bulk of buried hydrophobic residues without introducing concomitant strain into the folded structure might also provide a way to increase the stability of proteins.

DESIGN AND CONSTRUCTION OF AMINO ACID REPLACEMENTS

The general approach was to identify all cavities in T4 lysozyme and to seek amino acid replacements that would fill these cavities without at the same time introducing excessively short van der Waals contacts or requiring unfavorable side-chain conformations. The largest cavities within the refined structure of T4 lysozyme (14) were determined by the program of Connolly (15). In this procedure each atom is given an appropriate van der Waals radius (16) and a probe sphere of defined radius is rolled over this van der Waals surface (17). The volume of a cavity is defined as the volume enclosed by the surface traced out by the center of the probe sphere. Note that a "cavity" is a void space within the protein. In contrast, an internal site occupied by a solvent atom is not considered a cavity.

Hydrophobic residues comprising the walls of each cavity were examined (18) to determine whether the size of the cavity could be reduced by an appropriate amino acid substitution. The following criteria were adopted. (*i*) The substitution should replace a smaller hydrophobic side chain with a larger one. (*ii*) The introduced side chain should make no contact of less than 3.2 Å with the surrounding atoms, which were assumed to be fixed. (*iii*) The dihedral angles of the introduced side chain should be close to the values observed in well-determined protein structures (5, 19). (*iv*) The residue being substituted should have a solvent accessibility (17) of zero. The last criterion was included because we were interested in probing hydrophobic packing within the core of the protein and wanted to ensure that the residue being replaced was completely buried.

Two amino acid replacements suggested by the above screening procedure were Leu-133 → Phe and Ala-129 → Val. These replacements were expected to occupy the largest and second-largest cavities in the protein (Fig. 1), which have volumes of 39.0 $Å^3$ and 23.4 $Å^3$ calculated with a probe sphere of radius 1.2 Å.

Mutant proteins Leu-133 → Phe (L133F) and Ala-129 → Val (A129V) were constructed (6, 20), purified (21, 22), and crystallized (14). X-ray diffraction data were collected to 1.9-Å resolution (23) and the structures refined (24) with good stereochemistry to crystallographic residuals of ≈15.9% with

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*Present address: Department of Immunology, Research Institute of Scripps Clinic, 10666 North Torrey Pines Road, La Jolla, CA 92037.

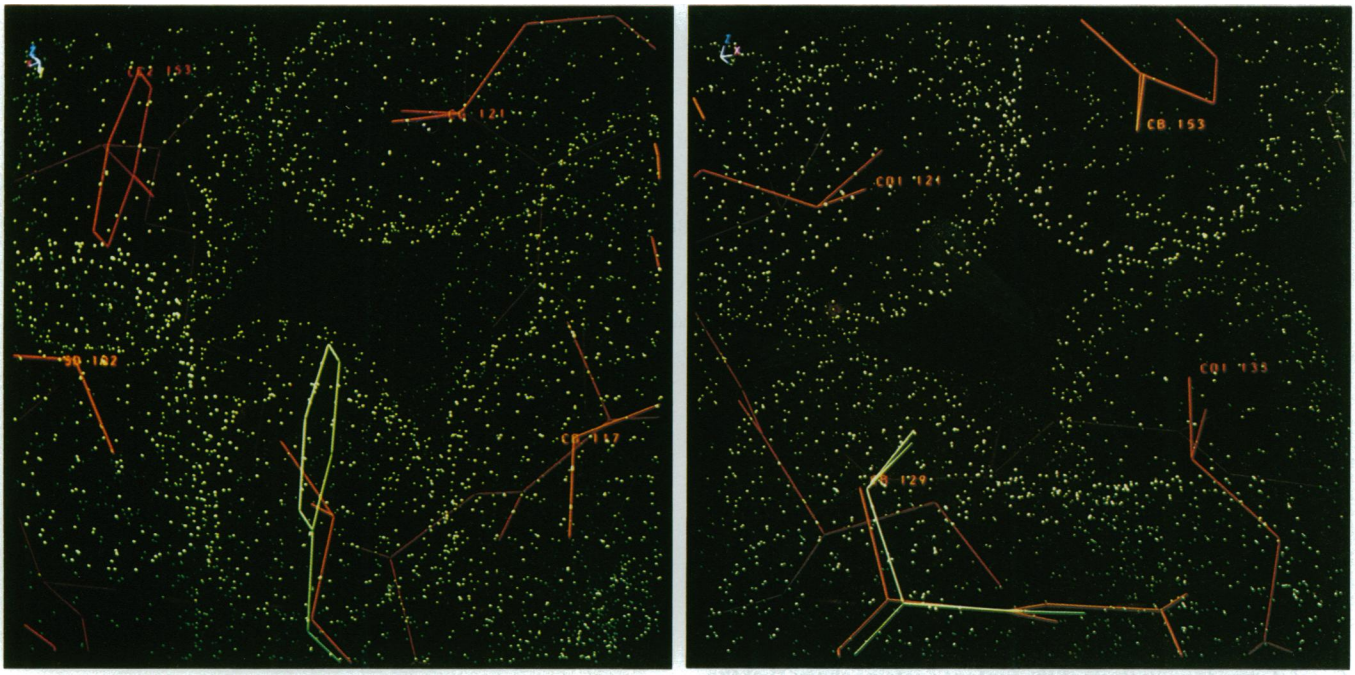

FIG. 1. (*Left*) Cavity within T4 lysozyme occupied by the Leu-133 → Phe replacement. The skeleton structure of wild-type lysozyme is shown in orange and the corresponding van der Waals surface is dotted. The amino acid designed to occupy the cavity (Phe-133) is shown in yellow (with coordinates taken from the mutant structure). (*Right*) Cavity occupied by the Ala-129 → Val replacement. Conventions are as in *Left*.

≈12,800 reflections. Coordinates will be deposited in the Brookhaven Data Bank.

RESULTS

The two mutant lysozymes have activities (25) that are indistinguishable from wild type. In the pH range 2–3, the reversible melting temperature (26) of L133F is about 1°C lower, and the second mutant, A129V, unfolds at a temperature about 2–3°C lower than wild type (Table 1).

The replacement of the leucine with the larger phenylalanine side chain is accomplished with very little change in the structure of the protein (Fig. 2). The largest shift is an adjustment in the side chain of Leu-121 (maximum atom shift of 0.68 Å). The difference in electron density between the second mutant, A129V, and wild-type lysozyme (Fig. 3*a*) is dominated by the density corresponding to the introduced valyl side chain but also includes positive and negative features indicating associated structural adjustments. The largest of these are translations of the δ-carbons of Leu-121 by 0.6 Å. Also residues 129–134, within the α-helix that includes Val-129, undergo a bodily translation of about 0.15 Å (Fig. 3*b*).

DISCUSSION

The results clearly show that there are cavities within the structure of T4 phage lysozyme that permit internal hydrophobic residues in the wild-type protein to be replaced with more-bulky counterparts. Furthermore, these substitutions can be accomplished with relatively minor adjustments in the structure of the protein. The most dramatic example is provided by the replacement Leu-133 → Phe, in which case the only significant change is an adjustment in the side chain of Leu-121 (Fig. 2*b*). Although it is commonly stated that the interiors of proteins are well-packed, the present results show that the packing density can be increased.

In both cases the mutations were designed to replace smaller nonpolar residues with larger ones. The anticipated additional hydrophobic stabilization can be estimated from the increase in solvent-accessible surface area. Each square angstrom of hydrophobic surface that is buried within the folded protein is expected to contribute ≈24 cal/mol toward the free energy of folding (12, 13). Although phenylalanine has three more carbon atoms than leucine, its side-chain solvent-accessible surface area is only 20–30 Å^2 greater (depending on the assumed side-chain conformation). Thus

Table 1. Thermodynamic parameters of mutant lysozymes

Protein	pH	t_m, °C	Δt_m, °C	ΔH, kcal/mol	ΔS, cal/(deg·mol)	$\Delta\Delta G$, kcal/mol
Wild type	1.99	41.4 ± 0.3	—	87 ± 10	278 ± 20	—
	2.86	52.1 ± 0.4	—			
L133F	1.99	40.7 ± 0.5	−0.7	83 ± 10	264 ± 20	−0.2
	2.85	51.3 ± 0.6	−0.8			−0.3
A129V	1.99	38.5 ± 0.5	−2.9	78 ± 10	252 ± 20	−0.8
	2.8	50.2 ± 0.6	−1.9			−0.7

The melting temperature, t_m, is the temperature of the midpoint of the reversible thermal denaturation transition in 0.2 M KCl (26). Δt_m is the difference in melting temperature between the mutant and wild-type lysozyme. The enthalpy of unfolding, ΔH, and the entropy, ΔS, were derived from van't Hoff analyses of the thermal denaturations and were measured at the melting temperatures of the proteins. $\Delta\Delta G$ is the difference between the free energy of stabilization of the mutant lysozymes and wild type (a negative value indicates that the mutant is less stable than wild type). Standard deviations are estimates based on the reproducibility of the observed data for these and other mutant lysozymes.

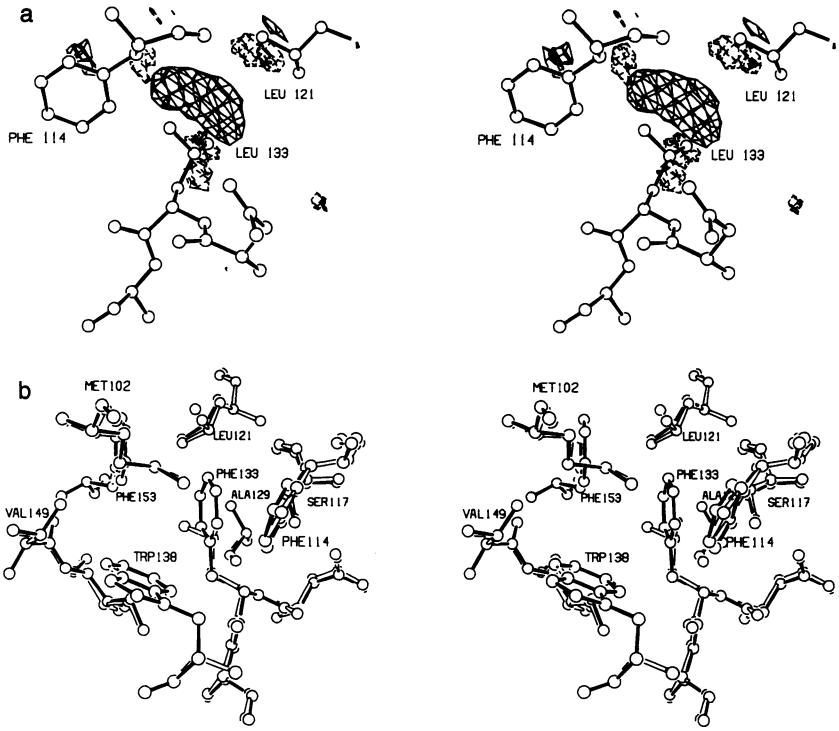

FIG. 2. (*a*) Map showing the difference in electron density between mutant L133F and wild-type lysozyme, superimposed on the wild-type structure. Difference amplitudes are ($F_{mut} - F_{WT}$), where F_{mut} and F_{WT} are the observed amplitudes for the respective crystals. Phases were calculated from a model of wild-type lysozyme similar to that described in ref. 14 but subject to additional refinement based on recollected x-ray data (unpublished results of K. Wilson, R. Faber, and B.W.M.). Contours were drawn at $\pm 3\sigma$, where σ is the root-mean-square (rms) density throughout the unit cell. Positive contours are drawn solid, negative contours broken. (*b*) Superposition of the refined structure of the mutant lysozyme L133F (solid bonds) on that of wild type (open bonds).

the hydrophobic stabilization expected for the Leu-133 → Phe replacement is only about 0.5–0.7 kcal/mol. This should be reflected in a slight increase (≈2°C) in the melting temperature of the mutant lysozyme. Why is this not observed?

Inspection of the wild-type and L133F crystal structures suggests that the additional hydrophobic stabilization is offset by strain introduced in the mutant protein. Table 2 lists the closest approaches and the estimated (27) van der Waals

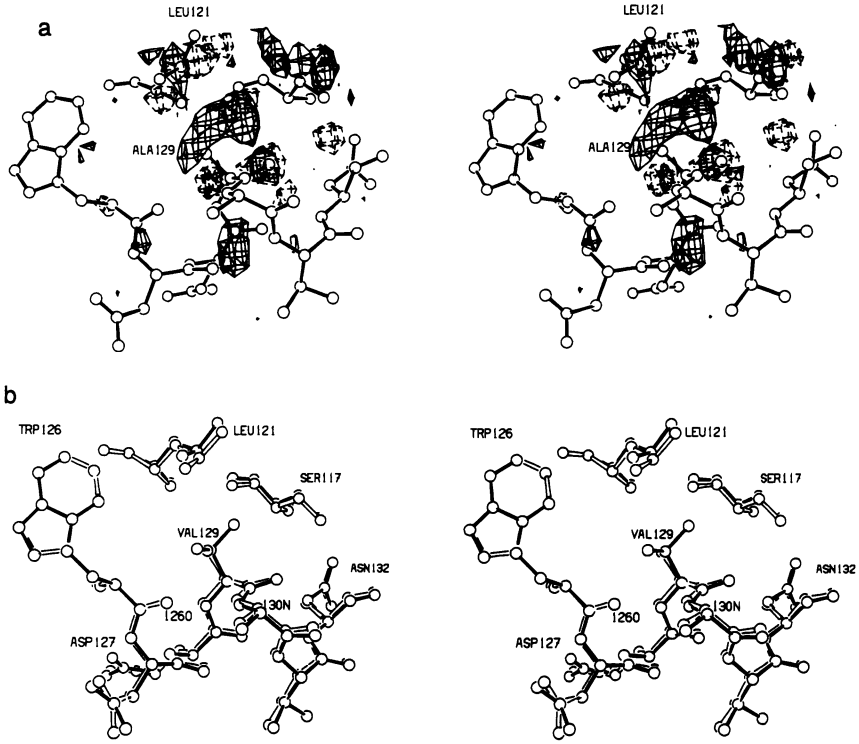

FIG. 3. (*a*) Difference electron density for mutant lysozyme A129V relative to wild type. All conventions and procedures are as for Fig. 2*a*. (*b*) Superposition of the structure of mutant lysozyme A129V (solid bonds) on that of wild type (open bonds).

Table 2. Contact distances and van der Waals interaction energies for substituted side chains

	Wild-type lysozyme			Mutant lysozyme		
Closest protein atom(s)	Side-chain atom	Distance to closest protein atom(s), Å	Interaction energy, kcal/mol	Side-chain atom	Distance to closest protein atom(s), Å	Interaction energy, kcal/mol
	Ala-129			*Val-129*		
O Arg-125	CB	3.50	−0.16	CB	3.55	−0.18
CD1 Leu-121		3.89	−0.18		3.89	−0.18
O Trp-126	—	—	—	CG1	2.93	1.31
N Ala-130					3.08	2.43
O Ser-117	—	—	—	CG2	3.62	−0.18
CG Leu-121					3.63	−0.09
	Leu-133			*Phe-133*		
O Ala-129	CB	3.45	−0.15	CB	3.54	−0.18
O Ala-129	CG	3.33	−0.07	CG	3.60	0.26
OG Ser-117		3.73	−0.19		3.84	0.25
O Ala-129	CD1	3.66	−0.19	CD1	3.07	0.68
OG Ser-117	CD2	3.58	−0.18	CD2	3.68	−0.06
CD1 Phe-153	—	—	—	CE1	3.66	0.22
CB Ser-117	—	—	—	CE2	3.65	−0.05
CB Ser-117	—	—	—	CZ	3.65	−0.05
CD2 Leu-121					3.85	−0.13

The table gives the closest distances and corresponding estimated van der Waals interaction energies (27) between atoms within the side chains of residues 129 and 133 and the surrounding atoms in the crystal structures of wild-type and the respective mutant lysozymes. Note that the table includes only the interactions with the atoms closest to the substituted side chains, not all van der Waals interactions.

interaction energies between the side-chain atoms of residue 133 and the surrounding protein atoms. Leu-133 in wild-type lysozyme has van der Waals interactions with its neighbors that are all favorable. In contrast, the side chain of Phe-133 in the mutant structure has several unfavorable close contacts, especially between atom CD1 and the carbonyl oxygen of Ala-129 (Table 2). These close contacts, in turn, require the side chain of Phe-133 to adopt an unfavorable rotational alignment about its C^{α}—C^{β} bond. In the mutant structure the χ_1 dihedral angle of Phe-133 is −87°. This is to be compared with $\chi_1 = -60°$, which is the low-energy trans conformation. Also, in protein structures in general, the average, and presumably the optimum, χ_1 for phenylalanyl (and leucyl) side chains in the g^+ conformation is close to −65° (5). In wild-type lysozyme, the value of χ_1 for Leu-133 is −80°. The estimated increase in strain energy corresponding to the change in χ_1 from −80° in wild-type lysozyme to −87° in L133F is ≈0.5 kcal/mol (27). It also appears that there is increased strain in the mutant structure in the form of deformation of one of the angles about the α-carbon of residue 133. The angle C—C^{α}—C^{β} of Leu-133 in wild-type lysozyme is 108.0°, and it decreases to 105.7° for Phe-133 in the mutant structure. These values are, respectively, 3.0° and 5.3° less than the low-energy value of 111.0° (24) and correspond to an estimated (27) increase in strain energy from 0.2 kcal/mol in wild type to 0.5 kcal/mol in the mutant structure. There also appears to be strain in the form of unfavorable van der Waals contacts (Table 2), although the overall magnitude of the van der Waals energy for a given residue is difficult to estimate reliably. The introduction of the additional atoms in the mutant side chain adds many favorable van der Waals interactions but introduces unfavorable interactions as well. These repulsive interactions, in particular, are very sensitive to possible errors in the atomic coordinates. One also has to take into account not only the van der Waals interactions within the folded mutant and wild-type proteins but also the van der Waals interactions of the respective unfolded side chains with solvent. What is clear, however, is that the sum of the different types of strain that are introduced into the mutant structure must exceed the hydrophobic and van der Waals stabilization, so that the L133F lysozyme is marginally less stable than wild-type lysozyme.

A similar situation occurs in the Ala-129 → Val variant. Valines within α-helices normally adopt the g^+ conformation with $\chi_1 \approx 180°$ (5, 19). In attempting to model the Val-129 structure, however, it was difficult to avoid close contacts between the valyl side chain and surrounding atoms, and the optimum value of χ_1 appeared to be 72°. This is exactly as observed in the crystal structure of the mutant protein (Fig. 3*b*). For an isolated valine within an extended polypeptide chain, a χ_1 value of 72° would correspond to a relatively low strain energy (≈0.2 kcal/mol). Val-129, however, is in an α-helix and the χ_1 value of 72° requires one of the γ-methyl groups to be wedged against the side of the helix. The CG1 atom makes unfavorable close contacts with both the carbonyl oxygen of Trp-126 and the nitrogen atom of Ala-130 (Table 2). Thus, although the replacement of a buried alanine with a buried valine results in an estimated increase of about 54 $Å^2$ in solvent-accessible surface area, which should stabilize the mutant structure by about 1.3 kcal/mol, this is offset by the apparent van der Waals strain energy (Table 2) associated with the unfavorable χ_1 value that Val-129 is forced to adopt. As with the Leu-133 → Phe substitution, the

Ala-129 → Val replacement also results in a protein that is marginally less stable than wild type.

For both replacements the largest observed structural changes in the protein (0.6–0.7 Å) occur within the side chain of Leu-121 (Figs. 2*b* and 3*b*). There is no evidence, however, for the introduction of strain in this residue. In particular, the side chain of Leu-121 maintains energetically favorable van der Waals interactions with the introduced amino acids (Table 2). This demonstration that a buried side chain can make facile adjustments in its position in response to a mutation will complicate algorithms designed to explore alternative packing arrangements within proteins (5). On the other hand, the main-chain atoms move very little and it is these essentially invariant atoms that are responsible for the energetically unfavorable contacts with the introduced side chains in the respective mutant structures (Table 2).

Taken together, the structural studies of the two mutant lysozymes described here provide new insights into the nature of the packing of protein interiors. Insofar as T4 lysozyme can be taken as a typical protein, it is seen that the packing within the cores of proteins is not designed to maximize the packing density alone. Rather, there remain cavities that could, in principle, be filled or partially filled by appropriate amino acid replacements. These amino acid replacements, however, are limited to the 20 naturally occurring amino acids. If only hydrophobic amino acids are allowed, as is normally observed for core replacements (8), the choice is even more restricted.

In this sense the hydrophobic cores of proteins are constructed from a small number of "quantized" building blocks whose volumes and shapes are determined by the identity of the amino acid. The shapes and placement of the building blocks can be varied by changes in the main-chain and side-chain dihedral angles, but these also have restricted values. The side-chain dihedral angles cluster within about ±10° (rms) of the low-energy minima (5, 19). An "extreme" rotational excursion of ±20° costs roughly 0.7 kcal/mol (27). Larger rotations toward more-eclipsed positions would introduce even more strain energy and so are avoided. Similarly, bond angles in well-refined protein structures cluster within about ±3° (rms) of their "ideal" values. An "extreme" bond angle deviation of 6° incurs a strain energy of about 0.7 kcal/mol (27).

On purely hydrophobic grounds the burial of an additional methyl or methylene group within the core of a protein contributes about 0.6 kcal/mol toward the stability of the protein (6, 9, 10, 28). It is striking that this favorable energy value (0.6 kcal/mol) is essentially equal to the unfavorable energy cost that residues are observed to pay for side-chain dihedral angle excursions, bond bending, and presumably other forms of strain as well. We suggest that this agreement is not a coincidence. It implies that the packing density of the protein has increased to the point that the further burial of methylene or other hydrophobic groups would require energy-equivalent perturbations of the protein stereochemistry. In other words the observed core structures of proteins reflect a compromise between the hydrophobic plus van der Waals effects, which will tend to maximize the core packing density, and the strain energy that would be incurred in eliminating all packing defects.

In terms of the use of genetic engineering to increase the stability of proteins, there are examples where the replacement of hydrophobic residues or the introduction of a nonpolar group within the core of a protein has resulted in increased thermostability (ref. 6; B. Malcolm, K. Wilson, B.W.M., J. Kirsch, and A. Wilson, unpublished results). The present results, however, suggest that mutations designed to fill existing cavities may be effective in some cases but are not likely to provide a general route to substantial improvement in protein stability.

We thank J. Wozniak, H. Nicholson, S. Dao-pin, L. H. Weaver, and D. L. Tronrud for help with mutagenesis, protein purification, data collection, and refinement procedures. We are also most grateful to Andrew Morton and Hale Nicholson for helpful comments on the manuscript. This work was supported in part by grants from the National Institutes of Health (GM21967 to B.W.M.; GM20195 to J. A. Schellman) and by a grant from the Lucille P. Markey Charitable Trust.

1. Hecht, M. H., Nelson, H. C. M. & Sauer, R. T. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 2676–2680.
2. Reidhaar-Olson, J. F. & Sauer, R. T. (1988) *Science* **241,** 53–57.
3. Alber, T., Dao-pin, S., Nye, J. A., Muchmore, D. C. & Matthews, B. W. (1987) *Biochemistry* **26,** 3754–3758.
4. Matthews, B. W. (1987) *Biochemistry* **26,** 6885–6888.
5. Ponder, J. W. & Richards, F. M. (1987) *J. Mol. Biol.* **193,** 775–791.
6. Matsumura, M., Becktel, W. J. & Matthews, B. W. (1988) *Nature (London)* **334,** 406–410.
7. Kellis, J. T., Jr., Nyberg, K., Sali, D. & Fersht, A. R. (1988) *Nature (London)* **333,** 784–786.
8. Lim, W. A. & Sauer, R. T. (1989) *Nature (London)* **339,** 31–36.
9. Richards, F. M. (1974) *J. Mol. Biol.* **82,** 1–14.
10. Chothia, C. (1975) *Nature (London)* **254,** 304–308.
11. Connolly, M. L. (1986) *Int. J. Peptide Protein Res.* **28,** 360–363.
12. Chothia, C. (1974) *Nature (London)* **248,** 338–339.
13. Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng.* **6,** 151–176.
14. Weaver, L. H. & Matthews, B. W. (1987) *J. Mol. Biol.* **193,** 189–199.
15. Connolly, M. L. (1983) *Science* **221,** 709–713.
16. McCammon, J. A., Wolynes, P. G. & Karplus, M. (1979) *Biochemistry* **18,** 927–942.
17. Lee, B. & Richards, F. M. (1971) *J. Mol. Biol.* **55,** 379–400.
18. Jones, T. A. (1982) in *Crystallographic Computing*, ed. Sayre, D. (Oxford Univ. Press, Oxford, U.K.), pp. 303–317.
19. McGregor, M. J., Islam, S. A. & Sternberg, M. J. E. (1987) *J. Mol. Biol.* **198,** 295–310.
20. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 488–492.
21. Alber, T. & Matthews, B. W. (1987) *Methods Enzymol.* **154,** 511–533.
22. Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E. & Dahlquist, F. W. (1989) *Methods Enzymol.* **177,** 44–73.
23. Schmid, M. F., Weaver, L. H., Holmes, M. A., Grütter, M. G., Ohlendorf, D. H., Reynolds, R. A., Remington, S. J. & Matthews, B. W. (1981) *Acta Crystallogr.* **A37,** 701–710.
24. Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. (1987) *Acta Crystallogr.* **A34,** 489–501.
25. Becktel, W. J. & Baase, W. A. (1985) *Anal. Biochem.* **150,** 258–263.
26. Becktel, W. J. & Baase, W. A. (1987) *Biopolymers* **26,** 619–623.
27. Levitt, M. (1974) *J. Mol. Biol.* **82,** 393–420.
28. Tanford, C. (1962) *J. Am. Chem. Soc.* **84,** 4240–4247.