in our model; and third, Horizon "island" may be younger than 110 million years ago. The third possibility is supported by the age of shallow water fossils (86.5 to 91 million years ago) on the summit of Horizon Guyot.
32. M. G. Debiche, A. Cox, D. C. Engebretson, *Geol. Soc. Am. Spec. Pap.* **207**, 1 (1987).
33. C. R. Newton, *Science* **242**, 385 (1988); *ibid.* **249**, 681 (1990).
34. G. D. Stanley, *Palaios* **3**, 170 (1988); *Can. Soc. Pet. Geol. Mem.* **13**, 766 (1989); _____ and M. T. Whalen, *J. Paleontol.* **63**, 800 (1990); G. D. Stanley and L. Beauvais, *ibid.* **64**, 352 (1990); G. D. Stanley, *Third Int. Symp. Shallow Tethys* (1990), abstract 64.
35. G. D. Stanley and T. E. Yancey, *Science* **249**, 680 (1990).
36. C. Newton, *Geology* **15**, 1126 (1987); E. A. Kay, personal communication.
37. C. H. Stevens, *Geology* **11**, 603 (1983).
38. Z. Ben-Avraham, *Sci. Am.* **69**, 291 (1981); A. Nur and Z. Ben-Avraham, in *The Evolution of the Pacific Ocean Margins*, Z. Ben-Avraham, Ed. (Oxford Univ. Press, New York, 1989), pp. 7–19.
39. An opposite argument by W. C. Fallow, that migration across the Pacific has steadily increased over Mesozoic and Cenozoic time due to narrowing of the Pacific, fails to take into account the effect of island stepping stones on dispersal [W. C. Fallow, *Am. J. Sci.* **283**, 166 (1983)].
40. V. G. Springer, *Smithson. Contrib. Zool. 367* (1982).
41. D. C. Engebretson, A. Cox, R. G. Gordon, *Geol. Soc. Am. Spec. Pap. 206* (1985).
42. L. J. Henderson and R. G. Gordon, *Eos* **62**, 1028 (1981).
43. A. Cox and R. B. Hart, *Plate Tectonics: How It Works* (Blackwell, Palo Alto, CA, 1986).
44. H. S. Ladd, W. A. Newman, N. F. Sohl, in *Proceedings of the Second International Coral Reef Symposium*, A. M. Cameron *et al.*, Eds. (Great Barrier Reef Committee, Brisbane, Australia), pp. 513–522.
45. E. L. Winterer and C. V. Metzler, *J. Geophys. Res.* **89**, 9969 (1984).
46. We thank E. A. Kay, W. Newman, G. Stanley, and three anonymous reviewers for constructive comments on an earlier draft of the article. R.W.G is grateful to the National Sea Grant Program (grant NA-81AA-D-00070) and the State of Hawaii, Department of Land and Natural Resources for financial support of this research. SOEST Contribution number 2691.

Research Article

Response of a Protein Structure to Cavity-Creating Mutations and Its Relation to the Hydrophobic Effect

A. E. ERIKSSON,* W. A. BAASE, X.-J. ZHANG, D. W. HEINZ, M. BLABER, E. P. BALDWIN, B. W. MATTHEWS†

Six "cavity-creating" mutants, Leu^{46} → Ala (L46A), L99A, L118A, L121A, L133A, and Phe^{153} → Ala (F153A), were constructed within the hydrophobic core of phage T4 lysozyme. The substitutions decreased the stability of the protein at pH 3.0 by different amounts, ranging from 2.7 kilocalories per mole (kcal mol^{-1}) for L46A and L121A to 5.0 kcal mol^{-1} for L99A. The double mutant L99A/F153A was also constructed and decreased in stability by 8.3 kcal mol^{-1}. The x-ray structures of all of the variants were determined at high resolution. In every case, removal of the wild-type side chain allowed some of the surrounding atoms to move toward the vacated space but a cavity always remained, which ranged in volume from 24 cubic angstroms (Å^3) for L46A to 150 Å^3 for L99A. No solvent molecules were observed in any of these cavities. The destabilization of the mutant Leu → Ala proteins relative to wild type can be approximated by a constant term (~2.0 kcal mol^{-1}) plus a term that increases in proportion to the size of the cavity. The constant term is approximately equal to the transfer free energy of leucine relative to alanine as determined from partitioning between aqueous and organic solvents. The energy term that increases with the size of the cavity can be expressed either in terms of the cavity volume (24 to 33 cal mol^{-1} Å^{-3}) or in terms of the cavity surface area (20 cal mol^{-1} Å^{-2}). The results suggest how to reconcile a number of conflicting reports concerning the strength of the hydrophobic effect in proteins.

IT IS GENERALLY AGREED THAT THE HYDROPHOBIC EFFECT IS the major factor in stabilizing the folded structures of globular proteins [see, for example, the recent reviews by Dill (*1*) and Sharp (*2*)]. Until recently, it has also been generally agreed that the strength of the hydrophobic effect, that is, the energy of stabilization provided by the transfer of hydrocarbon surfaces from solvent to the interior of a protein, is about 25 to 30 cal mol^{-1} Å^{-2} (*3*). However, some recent studies in which site-directed mutagenesis and protein denaturation were used suggest that the strength of the hydrophobic effect might be much greater. In a typical experiment, a hydrophobic residue within the core of a protein is substituted by a smaller hydrophobic residue and the resulting change in the stability of the folded versus the unfolded (or denatured) form of the protein is taken as a measure of the difference between the hydrophobic stabilization provided by the two amino acids. Such experiments carried out with different proteins (*4–6*) or at different sites within the same protein (7) have, however, given variable results. For example, Shortle *et al.* replaced each of the leucines in staphylococcal nuclease with alanine and found that the decrease in free energy of protein folding ranged from 1.6 to 5.8 kcal mol^{-1} (7). The latter value corresponds to stabilization of ~80 cal mol^{-1} Å^{-2}, a value about four times that estimated from solvent transfer experiments (*3, 8–10*). The reason for this discrepancy has not been resolved and remains the subject of debate. A principal difficulty in addressing

The authors are in the Institute of Molecular Biology, Howard Hughes Medical Institute, and Department of Physics, University of Oregon, Eugene, OR 97403.

*Present address: Department of Molecular Biology, Biomedical Center, Box 590, S-751 24 Uppsala, Sweden.
†To whom correspondence should be addressed.

Table 1. X-ray data collection and refinement statistics. Data were collected either by oscillation photography or with a Xuong-Hamlin area detector system. The cell dimensions of pseudo–wild-type lysozyme are $a = b = 60.9$ Å and $c = 96.8$ Å. R_{merge} gives the agreement between intensities measured on different films (*29*), R is the crystallographic residual giving the agreement between the refined structural model and the observed structure amplitudes, and $\Delta_{bond\ length}$ and $\Delta_{bond\ angle}$ give the average deviations of the bond lengths and bond angles in the final refined model from "ideal" values.

Parameter	L46A	L99A	L133A	L118A	L121A	F153A	L99A/F153A
			Data collection				
Cell dimensions							
a, b (Å)	60.9	60.9	61.3	61.1	61.0	61.1	61.4
c (Å)	96.4	96.8	96.2	96.9	96.8	95.5	95.6
Resolution (Å)	1.8	1.75	1.9	1.8	2.0	2.0	1.9
Unique reflections	14,062	14,445	10,953	16,392	13,324	10,935	12,333
Completeness of data (percent)	76	65	71	75	88	77	72
R_{merge}	0.050	0.074	0.082	0.061	0.035	0.066	0.058
			Refinement				
R	0.153	0.156	0.162	0.157	0.162	0.163	0.171
$\Delta_{bond\ length}$ (Å)	0.018	0.016	0.015	0.015	0.013	0.016	0.017
$\Delta_{bond\ length}$ (degrees)	2.4	2.0	2.3	2.0	2.0	2.2	2.2

this problem has been the lack of relevant structural data. How does a protein structure respond when a bulky hydrophobic residue such as leucine is replaced by a smaller residue such as alanine? Does the protein structure remain essentially unchanged (*5*), or is there structural rearrangement to avoid the creation of a cavity (*2, 11*)? If cavities are created, do they contain solvent (*5*)?

In this article we describe the high-resolution crystal structures and the thermal stabilities of six "cavity-creating" mutants in T4 lysozyme. Five of the mutants were created by replacements of Leu residues with Ala; the sixth is a Phe → Ala substitution. A double mutant combining the Phe → Ala substitution with one of the Leu → Ala replacements was also constructed. The changes in thermal stability associated with each of the mutations vary substantially from case to case but can be correlated with the size of the cavity that is created by the mutation. This result provides a rationalization for the variability that has been observed in the directed mutagenesis experiments and also reconciles the apparent differences in the hydrophobic stabilization estimated by mutagenesis and by solvent transfer experiments.

Experimental. The side chains of the six amino acids chosen for substitution, Leu^{46}, Leu^{99}, Leu^{118}, Leu^{121}, Leu^{133}, and Phe^{153}, are all buried within the protein and inaccessible to solvent except for Leu^{118}, which has 10 percent solvent accessibility. The Leu^{46} residue is within the amino-terminal domain; the other residues constitute part of the hydrophobic core within the carboxyl-terminal domain of T4 lysozyme.

Replacement of these bulky amino acids with alanine achieves a relatively large change in side-chain volume but avoids possible complications that might accrue from replacement with the amino acid glycine. Replacements with other amino acids have been made but are more complicated to analyze than the Ala replacements (*12*).

The mutant L133A was created by using the gene for wild-type lysozyme as a template. All of the other mutants were constructed with the gene for a pseudo–wild-type lysozyme, Cys^{54} → Thr/Cys^{97} → Ala (C54T/C97A, or WT*), in which the two Cys residues were replaced with Thr and Ala, respectively (*13*). This cysteine-free lysozyme has properties similar to the wild type but displays better reversibility in thermal denaturation experiments (*13, 14*). The full identifications of these mutants are, for example, L46A/C54T/C97A, but are referred to below as L46A, L99A, L118A, L121A, F153A, and L99A/F153A.

Methods for generation and purification of the mutants were as described (*4, 15*). All of the proteins were judged to be at least 95 percent pure by the analytical reversed-phase high-performance liquid chromatography (HPLC) and SDS–polyacrylamide gel electrophoresis. Crystals were grown from ~2.2 M phosphate solutions at pH ~6.7 by batch or hanging drop methods (*16*).

Prior to x-ray data collection, the crystals were equilibrated in a standard mother liquor containing 1.05 M K_2HPO_4, 1.26 M $NaHPO_4$, 0.23 M NaCl, and 1.4 mM 2-mercaptoethanol at pH 6.7. Procedures for data collection and crystallographic refinement (*17*) were similar to those used for other lysozyme mutants (*18*). Essential statistics are summarized in Table 1. Coordinates will be deposited in the Brookhaven Protein Data Bank.

Thermal unfolding experiments for all of the proteins were carried out in 0.025 M potassium chloride, 0.020 M potassium phosphate, pH 3.01, with an in-cell probe as previously described (*19*). Protein concentrations were between 0.01 and 0.03 mg ml^{-1} in all cases. These conditions were chosen to optimize reversibility, which was greater than 95 percent except for the dual cavity mutant, L99A/F153A, in which case it was 80 percent. In this buffer the dual cavity mutant was only ~50 percent folded at room temperature and was unfolded for the greatest length of time during melting. Circular

Table 2. Thermodynamic data for mutant lysozymes; T_m is the melting temperature, ΔH is the change in enthalpy of unfolding at T_m, and $\Delta\Delta G$ is the free energy of unfolding. All of the measurements were performed at pH 3.01. For L133A, $\Delta\Delta G$ is the difference between the free energy of folding of the mutant and wild-type protein. In all other cases, $\Delta\Delta G$ is the difference between the mutant and the pseudo–wild-type lysozyme, WT*, from which it was constructed. Because of the wide range of melting temperatures, isothermal $\Delta\Delta G$ values were calculated at 51.8°C with the use of a thermodynamic model (*30*) that includes a constant change in heat capacity, ΔC_p, estimated in this case to be 2.5 kcal mol^{-1} $degree^{-1}$. The accuracy of the $\Delta\Delta G$ values is limited by choice of this model and of ΔC_p itself (*31*). For the plots in Fig. 2, calculation of $\Delta\Delta G$ values with ΔC_p ranging from 1.5 to 4.0 kcal mol^{-1} $degree^{-1}$ had almost no effect (less than 4 percent) on the intercept values, but the slopes were more sensitive to the choice of ΔC_p, which varied ±20 percent over this range. Thus, for the intercept values, random error is dominant, which we estimate from the scatter of the points to be ±10 percent, whereas for the slopes the uncertainty associated with the choice of the constant ΔC_p model and with ΔC_p itself is more significant.

Lysozyme	T_m (°C)	ΔH (kcal mol^{-1})	$\Delta\Delta G$ (kcal mol^{-1})
WT	53.5	130.0	0.7
WT*	51.8	118.8	0.0
L46A	43.2	89.8	−2.7
L118A	39.6	75.5	−3.5
L121A	42.5	81.0	−2.7
L99A	36.1	79.6	−5.0
L133A	42.9	91.5	−3.6
F153A	39.5	74.9	−3.5
L99A/F153A	10.0	6.0	−8.3

dichroism (CD) data at 223 nm taken as a function of temperature were analyzed by means of a two-state van't Hoff procedure to give the temperature of melting, T_m, and enthalpy of unfolding, ΔH, at T_m (Table 2). In the case of the dual mutant, the folded baseline was taken as that of WT*, the background in which these mutations were constructed.

Results. The crystal structures of four representative single mutant proteins in the vicinity of the respective amino acid substitutions are shown in Fig. 1. In each case the structure relaxes somewhat in response to the mutation, although the amount of structural change differs significantly from case to case. Surrounding atoms tend to move slightly toward the space vacated by the removed side chain. Typically, these movements are a few tenths of an angstrom and are therefore better characterized as slight adjustments rather than repacking. Two extreme examples are provided by L99A and F153A. When Leu[99] is replaced by Ala the structure hardly changes at all. The largest shifts are in atoms within the side chains of Val[87] and Tyr[88], which shift, respectively, 0.4 Å away from and 0.5 Å toward the mutation site (Fig. 1B). The backbone atoms of residues 87 and 88 also move by ~0.2 Å. Other than this, the mutant and wild-type structures are virtually identical.

In contrast, when Phe[153] is replaced with Ala the structural adjustments are substantially larger (Fig. 1D). One end of the α helix that includes residue 153 moves ~0.8 Å toward the cavity created by the removal of the Phe side chain. The maximal shift (1.0 Å) is for the carbonyl oxygen of residue 153. The distal region of the side chain of Met[102] also moves 0.6 Å toward the space vacated. The effect of these changes is to reduce in size, but not to eliminate, the putative cavity created by the Phe → Ala replacement. The side chain of Leu[121] is in contact with that of Phe[153] (Fig. 1D), and when Leu[121] is replaced with Ala the structural changes are relatively large and, in many respects, similar to those seen for the Phe[153] → Ala replacement. The structural changes in the double mutant L99A/F153A are essentially the combination of those seen in the respective single mutants. Because mutant F153A has relatively large structural changes and L99A has very few changes, the structure of the double mutant is similar to that for F153A, except for the loss of the side-chain atoms of Leu[99] (*20*).

In the case of L46A, The Cδ1 atom of Ile[27] moves 0.6 Å and some backbone atoms in the vicinity of residues 45 and 46 and 54 to 56 move up to 0.4 Å (Fig. 1A). The structural changes observed for L133A are somewhat larger and more distributed (Fig. 1C). The largest shift, 0.8 Å, is in the β-carbon of Leu[133], the residue that is substituted by Ala. Backbone atoms within the short α helix that includes residues 108 to 113 can also move as much as 0.6 to 0.7 Å. Atoms within the side chains of Met[106], Phe[114], Leu[118], and Leu[121] can move as much as 0.4 to 0.5 Å.

Energetics of cavity-forming mutations and the hydrophobic effect. The mutant structures described here, four of which are

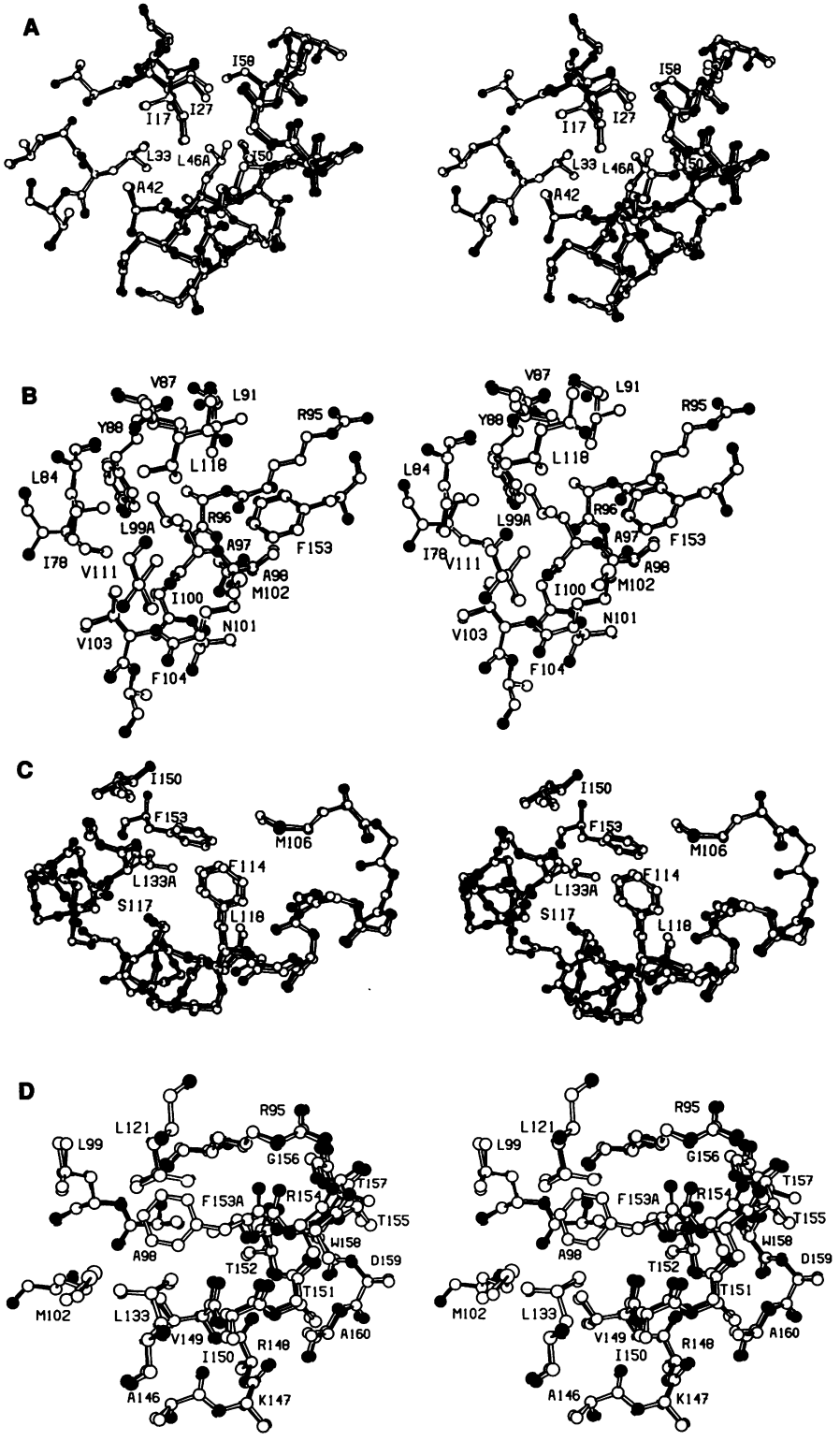

Fig. 1. Superposition of the structures of cavity-creating mutant lysozymes (solid bonds) on the structure of wild-type (WT) or pseudo–wild-type lysozyme (WT*) (open bonds) in the vicinity of the amino acid substitution (*32*). Oxygen atoms are drawn as solid circles, nitrogen as half-solid circles, and carbon as open circles. (**A**) L46A versus WT*. (**B**) L99A versus WT*. (**C**) L133A versus WT. (**D**) F153A versus WT*.

shown in Fig. 1, A to D, provide examples of the changes that are likely to occur in a typical protein structure when a bulky buried hydrophobic residue is replaced by a smaller one. The results do not support the liquid oil drop model of a protein interior in which the hydrophobic residues segregate themselves into a liquid-like cluster from which water and other polar groups are excluded. Rather, the core is seen to include some parts that are relatively rigid, and other parts that are more flexible, but not disordered. The protein structure does not remain completely invariant, but neither does it repack so as to fill the space vacated by the substituted amino acid. Side chains that contact the bulky residue in the parent structure usually, but not always, tend to move toward the cavity that is created when the bulky residue is removed. Maximal shifts of 0.4 to 1.0 Å were observed but varied from case to case. Shifts in backbone atoms tended to be smaller (up to ~0.6 Å) than side-chain atoms, although in the case of F153A backbone shifts up to 1.0 Å were observed.

In every case a cavity remained in the mutant structure. As summarized in Table 3, the size of this cavity varies substantially from case to case. In wild-type T4 lysozyme there is a cavity ("cavity I"), the center of which is within 3.5 to 4.5 Å of Leu^{99}, Met^{102}, Phe^{144}, Leu^{118}, Leu^{121}, Ser^{117}, and Leu^{133} (Fig. 1, B and C). A second, somewhat smaller cavity (cavity II) also occurs in the vicinity of Leu^{121}, Ala^{130}, Leu^{133}, and Phe^{153}. Leu^{99}, Leu^{118}, Leu^{121}, Leu^{133}, and Phe^{153} are five of the substituted residues. The effect of the mutations L99A, L118A, F153A, and the double mutant is to increase the size of this preexisting cavity (Table 3). In L121A and L133A, cavities I and II coalesce to form a single connected cavity. Mutant L46A simply creates a new cavity within the amino terminal domain (cavity III). In no case is there evidence to suggest that an ordered water molecule occupies the cavity that is created.

Another way to assess the consequences of a given mutation is to compare the size of the actual cavity in the mutant structure (V_c) with the cavity that would have been formed if the protein structure had remained exactly the same as in wild type. This model cavity volume (V_{model}) is also included in Table 3. In general V_c is less than V_{model}, but in the case of L99A the volume of the resultant cavity is actually slightly greater (17 $Å^3$) than the model.

The changes in protein stability associated with the six cavity-creating mutations show substantial variation (Table 2). The most destabilizing replacement was Leu^{99} → Ala, which decreased stability by 5.0 kcal mol^{-1}. The replacement of Leu^{46} with Ala was least destabilizing (2.7 kcal mol^{-1}). This variability is very much in line with Leu → Ala replacements reported elsewhere (*5, 7*).

It should be noticed in Table 3 that the Leu^{99} → Ala replacement is the one that causes the greatest increase of cavity volume within the folded protein. The replacement Leu^{46} → Ala created the smallest cavity. It appears that the Leu → Ala replacements destabilize the protein not only because of the reduction in hydrophobic stabilization of Ala relative to Leu, but also because there is an energetic cost associated with the creation of a cavity in the folded protein. Roughly speaking, the larger the cavity that is created, the more destabilizing is the replacement.

The Leu^{99} → Ala replacement is the one for which the protein structure remains practically unchanged (Fig. 1B). Apparently the protein structure in the vicinity of Leu^{99} is relatively rigid and is unable to relax in response to the Ala replacement. In the cases of L133A and F153A the protein structure can relax somewhat in response to the mutation. As a result the size of the cavity is reduced. This reduction in turn appears to offset the potential energetic costs of these cavity-creating replacements. In the case of L46A, the structural changes are not particularly large, but nevertheless the size of the cavity created by this mutant is the smallest of the six examples.

These observations are consistent with the results noted for other mutant protein structures. In particular, mutant proteins can relax or adapt their structures to ameliorate the consequences of potentially destabilizing lesions (*6, 18, 21–23*). Also, the most destabilizing replacements tend to occur in the most rigid parts of a protein structure (*24*), presumably because in such cases it is energetically costly for the protein structure to adjust in response to the mutation.

The relation between increase in cavity size associated with a given replacement and reduction in protein stability is shown in Fig. 2. Five of the replacements are of the form Leu → Ala. We also included the Phe^{153} → Ala replacement because a Phe side chain is approximately the same size as a Leu and is expected to have approximately the same hydrophobic strength (*9, 10*). In addition, we constructed the lysozyme variant Leu^{153} and showed that its structure and thermal stability are similar to the Phe^{153} protein (*12*). Therefore, in this case we know that the Phe^{153} → Ala result is similar to that which would be obtained in comparing the Leu^{153} with the Ala^{153} variant. The inclusion of the double mutant L99A/F153A in Fig. 2 provides an example where the overall cavity volume is large (207 $Å^3$; equivalent to a 5.9 Å cube) and the protein is quite unstable. We take this double mutant to be approximately representative of a protein in which two leucines have each been replaced with alanine.

Subject to these and other limitations (see below), Fig. 2 suggests that the decrease in protein stability associated with a Leu → Ala replacement consists of a constant energy term of 1.9 kcal mol^{-1} plus a second energy term that depends on the size of the cavity created by the substitution. The magnitude of the constant energy term agrees remarkably well with values of 1.7 to 1.9 kcal mol^{-1} for the difference in hydrophobicity of leucine and alanine estimated by transfer from water to ethanol (*8*), octanol (*9*), or *N*-methylacetamide (*10*). A value of 1.9 kcal mol^{-1} for leucine relative to alanine

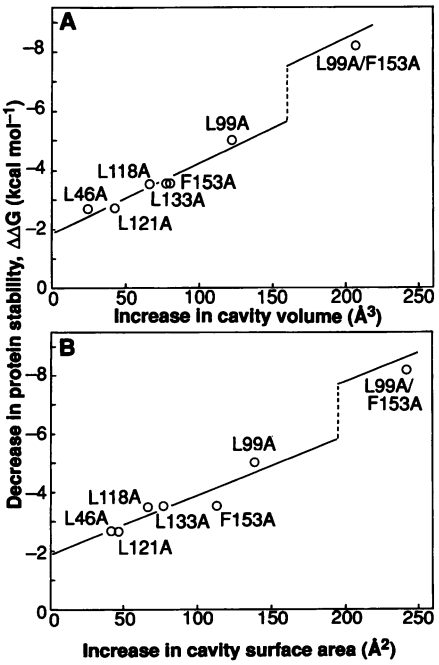

Fig. 2. (**A**) Change in the free energy of unfolding ($\Delta\Delta G$) of mutant lysozymes relative to wild type (from Table 2) plotted as a function of the cavity volume created by the amino acid substitution or substitutions (from Table 3). The straight line was fitted by least squares to the data for the six single mutants by the equation $\Delta\Delta G = a + b\,\Delta V$ where $a = -1.9$ kcal mol^{-1}, $b = -0.024$ kcal mol^{-1} $Å^{-3}$, and ΔV is the increase in cavity volume. In the vicinity of the double mutant an additional −1.9 kcal mol^{-1} is added to $\Delta\Delta G$ to reflect the fact that in this case both a Leu and a Phe (which is here regarded as equivalent to a second Leu) have been replaced with Ala. See legend to Table 2 for analysis of errors in $\Delta\Delta G$. In addition, there are other uncertainties in both the slope and intercept (see discussion). (**B**) Change in $\Delta\Delta G$ of mutant lysozymes relative to wild-type plotted as a function of the increase in cavity surface area created by the amino acid substitution or substitutions. The equation of the straight line that best fits the data for the single mutants is $\Delta\Delta G = c + d\Delta A$ where $c = -1.9$ kcal mol^{-1}, $d = -0.020$ kcal mol^{-1} $Å^{-2}$, and ΔA is the increase in cavity surface area. In the vicinity of L99A/F153A an additional −1.9 kcal mol^{-1} has been added to $\Delta\Delta G$ to reflect the fact that the double mutant is roughly equivalent to two Leu-to-Ala replacements.

corresponds to ~25 cal mol^{-1} Å^{-2}. The cavity-dependent energy term, at least for the Leu → Ala (or Phe → Ala) replacements in T4 lysozyme, is 24 cal mol^{-1} Å^{-3}. Expressed in terms of area, the energy cost is 20 cal mol^{-1} for each square angstrom of cavity surface created.

This result suggests a way to reconcile the different values for the hydrophobic strength obtained, on the one hand, by solvent transfer experiments, and, on the other hand, by directed mutagenesis. The analysis in Fig. 2 suggests that the change in energy associated with the replacement of a buried leucine with an alanine consists of two parts. The first part is a constant and is presumed to depend only on the identities of the two amino acids being compared, in this instance Leu (or Phe) and Ala. Physically this energy term can be considered as the difference in energy required to desolvate (that is, transfer from solvent to protein interior) a leucine relative to an alanine. The second part of the change in protein stability associated with the Leu → Ala replacements (Fig. 2, A and B) depends on the context within the three-dimensional structure and the way in which the protein structure adjusts in response to the substitution. Two extreme situations might exist. In one case a Leu → Ala replacement is constructed and the protein structure remains completely unchanged (compare with L99A). In this situation the size of the created cavity is large and the mutant (Ala) protein is maximally destabilized. In the other extreme, the protein structure relaxes in response to the Leu → Ala substitution, fills the space occupied by the Leu side chain and so avoids the formation of any cavity whatsoever. In this case, and in the absence of any other energy terms that might come into play, the decrease in energy of the mutant protein relative to wild type would reduce to the constant energy term described above, that is, ~1.9 kcal mol^{-1}.

In order to explain the observed changes in protein stability caused by cavity-creating mutations it is not necessary to suggest that the hydrophobic effect should be counted twice, once for the residue removed and once for the cavity created (*5*). Neither is it necessary to argue that the strength of the hydrophobic effect needs to be revised from the accepted value of about 25 to 30 cal mol^{-1} Å^{-2} to a new value of 43 to 47 cal mol^{-1} Å^{-2} (*11, 25*).

When a bulky internal side chain such as leucine is replaced with alanine, many favorable van der Waals contacts in the folded protein are often lost. The creation of a cavity would remove favorable van der Waals interactions from the folded protein, and we presume that

Table 3. Size of the created cavities. Cavity volumes (V_c) and surface areas (S_c) were calculated from the refined coordinates of the different structures by using the program of Connolly (*27*). The model cavity volume (V_{model}), for example for L99A, was obtained by taking the coordinates for WT* lysozyme, truncating the Leu^{99} coordinates to Ala, and calculating the volume of the resultant cavity assuming no change in structure. The cavity surface is the area swept out by a sphere of radius 1.2 Å as it rolls over the cavity surface. The volume of the cavity is the volume contained within the cavity surface (that is, the total volume that can be occupied by the 1.2 Å radius sphere). The increases in cavity volume and area are the overall increases associated with the mutation. Mutant L133A is calculated relative to WT. All of other mutants are relative to pseudo–wild-type lysozyme, WT* (see text). Cavities I and II are located within the hydrophobic core of the carboxyl-terminal domain. They are present in L46A with sizes essentially the same as in wild type but were ignored because the mutation is in the amino-terminal domain.

Lysozyme	Cavity	Center of cavity			Observed cavity volume V_c (Å^3)	Model cavity volume V_{model} (Å^3)	Increase in cavity volume [V_c(mut)-V_c(WT)] (Å^3)	Cavity surface area S_c (Å^2)	Increase in area [S_c(mut)-S_c(WT)] (Å^2)
		X (Å)	Y (Å)	Z (Å)					
WT	I	28.2	8.9	0.7	40.8			62.7	
	II	32.4	4.0	−5.0	23.1			43.5	
					63.9			106.2	
WT*	I	28.0	8.9	0.6	33.5			52.4	
	II	32.1	4.0	−4.9	20.4			39.5	
					53.9			91.9	
L46A	III	38.8	18.4	24.8	24.3	47.8	24.3	42.2	42.2
L118A	I	26.3	8.4	1.2	105.1	121.6*	71.6	127.2	74.8
	II	32.0	3.9	−4.9	15.3	20.4	−5.1	31.5	−8.0
					120.4	142.0	66.5	158.7	66.8
L121A	I + II	29.9	6.1	−2.4	96.9	148.7	43.0	139.2	47.3
L99A	I	27.3	6.8	3.3	149.5	138.7	116.0	182.3	129.9
	II	32.2	4.3	−4.8	26.9	6.5	47.7	8.2	8.2
					176.4	159.1	122.5	230.0	138.1
L133A	I + II	31.0	7.4	−2.3	141.7	172.6	77.8	183.3	77.1
F153A	I	30.2	6.3	0.7	100.7	135.2	67.2	151.9	99.5
	II	32.2	4.4	−4.6	32.1	20.4	11.7	53.7	14.2
					132.8	155.6	78.9	205.6	113.7
L99A/F153A	I	28.8	5.9	2.4	221.6	238.7	188.1	271.2	218.8
	II	32.3	4.6	−4.6	39.3	20.4	18.9	62.4	22.9
					260.9	259.1	207.0	333.6	241.7

*Leu^{118} is 10 percent solvent-exposed. As a result, in the model cavity calculation with a probe sphere of radius 1.2 Å the cavity connects with the exterior of the protein. In order to obtain a meaningful cavity volume, dummy atoms were placed at van der Waals distance from the solvent-accessible surface of Leu^{118}.

this is the physical basis for the cavity-dependent part of the destabilization associated with cavity-creating mutants. Readjustments of the protein that reduce the size of the cavity presumably increase the overall compactness of the protein, add new van der Waals interaction energy, and so tend to restore protein stability.

With the limited set of data available at present it is not possible to make a meaningful distinction as to whether the energy associated with the creation of a cavity is better described by the dependence on volume (Fig. 2A) or on surface area (Fig. 2B). Numerically, the value of 20 cal mol^{-1} $Å^{-2}$ is roughly comparable with the values of 26 to 31 cal mol^{-1} $Å^{-2}$ for the creation of a cavity against surface tension in organic solvents such as ethanol, isopropanol, or hexane (*26*).

It also should be noted that the calculation of cavity volume is nontrivial. We have estimated cavity volume and surface area by rolling a sphere of radius 1.2 Å over the surface of the cavity (*27*). If the data shown in Fig. 2 are recalculated with a probe radius of 1.4 Å, the slopes are essentially unchanged but, for both figures, the intercept increases to 2.1 kcal mol^{-1}. A more fundamental problem is that there may be cavities or packing defects in the wild-type structure adjacent to the mutated residue that are too small to be detected by the 1.2 Å probe (*28*). The volume of such cavities can be estimated as $(V_{model} - V_{side\ chain})$, where V_{model} is as defined previously (see Table 3) and $V_{side\ chain}$ is the volume of the side chain that is removed (47.8 to 48.6 $Å^3$ for the Leu → Ala replacements; 78.9 $Å^3$ for Phe^{153} → Ala). Thus, the increase in cavity volume associated with each mutant can be estimated as $(V_c - V_{model} + V_{side\ chain})$. Replotting Fig. 2A on this basis yields an intercept of 2.3 kcal mol^{-1} and a slope of 33 cal $Å^{-3}$. The estimate of 33 cal $Å^{-3}$ attempts to explicitly allow for packing defects in the wild-type structure but ignores possible defects that may be introduced in the mutant structure as a consequence of the amino acid replacement. The estimate of 24 cal $Å^{-3}$ from Fig. 2A assumes that the effect of packing defects present in the wild-type structure in the vicinity of the substituted residue would cancel with defects introduced in the mutant structure. As well as questions concerning the best method to estimate cavity volume, there are other reasons to expect that the structure and stabilities of "cavity-creating" mutants would be modified by factors that have not been considered here. For example, the entropy cost of transferring the amino acid in question from solvent to the interior of the protein can vary from case to case. There is also the possibility that the residue being substituted may be under strain in the folded protein. Effects of mutations on the unfolded protein might also be important. For these reasons one cannot anticipate that there would be a strict linear relation between $\Delta\Delta G$ and cavity size. At best, one might expect a general trend, as seen in Fig. 2. The hope, however, is that the overall principles suggested by the data in Fig. 2 would be supported by additional data and would provide a general framework within which to quantitate the strength of the hydrophobic effect and the energy cost of cavity creation in proteins.

REFERENCES AND NOTES

1. K. A. Dill, *Biochemistry* **29**, 7133 (1990).
2. K. A. Sharp, *Curr. Biol.* **1**, 171 (1991).
3. C. Tanford, *The Hydrophobic Effect* (Wiley, New York, ed. 2, 1980); C. Chothia, *Nature* **248**, 338 (1974); F. M. Richards, *Annu. Rev. Biophys. Bioeng.* **6**, 151 (1977).
4. M. Matsumura, W. J. Becktel, B. W. Matthews, *Nature* **334**, 406 (1988); M. Matsumura, J. A. Wozniak, S. Dao-pin, B. W. Matthews, *J. Biol. Chem.* **264**, 16059 (1989).
5. J. T. Kellis, K. Nyberg, D. Sali, A. R. Fersht, *Nature* **333**, 784 (1988); J. T. Kellis, K. Nyberg, A. R. Fersht, *Biochemistry* **28**, 4914 (1989).
6. W. S. Sandberg and T. C. Terwilliger, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1706 (1991).
7. D. Shortle, W. E. Stites, A. K. Meeker, *Biochemistry* **29**, 8033 (1990).
8. Y. Nozaki and C. Tanford, *J. Biol. Chem.* **246**, 2211 (1971).
9. J.-L. Fauchere and V. Pliska, *Eur. J. Med. Chem. Chim. Ther.* **18**, 369 (1983).
10. S. Damodaran and K. B. Song, *J. Biol. Chem.* **261**, 7220 (1986).
11. K. A. Sharp, A. Nicholls, R. F. Fine, B. Honig, *Science* **252**, 106 (1991); K. A. Sharp, A. Nicholls, R. Friedman, B. Honig, *Biochemistry* **30**, 9686 (1991).
12. A. E. Eriksson *et al.*, unpublished results.
13. M. Matsumura, W. J. Becktel, M. Levitt, B. W. Matthews, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6562 (1989).
14. R. Wetzel, L. J. Perry, W. A. Baase, W. J. Becktel, *ibid.* **85**, 401 (1988).
15. T. Alber and B. W. Matthews, *Methods Enzymol.* **154**, 511 (1987); D. C. Muchmore, L. P. McIntosh, C. B. Russell, D. E. Anderson, F. W. Dahlquist, *ibid.* **177**, 44 (1989); A. R. Poteete, S. Dao-pin, H. Nicholson, B. W. Matthews, *Biochemistry* **30**, 1425 (1991).
16. L. H. Weaver and B. W. Matthews, *J. Mol. Biol.* **193**, 189 (1987); A. E. Eriksson, W. A. Baase, B. W. Matthews, unpublished results.
17. D. E. Tronrud, L. F. Ten Eyck, B. W. Matthews, *Acta Crystallogr.* **A43**, 489 (1987).
18. S. Dao-pin, T. Alber, W. A. Baase, J. A. Wozniak, B. W. Matthews, *J. Mol. Biol.* **221**, 647 (1991); details will be published elsewhere (A. E. Eriksson *et al.*, unpublished results).
19. S. Dao-pin, W. A. Baase, B. W. Matthews, *Proteins* **7**, 198 (1990).
20. A. E. Eriksson *et al.*, unpublished results.
21. B. W. Matthews, *Biochemistry* **26**, 6885 (1987).
22. T. Alber *et al.*, *Science* **239**, 631 (1988).
23. D. E. McRee *et al.*, *J. Biol. Chem.* **265**, 14234 (1990).
24. T. Alber, S. Dao-pin, J. A. Nye, D. C. Muchmore, B. W. Matthews, *Biochemistry* **26**, 3754 (1987).
25. L. R. DeYoung and K. A. Dill, *J. Phys. Chem.* **94**, 801 (1990).
26. The method of calculation follows that given by A. Fersht, in *Enzyme Structure and Mechanism* (Freeman, San Francisco, 1977). Values of the surface tension were from R. C. Weast, Ed., *Handbook of Chemistry and Physics* (CRC Press, Cleveland, ed. 57, 1977).
27. M. L. Connolly, *Science* **221**, 709 (1983).
28. We thank one of the referees for emphasizing this problem.
29. M. F. Schmid *et al.*, *Acta Crystallogr.* **A37**, 701 (1981).
30. J. F. Brandts and L. Hunt, *J. Am. Chem. Soc.* **89**, 4826 (1967); W. J. Becktel and J. A. Schellman, *Biopolymers* **26**, 1859 (1987).
31. S. Dao-pin, D. E. Anderson, W. A. Baase, F. W. Dahlquist, B. W. Matthews, *Biochemistry* **30**, 11521 (1991).
32. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

33. We are most grateful to H. Nicholson, S. Dao-pin, L. Weaver, and D. Tronrud for advice and help with procedures of mutagenesis, data collection, and crystallographic refinement. We thank J. A. Schellman, R. L. Baldwin, J. Hurley, and A. Morton for helpful discussions and K. Sharp for communicating manuscripts prior to publication. We also thank J. Wozniak, S. Pepiot, and J. Lindstrom for expert technical assistance. A.E.E. thanks the Swedish Natural Science Research Council for support. Public Health Service postdoctoral fellowships to M.B. (no. GM13709) and E.P.B. (no. GM12989) are also gratefully acknowledged. Supported in part by NIH grant no. GM21967 and the Lucille P. Markey Charitable Trust.

18 September 1991; accepted 6 December 1991