

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/51310853>

The response of T4 lysozyme to large-to-small substitutions within the core and its relation to the hydrophobic effect

ARTICLE *in* PROTEIN SCIENCE · JANUARY 1998

Impact Factor: 2.85 · DOI: 10.1002/pro.5560070117 · Source: PubMed Central

CITATIONS

153

READS

22

4 AUTHORS, INCLUDING:



[Enoch Post Baldwin](#)
University of California, Davis
38 PUBLICATIONS 2,219 CITATIONS

[SEE PROFILE](#)

The response of T4 lysozyme to large-to-small substitutions within the core and its relation to the hydrophobic effect

JIAN XU,¹ WALTER A. BAASE, ENOCH BALDWIN,² AND BRIAN W. MATTHEWS

Institute of Molecular Biology, Howard Hughes Medical Institute and Department of Physics,
University of Oregon, Eugene, Oregon 97403

(RECEIVED June 12, 1997; ACCEPTED August 28, 1997)

Abstract

To further examine the structural and thermodynamic basis of hydrophobic stabilization in proteins, all of the bulky non-polar residues that are buried or largely buried within the core of T4 lysozyme were substituted with alanine. In 25 cases, including eight reported previously, it was possible to determine the crystal structures of the variants. The structures of four variants with double substitutions were also determined. In the majority of cases the “large-to-small” substitutions lead to internal cavities. In other cases declivities or channels open to the surface were formed. In some cases the structural changes were minimal (mainchain shifts $\leq 0.3 \text{ \AA}$); in other cases mainchain atoms moved up to 2 \AA . In the case of Ile 29 \rightarrow Ala the structure collapsed to such a degree that the volume of the putative cavity was zero. Crystallographic analysis suggests that the occupancy of the engineered cavities by solvent is usually low. The mutants Val 149 \rightarrow Ala (V149A) and Met 6 \rightarrow Ala (M6A), however, are exceptions and have, respectively, one and two well-ordered water molecules within the cavity. The Val 149 \rightarrow Ala substitution allows the solvent molecule to hydrogen bond to polar atoms that are occluded in the wild-type molecule. Similarly, the replacement of Met 6 with alanine allows the two solvent molecules to hydrogen bond to each other and to polar atoms on the protein. Except for Val 149 \rightarrow Ala the loss of stability of all the cavity mutants can be rationalized as a combination of two terms. The first is a constant for a given class of substitution (e.g., -2.1 kcal/mol for all Leu \rightarrow Ala substitutions) and can be considered as the difference between the free energy of transfer of leucine and alanine from solvent to the core of the protein. The second term can be considered as the energy cost of forming the cavity and is consistent with a numerical value of $22 \text{ cal mol}^{-1} \text{ \AA}^{-3}$. Physically, this term is due to the loss of van der Waal’s interactions between the bulky sidechain that is removed and the atoms that form the wall of the cavity. The overall results are consistent with the prior rationalization of Leu \rightarrow Ala mutants in T4 lysozyme by Eriksson et al. (Eriksson et al., 1992, *Science* 255:178–183).

Keywords: alanine; cavities; core packing; hydrophobic effect; T4 lysozyme

Globular proteins have a general pattern of folding in which the hydrophobic residues are in the interior while the polar residues are on the surface, suggesting that the hydrophobic effect drives protein folding (Bernal, 1939; Kauzmann, 1959; Dill, 1990). The contributions of individual hydrophobic residues to the overall folding free energy, however, remain controversial (Sharp et al., 1991). Several studies in which larger interior hydrophobic side-chains were replaced with smaller ones have shown that the loss in stability resulting from the same type of the replacement, e.g.,

Leu \rightarrow Ala, varies from case to case and is generally larger than predicted from solvent transfer scales (Kellis et al., 1989; Shortle et al., 1990; Eriksson et al., 1992; see also Yutani et al., 1987; Matsumura et al., 1988; Takano et al., 1995, 1997). This suggests that factors other than solvent transfer contribute to the overall free energy of folding.

Eriksson et al. (1992) found a correlation between $\Delta\Delta G$, the change of free energy of unfolding between mutant and wild-type protein, and the cavity size created, for a series of Leu \rightarrow Ala replacements at different sites in T4 lysozyme. The overall loss in stability could be separated into two terms, one associated with the size of the cavity, and a constant energy term close in value to the difference between the free energy of transfer from water to octanol of leucine and alanine (Fauchère & Pliška, 1983). If this result was general it would imply that “large-to-small” substitutions of other amino acids might be used to estimate the energy of transfer of these amino acids from the interior of the protein to solvent. To

Reprint requests to: Brian W. Matthews, Institute of Molecular Biology, University of Oregon, Eugene, OR 97403; e-mail: brian@uoxray.uoregon.edu.

¹Present address: Department of Molecular Biology/MB-13, The Scripps Research Institute, 10666 N Torrey Pines Road, La Jolla, CA 92037.

²Present address: Molecular and Cellular Biology, 220 Briggs Hall, University of California, Davis, CA 95616.

Table 1. Overview of the bulky non-polar residues in T4 lysozyme and the results of the alanine substitutions that have been made^a

Amino acid	Fractional solvent accessibility		Mutant	$\Delta\Delta G$ (kcal/mol)	Increase in cavity volume (\AA^3)	Comment
	Main chain	Side chain				
Leu 7	0.0	0.0	L7A	-2.7		Non-isomorphous crystals
Leu 13	0.56	0.13	None			Partially solvent exposed
Leu 15	0.19	0.49	None			Partially solvent exposed
Leu 32	0.58	0.43	None			Partially solvent exposed
Leu 33	0.64	0.01	L33A	-3.6		Crystals not obtained
Leu 39	0.15	0.50	L39A ^b	-0.9 ^b		Partially solvent exposed
Leu 46	0.0	0.0	L46A ^c	-2.7	24	Cavity
Leu 66	0.0	0.03	L66A	-3.9	—	Crystallization not attempted
Leu 79	0.27	0.42	None			Solvent exposed
Leu 84	0.00	0.00	L84A	-3.8	—	Large cavity with channel to surface
Leu 91	0.16	0.00	L91A	-3.1	—	Small non-isomorphous crystals
Leu 99	0.00	0.00	L99A ^c	-5.0	123	Cavity
Leu 118	0.01	0.09	L118A ^c	-3.5	67	Cavity
Leu 121	0.00	0.01	L121A ^c	-2.7	21 ^d	Cavity
Leu 133	0.00	0.01	L133A ^c	-3.6	78	Cavity
Leu 164	0.50	0.50	None			Partially solvent exposed
Ile 3	0.11	0.17	I3A ^e	-1.1 ^e	—	Crystals not obtained
Ile 9	0.81	0.52	None			Solvent exposed
Ile 17	0.53	0.02	I17A ^f	-2.7	—	Crevice formed
Ile 27	0.0	0.0	I27A	-3.1	17	Cavity
Ile 29	0.04	0.01	I29A	-2.7	0.0	No detectable cavity
Ile 50	0.79	0.07	I50A ^g	-2.0	—	Crevice formed
Ile 58	0.11	0.0	I58A	-3.2	18	Cavity
Ile 78	0.00	0.00	I78A	-1.6	—	Small non-isomorphous crystals
Ile 100	0.00	0.08	I100A	-3.4	28	Cavity
Ile 150	0.00	0.13	None			Partially solvent exposed
Val 57	0.03	0.60	None			Partially solvent exposed
Val 71	0.01	0.07	V71A	-1.5	—	Small non-isomorphous crystals
Val 75	0.00	0.34	None			Partially solvent exposed
Val 87	0.00	0.04	V87A	-1.7	31	Cavity
Val 94	0.00	0.26	V94A	-1.8	—	Surface indentation
Val 103	0.12	0.05	V103A	-2.2	23	Cavity
Val 111	0.00	0.00	V111A	-1.3	11	Cavity
Val 131	0.12	0.86	V131A ^h	0.26 ^h	—	Solvent exposed
Val 149	0.00	0.00	V149A	-3.2	3 ⁱ	Solvent bound within cavity
Met 1	0.61	0.14	None			Partially solvent exposed
Met 6	0.00	0.00	M6A	-1.9	10 ⁱ	Solvent bound within cavity
Met 102	0.01	0.03	M102A ^j	-3.3	65 ^j	Cavity
Met 106	0.64	0.24	M106A ^j	-2.3	—	Surface indentation
Met 120	0.00	0.21	M120A ^k	-0.20 ^k	—	Surface indentation
Phe 4	0.03	0.52	None			Partially solvent exposed
Phe 67	0.10	0.10	F67A	-1.9	—	Crevice formed
Phe 104	0.78	0.06	F104A ^j	-3.1	—	Crevice formed
Phe 114	0.17	0.23	None			Partially solvent exposed
Phe 153	0.01	0.02	F153A ^c	-3.5	79	Cavity

^aThe table includes all leucines, isoleucines, valines, methionines, and phenylalanines in T4 lysozyme. Alanine substitutions were made at all buried sites. Substitutions were, in general, not made at solvent exposed or partially solvent exposed sites.

^bHeinz et al. (1994).

^cEriksson et al. (1992).

^dBaldwin et al. (1996).

^eData from Matsumura et al. (1988) at pH 2.0. Mutant made in the wild-type protein.

^fUnpublished results of R. DuBose.

^gUnpublished results of R. Kuroki.

^hDao-pin et al. (1990).

ⁱVolume calculated in the presence of bound solvent.

^jE.B. and B.W.M., unpublished.

^kBlaber et al. (1995).

test this idea, we have replaced different types of non-polar residues within the core of T4 lysozyme with alanine. To some extent the analysis is limited by the types and locations of the different types of amino acids that are available within the structure of the protein. Also some of the mutations result in crevices or declivities in the molecular surface rather than internal cavities. Nevertheless, the results generally support the rationalization of Eriksson et al. (1992). In two mutants one or more water molecules were observed within the cavity hydrogen-bonded to the protein, showing that in these cases the cavities are somewhat polar.

Results

In order to provide an overview of the entire experiment, Table 1 lists all the leucines, isoleucines, valines, phenylalanines and methionines in T4 lysozyme. The table summarizes the changes in stability, success in obtaining crystals and, where known, the increase in cavity volume resulting from substitution with alanine. In order to be comparable with the leucine-to-alanine data of Eriksson et al. (1992), changes in the free energy of unfolding ($\Delta\Delta G$) were measured at pH 3.0. Additional details are given in Table 2.

Table 2. Thermostabilities of mutant lysozymes^a

Proteins	ΔT_m (°C)	ΔH (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
WT*	—	113	0
I17A	-8.4	87	-2.7
I27A	-10.1	76	-3.1
I29A	-8.2	85	-2.6
I50A	-5.8	94	-2.0
I58A	-10.4	80	-3.2
I78A	-4.7	105	-1.6
I100A	-10.7	85	-3.4
V71A	-4.7	108	-1.5
V87A	-4.9	102	-1.7
V94A	-5.0	94	-1.8
V103A	-6.6	94	-2.2
V111A	-3.7	100	-1.3
V149A	-11.0	66	-3.2
M6A	-5.7	95	-1.9
M106A	-7.1	89	-2.3
F67A	-5.7	101	-1.9
F104A	-9.7	82	-3.1
L7A	-8.1	90	-2.6
L33A	-12.3	67	-3.6
L66A	-13.4	69	-3.9
L84A	-13.4	67	-3.9
L91A	-9.7	85	-3.1

^aAll the measurements were made in 25 mM KCl, 20 mM potassium phosphate, pH 3.0. These conditions were chosen to be consistent with the prior analysis of leucine to alanine substitutions by Eriksson et al. (1992). ΔT_m is the change of melting temperature of the mutant relative to that of the wild-type (51.65°C). ΔH is the enthalpy of unfolding at the T_m of the mutant. $\Delta\Delta G$, the difference between the free energy of unfolding of the mutant and wild-type proteins, was estimated at 44°C using a thermodynamic model (Brandts & Hunt, 1967; Becktel & Schellman, 1987), which includes a constant change in heat capacity, ΔC_p , estimated in this case to be 1.8 kcal/mol deg. The estimated error in ΔT_m and $\Delta\Delta G$ is about 0.2°C and 0.15 kcal/mol, respectively. The uncertainty in $\Delta\Delta G$ increases to about ± 0.3 kcal/mol for the least stable proteins.

The mutants that gave crystals suitable for X-ray analysis were grown under conditions similar to those for wild-type (i.e., ~2 M phosphate solutions, pH ~6.7) (Eriksson et al., 1992, 1993). The locations of these mutations are shown in Figure 1. Other conditions were explored for those mutants that did not crystallize. Data collection and refinement statistics are given in Table 3. Coordinates for these refined structures have been deposited in the Brookhaven Data Bank (ID Codes 235L-251L).

In many mutant structures, even those isomorphous with wild-type, T4 lysozyme can exhibit "hinge-bending" displacements. For this reason, the domain within which the mutant in question is located is always superimposed on the same domain in the wild-type structure prior to determining the magnitudes of any structural changes.

In the following paragraphs a brief description is given of the structural adjustments associated with each variant. The α -helices in T4 lysozyme are identified as follows: A (3–10), B (39–50), C (60–79), D (82–90), E (93–106), F (108–113), G (115–123), H (126–134), I (137–141), and J (143–155) (Nicholson et al., 1991).

Isoleucine 17 to alanine

Isoleucine 17 is located within the three short β -strands of the N-terminal domain. The sidechain is not fully buried, being partly covered by the mobile residues primarily on the surface, Lys 43 and Lys 16, as well as some that are relatively well-ordered such as Leu 39 and Tyr 25. Upon replacement of Ile 17 by alanine the backbone atoms change very little. The sidechains of Lys 43 and Lys 16 may move but this is uncertain because their thermal factors remain high in the mutant structure (81 \AA^2 and 61 \AA^2 ,

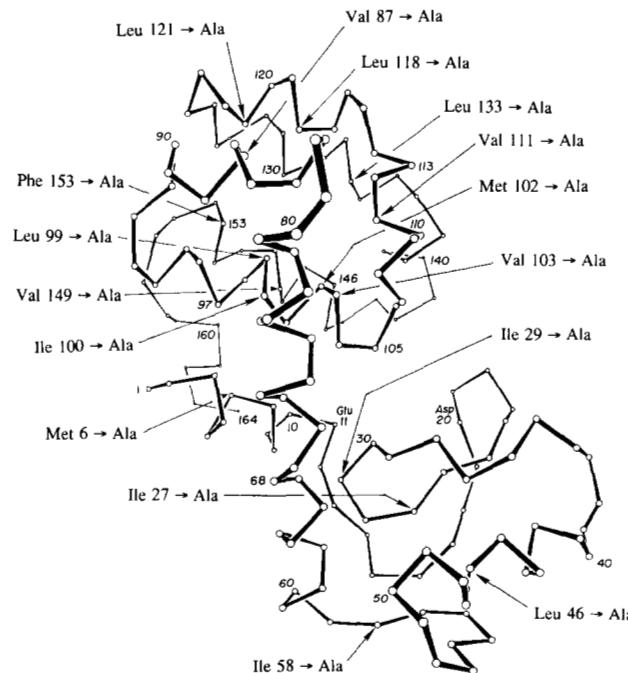


Fig. 1. Schematic drawing of the α -carbon backbone of T4 lysozyme showing the locations of those substitutions confirmed by crystallographic analysis to be internal and/or to form cavities. Note that the figure does not include all the mutants generated (Table 1), nor all those analyzed thermodynamically (Table 2) or crystallographically (Table 3).

Table 3. Crystallographic data collection and refinement statistics^a

Mutant lysozyme	Data collection				Refined model			
	Cell dimensions		Resolution (Å)	Data completeness (%)	R_{merge} (%)	R (%)	Δ_{bond} (Å)	Δ_{angle} (°)
I17A	61.1	96.3	1.8	71	—	16.0	0.012	1.9
I27A	60.8	97.0	1.7	91	3.2	16.0	0.012	1.8
I29A	61.0	96.6	1.75	90	2.1	15.9	0.012	1.8
I50A	60.9	97.5	1.8	87	3.6	16.2	0.015	2.0
I58A	60.8	97.0	1.7	89	2.6	16.3	0.012	1.8
I100A	60.8	97.2	1.7	62	5.4	15.7	0.014	2.0
V87A	60.7	96.7	1.9	90	3.2	16.0	0.013	1.8
V103A	61.0	97.1	1.9	87	4.0	16.4	0.015	2.0
V111A	61.0	96.7	1.9	91	3.2	15.6	0.013	1.8
V149A	60.8	97.0	1.7	87	3.3	16.4	0.017	2.0
M6A	60.9	96.6	1.8	91	3.0	16.2	0.014	1.9
M106A	60.9	96.5	1.9	67	4.7	15.3	0.015	2.0
F67A	60.8	97.0	1.9	89	3.9	15.2	0.015	2.0
L84A	60.9	97.1	1.75	88	2.6	15.8	0.015	1.9
I27A/I29A	61.0	97.0	1.9	93	3.8	15.7	0.015	2.0
I27A/I58A	60.9	96.9	1.9	93	4.3	15.3	0.015	1.9
I29A/I58A	61.0	96.8	1.8	90	4.2	15.9	0.014	1.8
L121A/L133A	61.1	96.4	2.6	93	6.8	15.0	0.014	2.0

^aThe cell dimensions of wild-type lysozyme are $a = b = 61.2 \text{ \AA}$, $c = 96.8 \text{ \AA}$. All crystals have space group P3₂1, the same as that of wild-type lysozyme. R_{merge} gives the average agreement between independently measured diffraction intensities; R is the crystallographic residual for the refined model, and Δ_{bond} and Δ_{angle} are the average discrepancies of the bond lengths and angles from "ideal" stereochemistry.

respectively) (Fig. 2). One of the unique features of this mutant is that an obvious crevice forms at the mutation site, connected to bulk solvent through an opening on the surface. There is no crystallographic evidence suggesting ordered water molecules within the crevice. One wall of the crevice is formed by the highly mobile sidechains of Lys 16 and Lys 43.

Isoleucine 27 to alanine

Isoleucine 27 is located in the middle strand of three short anti-parallel β -strands and is completely inaccessible to solvent. The electron density map (Fig. 3A) and the refined structure of the Ile 27 \rightarrow Ala mutant (Fig. 3B) show that the backbone atoms

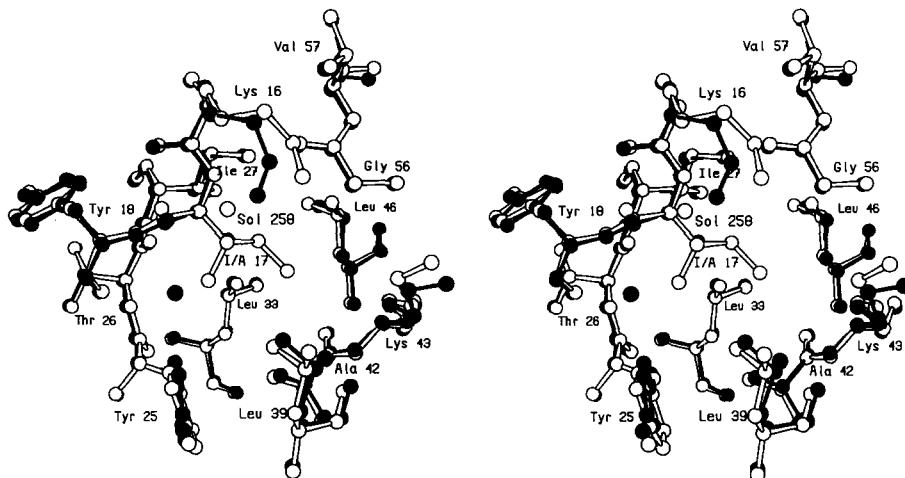


Fig. 2. Superposition of the refined structures of mutant I17A (solid bonds) on WT* lysozyme (open bonds). In order to eliminate possible artefacts due to "hinge-bending," the two sets of coordinates were superimposed based on the mainchain atoms within the domain that included the mutation, here residues 1–60.

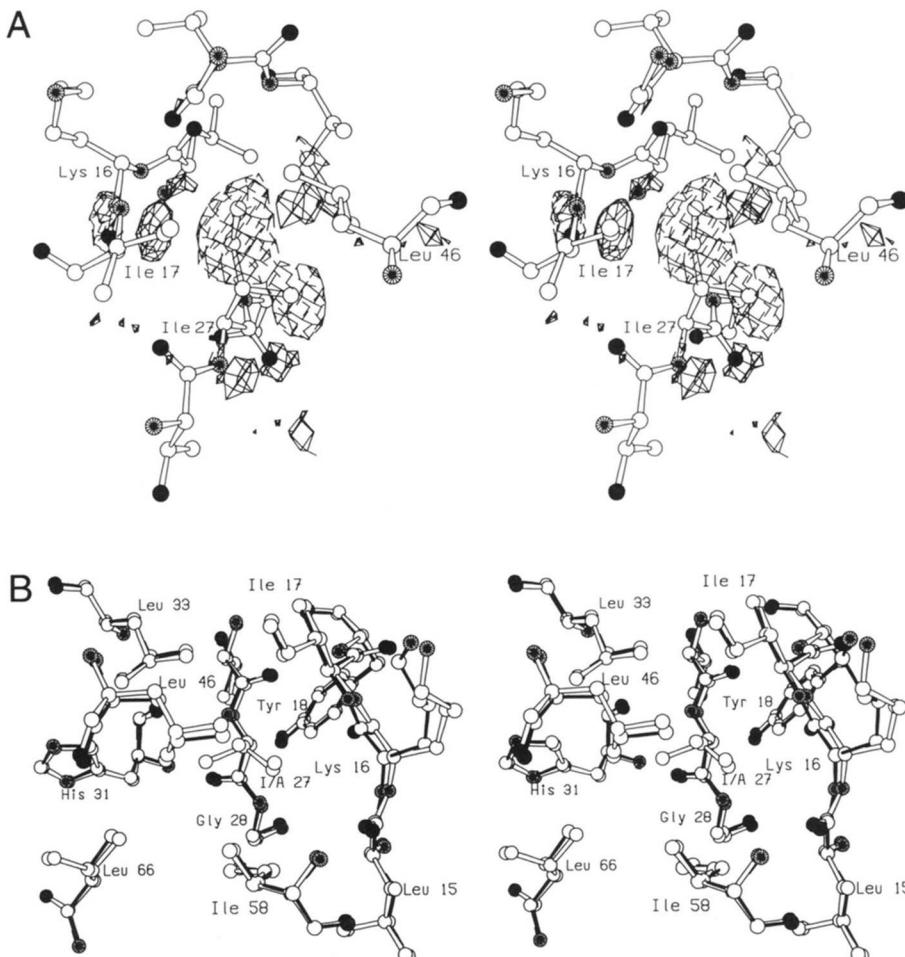


Fig. 3. **A:** Map showing the difference in electron density between mutant I27A and WT* lysozyme. Amplitudes ($F_{\text{mut,obs}} - F_{\text{WT,obs}}$) where $F_{\text{mut,obs}}$ and $F_{\text{WT,obs}}$ are the observed structure amplitudes for the mutant and the wild-type lysozyme, respectively; phases calculated from the refined model structure of the WT* protein. Positive contours (solid) and negative contours (broken) drawn at $+4\sigma$ and -4σ , where σ is the RMS density throughout the unit cell. The map is superimposed on the model for WT* with oxygen atoms drawn solid, nitrogen with "spokes" and carbon as open circles. **B:** Superposition of mutant I27A (solid bonds) on WT* (open bonds) based on residues 1–60.

of residues 14–17 move toward the space vacated by the deleted atoms, the largest shift ($\sim 0.4 \text{ \AA}$) occurring for Lys 16. The backbone and sidechain atoms of Ile 17 shift about 0.16. The Ile 58 sidechain rotates slightly toward the cavity (0.2 to 0.4 \AA). Other backbone movements ($=0.3 \text{ \AA}$) toward the cavity occur at the amino-terminus of helix B (Ile 50) and at the β -turn that includes residues 28–30. In addition, the sidechain of Leu 46 shifts by 0.6 \AA as a result of mainchain atom and dihedral angle adjustments.

Isoleucine 29 to alanine

Ile 29 is within a sharp turn at one end of the irregular anti-parallel β structure of T4 lysozyme. The difference density map (Fig. 4A) shows several strong positive and negative peaks around the mutation site, suggesting that there are significant adjustments in the structure of this mutant. This is also seen in the shift plot (Fig. 5B) and in the larger overall coordinate changes for this variant

(Fig. 4B; Table 4). The most significant changes occur within the loop (residues 11–16) which collapses toward the site of replacement. In particular, the backbone atoms of Leu 13 move 1.0 \AA and the sidechain swings in to largely fill the space vacated by Ile 29. The sidechain torsional angles of Ile 29 change from $\chi_1 = 167^\circ$ and $\chi_2 = 69^\circ$ in the wild-type structure to $\chi_1 = -163^\circ$ and $\chi_2 = 111^\circ$ in the mutant. In addition to the large-scale movement of Leu 13, Phe 67 tilts with one side of the ring shifting 0.5 \AA closer to Leu 13. The backbone atoms of residues 28 and 29 also move about 0.5 \AA so that, in total, the potential cavity that might be made by this mutation is, in fact, completely eliminated (Table 5). As will be apparent from Figure 5B, this mutation causes a significant "hinge-bending" displacement. It appears that the long interdomain helix that connects the two domains of the molecule (Fig. 1) acts as a lever. The movement of Phe 67, which is within the helix, toward the site vacated by Ile 29, causes the lever-arm to move in such a way that the active site cleft is slightly opened.

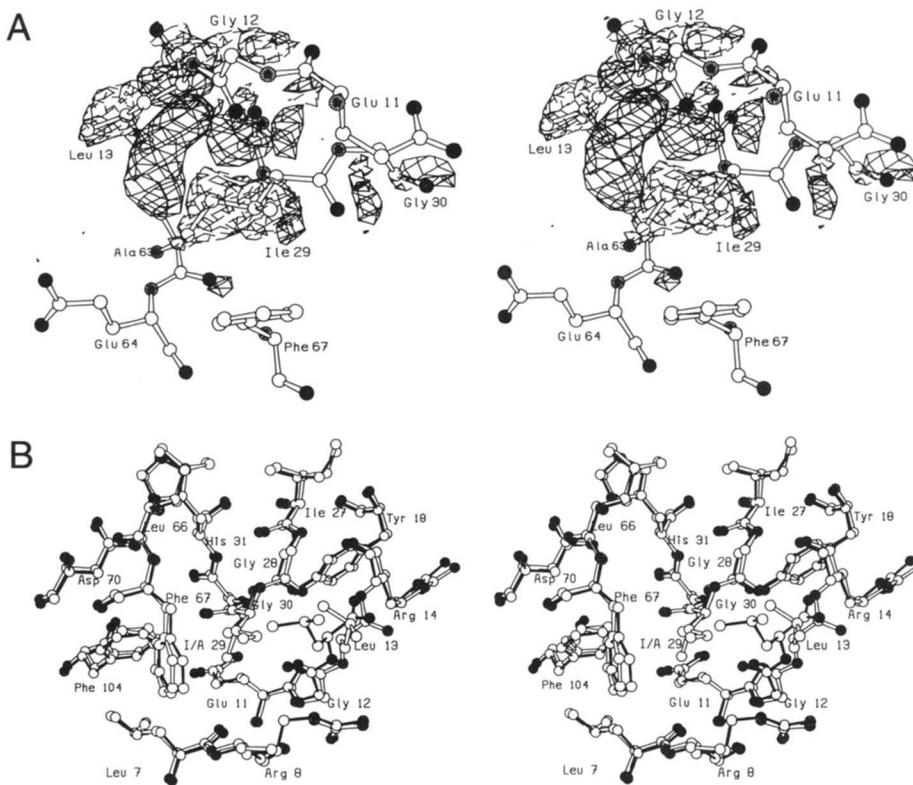


Fig. 4. **A:** Difference density map for mutant I29A. All procedures and conventions as in Figure 3A. **B:** Superposition of mutant I29A (solid bonds) on WT* (open bonds) based on residues 1–60.

Isoleucine 50 to alanine

Ile 50 is partially buried and substitution of alanine leaves a crevice on the surface without significant structural adjustments except that the surrounding sidechains move slightly toward the space vacated by the Ile 50 sidechain (Fig. 6). Unlike Ile 17 → Ala, the crevice formed in the Ile 50 → Ala mutant is more shallow with a wider opening to the surface.

Isoleucine 58 to alanine

Isoleucine 58 is located within the short loop connecting helices B and C in the amino-terminal domain and is fully buried. The difference map for the Ile 58 → Ala mutant (Fig. 7A) shows the loss of the ethyl and methyl groups, as well as associated structural adjustments, the largest of which (1.0 Å) is in the sidechain of Ile 27 which rotates 80° relative to its wild-type conformation (Fig. 7B). $C^{\gamma 1}$ of Ile 27 moves by 0.9 Å toward the space vacated by the isoleucyl sidechain and $C^{\gamma 2}$ moves 0.7 Å. In doing so, $C^{\delta 1}$ of Ile 27 moves 1.0 Å away to avoid a stearic clash with Ala 58. In addition, the β -carbon of the newly-introduced Ala 58 moves 0.6 Å toward the newly-created cavity.

Isoleucine 100 to alanine

Ile 100 is in the middle of helix E, the “central helix” (Fig. 1), but is slightly (6%) exposed to solvent. In the crystal, however, the “exposed” region is in contact with a hydroxyethyl disulfide molecule, “HED 170” (i.e., the oxidized form of β -mercaptoethanol)

which binds between two lysozyme molecules at a crystal contact (Bell et al., 1991). Since helix E is largely buried within the carboxy-terminal domain, it is expected to be held relatively rigidly and not especially susceptible to structural change. Indeed, in the difference map (Fig. 8A), the most obvious feature other than the density change due to the loss of Ile 100, corresponds to a shift in one of the sulfur atoms of HED 170. The refined mutant structure (Fig. 8B) does show that one turn of helix C, comprising residues from Ala 74 to Ile 78, rotates essentially about the axis of the helix such that C^{α} of Val 75 moves 0.4 Å and C^{α} of Arg 76 moves 0.6 Å. In association with this mainchain rotation, the sidechain of Val 75 shifts 0.9 Å, bringing it closer to the cavity. Sidechain atoms of Tyr 88 move about 0.5 Å toward the cavity while the sidechains of Leu 99 and Ala 100 rotate 0.25 Å away.

Valine 87 to alanine

Valine 87 is located in the middle of helix D. Together with Leu 118, Leu 121, and Leu 99, it constitutes part of the hydrophobic core within the carboxy-terminal domain. The difference map (Fig. 9A) shows several density features associated with helix D in addition to the truncation of the valyl sidechain, suggesting that some adjustments have been made by this helix to fill the vacated space. This is confirmed from the refined model structure (Fig. 9B) in which helix D shifts its backbone atoms 0.35 Å, on average. The largest shift for sidechain atoms occurs at residue Gln 122 where the sidechain atoms rotate up to 1.9 Å to partially fill the putative cavity. The neighboring hydrophobic residues Leu 118, Leu 91, Leu 99, and Leu 121 are practically unchanged.

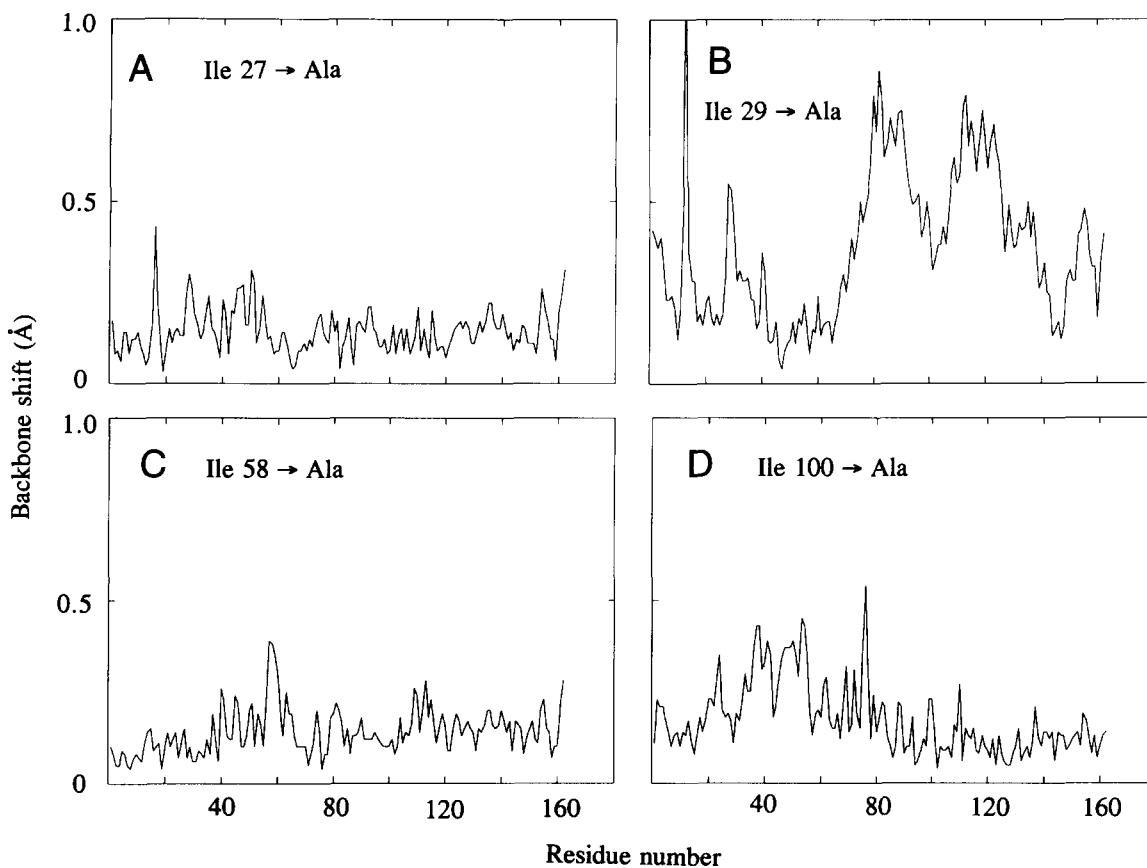


Fig. 5. Shift plots showing the difference between the backbones of representative cavity-creating mutant lysozyme and WT*. Each mutant structure was superimposed on the wild-type so as to minimize the RMS discrepancy between the respective backbone atoms in the domain where the mutation occurs, i.e., the carboxy-terminal domain for mutants at positions 87, 100, 111, and 149, and amino-terminal domain for those at positions 27, 29, and 58. The value plotted for each residue is the average discrepancy for the C^α, N, C, and O atoms and the arrowheads show the location of the mutation. **A:** I27A. **B:** I29A. **C:** I58A. **D:** I100A.

Apart from the above shifts, the overall mutant structure is very similar to wild-type.

Valine 103 to alanine

The difference Fourier map (Fig. 10A) shows density corresponding to the truncated valyl sidechain. Additional features suggest that the backbone segment between Ala 74 and Val 75 shifts toward the core. In the refined model of the mutant, it is seen that the Val 111 sidechain moves 0.6 Å toward the space vacated by Val 103 (Fig. 10B). In concert, the sidechain of Met 106 changes its rotamer angles to maintain optimal contacts with Val 111. These adjustments cause the sulfur atom to move about 1.1 Å. In wild-type lysozyme, the torsional angles χ_1 , χ_2 , and χ_3 of Met 106 are -79° , -164° , and 113° , respectively. In the mutant they change to 66° , -176° , and 74° . The remainder of the mutant structure is, in general, very similar to wild-type.

Valine 111 to alanine

Helix F, which includes Val 111, is short and relatively mobile. As shown in Figure 11A, there are several strong difference-density features associated with the mainchain atoms within this helix as

well as with the sidechains of Met 120 and Met 106. The refined model (Fig. 11B) suggests that, in association with removal of the two methyl groups, some mainchain atoms within helix F shift up to 0.5 Å toward the putative cavity. The sulfur atom within the sidechain Met 102 shifts 1.1 Å and largely fills in the space vacated by Val 111. In order to accommodate this motion, C^ε of Met 102 rotates away to avoid a potential steric clash. In total, the adjustments substantially reduce the volume of the cavity that would otherwise have been generated.

Valine 149 to alanine

Val 149 is within helix 143–155 in the carboxy-terminal domain. One of the methyl groups of the valine is in a hydrophobic environment that includes the sidechains of Phe 153, Met 102, Trp 138, and Ala 98 (Fig. 12A). The other methyl group is in a more polar environment that includes Asp 10, Arg 145, Asn 101, the hydroxyl of Tyr 161 and an internal solvent molecule (Fig. 1G). The initial difference density map (Fig. 12A) showed negative density for the deleted valyl sidechain at position 149 and an adjacent positive peak that corresponds to an additional solvent molecule bound to the mutant structure (see below). Otherwise, there are virtually no other difference features indicating that the mutant structure is essentially unchanged from the wild-type protein. This is con-

Table 4. Cavities in mutant T4 lysozymes^a

Protein	Cavity	Center of cavity			Observed cavity volume V_c (\AA^3)	Increase in cavity volume [$V_c(\text{mut}) - V_c(\text{WT}^*)$] (\AA^3)	Model cavity volume V_{model} (\AA^3)
		X (\AA)	Y (\AA)	Z (\AA)			
WT*	I	28.0	8.9	0.6	34	—	—
	II	32.1	4.0	-4.9	<u>20</u> 54	—	—
I27A		43.9	20.0	23.5	17	17	22
I29A		—	—	—	0	0	0 ^c
I58A		41.3	15.5	22.9	18	18	50
I100A	I	28.0	8.8	0.3	33	28	23
	II	32.1	3.8	-4.9	<u>20</u> 29	—	—
		33.4	6.6	12.7	<u>29</u> 82	—	—
V87A	I	28.0	8.6	0.6	39	31	23
	II	32.2	4.1	-4.9	<u>23</u> 13	—	—
		25.4	1.6	0.2	<u>10</u> 85	—	—
		27.0	4.4	1.8	<u>10</u> 85	—	—
V103A	I	28.3	8.6	0.6	41	23	12
	II	32.1	4.0	-4.9	<u>19</u> 17	—	—
		27.4	8.1	7.0	<u>17</u> 77	—	—
V111A	I	27.8	9.1	1.2	38	11	15
	II	32.3	4.1	-4.9	<u>27</u> 65	—	—
V149A (without water)	I	28.1	8.9	0.6	35 ^b	17	0
	II	32.2	4.0	-4.9	<u>21</u> 14	—	—
		38.6	9.9	4.1	<u>14</u> 71	—	—
V149A (with water)	I	28.1	8.9	0.6	35 ^b	3	—
	II	32.2	4.0	-4.9	<u>21</u> 57	—	—
M6A (without water)	I	27.9	8.7	0.5	39 ^b	68	71
	II	32.2	3.9	-4.8	<u>25</u> 58	—	—
		37.7	4.5	5.5	<u>58</u> 122	—	—
M6A (with water)	I	27.9	8.7	0.5	39 ^b	10	—
	II	32.2	3.9	-4.8	<u>25</u> 64	—	—
M102A	I	29.5	9.5	1.0	83	65	109
	II	32.2	4.5	-4.7	<u>35</u> 119	—	—
I27A/I29A		43.8	19.8	23.7	21	21	22 ^c
I27A/I58A		41.4	16.2	22.9	27	27	95
I29A/I58A		41.4	15.6	23.2	12	12	50 ^c
L121A/L133A	I & II	27.2	4.1	0.5	8	126	234
		30.5	6.5	-2.6	<u>172</u> 180	—	—

^aCavity volumes (V_c) were calculated as in Materials and methods. There are four internal solvent molecules which are observed in the structure of WT* lysozyme and are retained in each of the mutants. These are considered to be an integral part of the lysozyme structure and were left in place during all the cavity calculations. Except for the special consideration of V149A and M6A (see below) all other solvent molecules were deleted. The "increase in cavity volume" refers to the increase in the total volume of cavities in the mutant structure compared with that of the wild-type structure. This calculation is carried out for the domain within which the mutation occurs. In wild-type lysozyme there are two cavities in the carboxy-terminal domain, identified in the table as I and II, with a total volume of 54 \AA^3 . No cavities are detected in the amino-terminal domain. ΔV_{model} is the volume of the putative cavity that would be generated by truncating the targeted sidechain to an alanine without changing the rest of the structure from that of wild-type lysozyme.

^bMutants V149A and M6A bind, respectively, one and two solvent molecules at or close to the site of the substitution. The volume of the resultant cavity was calculated both without and with these solvent molecules present. With water present, no cavity was apparent at the site of the mutation in either case.

^cModel cavity for Ile 29 connects to the surface.

Table 5. Stability and cavity volume data for alanine substitutions in barnase and ribonuclease-S^a

Mutant	$\Delta\Delta G$ (kcal/mol)	Cavity volume (Å ³) (Buckle et al., 1996; Varadarajan & Richards, 1992)	Cavity volume (Å ³) (Connolly, 1983)
Barnase Leu 14 → Ala	-4.3	89	24
Barnase Ile 76 → Ala	-3.2	31	0
Barnase Ile 88 → Ala	-4.0	91	21
Barnase Ile 96 → Ala	-1.9	83	44
R-Nase-S Met 13 → Ala	-4.3	24	—

^aThe table summarizes stability and cavity volume data for Ile → Ala and Leu → Ala mutations in barnase (Buckle et al., 1996) and a Met → Ala replacement in ribonuclease-S (Varadarajan & Richards, 1992). The cavity volume calculations of Buckle et al. (1996) used the program VOIDOO (Kleywegt & Jones, 1994) with a probe radius of 0.8 Å; those of Varadarajan and Richards (1992) used a Voronoi analysis. The alternative volumes given in the last column were obtained using the program of Connolly (1983) with a probe radius of 1.2 Å. For barnase, each volume is the average for the three molecules in the crystallographic asymmetric unit.

firmed by superimposing the mutant structure on wild-type (Fig. 12B). The largest sidechain motion is only 0.3 Å for O^{δ1} of Asn 101. Otherwise, the sidechain displacement in the vicinity is 0.2–0.25 Å. The refinement confirms that the positive density feature seen in Figure 12A corresponds to an ordered water molecule which makes hydrogen bonds with O^{δ2} of Asp 10 (2.3 Å), N^ε of Arg 145 (2.9 Å), N^{δ2} and O^{δ1} of Asn 101 (3.0 Å and 3.1 Å) and O^γ of Tyr 161. These polar sidechains are fully buried and are involved in an extensive hydrogen-bonded network.

Methionine 6 to alanine

Met 6 is located in the “waist” of T4 lysozyme (Fig. 1) but can be considered as part of the carboxy-terminal domain. The replacement of Met 6 with isoleucine was previously shown to cause the protein to crystallize in a different space group as well as that of

the native crystal (Faber & Matthews, 1990). In the present study Met 6 → Ala crystallized isomorphously with the wild-type protein. Prominent negative difference electron density peaks confirm the truncation of the methionine sidechain in the mutant (data not shown). Two distinct positive peaks near the sites where S^δ and C^ε atoms of Met 6 were located in the WT* protein suggest the binding of two new ordered water molecules (Fig. 13). Refinement shows these solvent molecules to have crystallographic thermal factors of 27 Å and 28 Å, respectively, indicating that they are well localized with high occupancy. The first water molecule is hydrogen bonded to O^γ of Tyr 161 (2.9 Å), O^γ of Thr 152 (3.3 Å) and the second water molecule (2.6 Å) which in turn is hydrogen bonded to N^{δ2} of Asn 101 (3.1 Å) and the carbonyl oxygen of Ala 97 (3.2 Å). In the mutant, O^γ of Tyr 161 moves 0.3 Å and Trp 158 moves up to 0.9 Å toward the space vacated by Met 6. Otherwise, the mutant structure is very similar to wild-type.

The observation that solvent binds within the cavity created by the Met 6 → Ala substitution, as well as in the case of the Val 149 → Ala substitution discussed above, shows that the regions enclosing these sidechains are somewhat non-polar (see Matsumura et al., 1988, 1989; Varadarajan & Richards, 1992; Otzen et al., 1995; Takano et al., 1995).

Phenylalanine 67 to alanine

Another case where a crevice results from a large-to-small substitution is the mutant Phe 67 → Ala. Phe 67 is at the interface between the two domains and is within the interdomain helix. Removal of the phenyl moiety causes some small adjustments of the surrounding sidechains towards the center of the mutation site (Fig. 14). An ordered water molecule (Sol 327 in Fig. 14) is observed near the position previously occupied by C^{ε1} of Phe 67. This solvent molecule hydrogen bonds to the backbone carbonyl oxygen of Phe 4 as well as to another solvent (Sol 221 in Fig. 14) which is present in the WT protein structure.

Phe 67 has been identified as one of the sidechains that changes conformation in association with interdomain, hinge-bending, motion (Zhang et al., 1995). In the present context, however, the Phe 67 → Ala substitution did not cause a significant hinge-bending displacement.

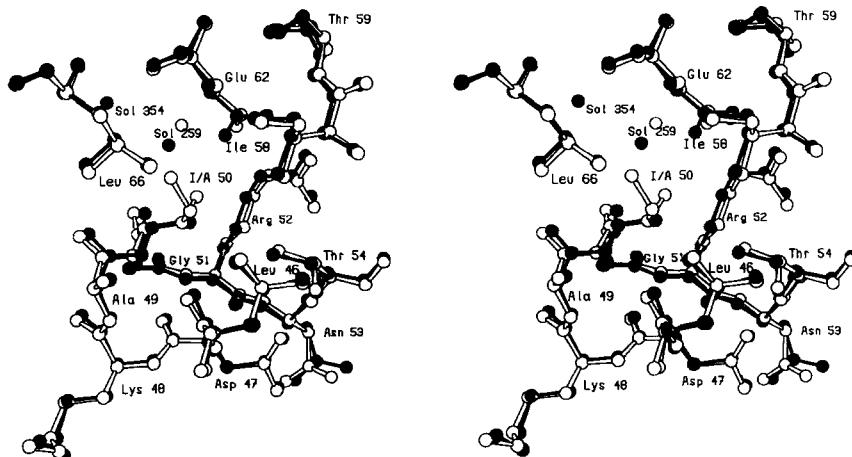


Fig. 6. Superposition of mutant I50A (solid bonds) on WT* (open bonds) based on residues 1–60.

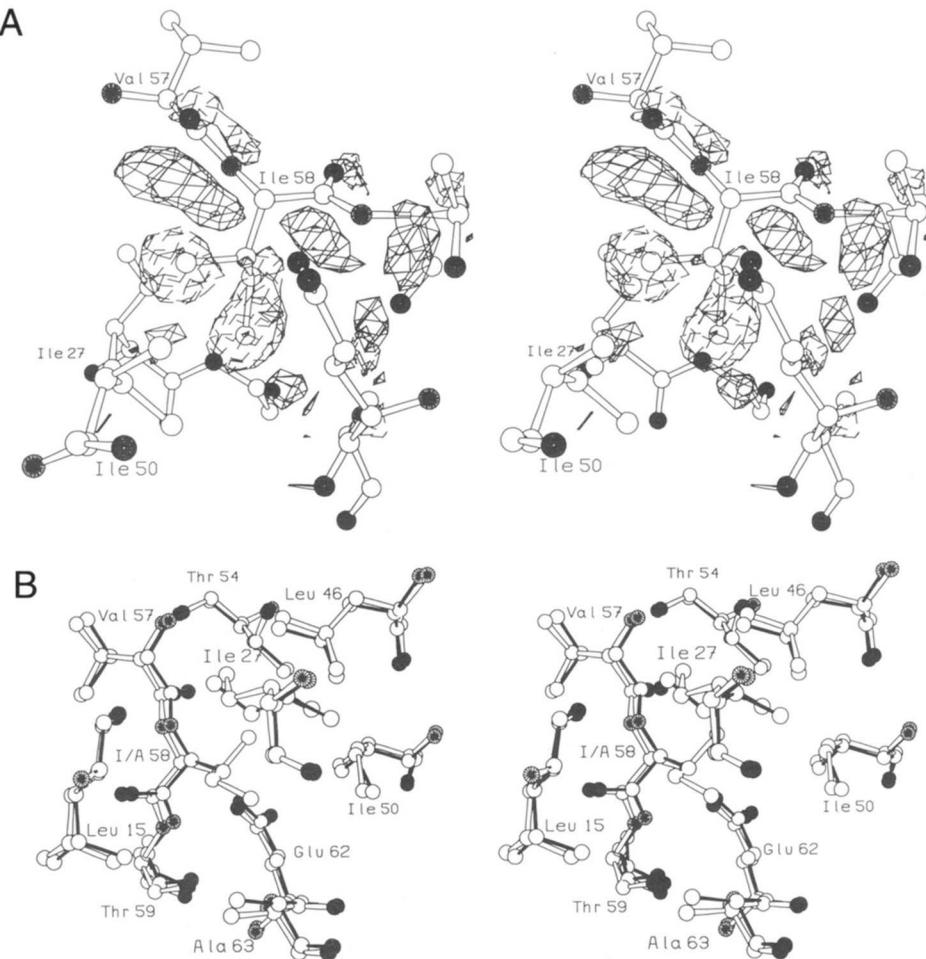


Fig. 7. **A:** Difference density map for mutant I58A. All procedures and conventions as in Figure 3A. **B:** Superposition of mutant I58A (solid bonds) on WT* (open bonds) based on residues 1–60.

Leucine 84 to alanine

The Leu 84 → Ala mutant differs from any of the other large-to-small substitutions in that it generates neither an internal cavity nor a crevice but rather a channel with both ends open to the protein surface. Leu 84 is fully excluded from the bulk solvent having its principal contacts with the “floppy helix,” helix F, as well as the sidechains of Glu 108, Val 111, and Leu 118. Upon truncation of leucine sidechain, there are large structural rearrangements within helix F. These are the largest changes among all the single substitutions discussed herein. Within helix F the electron density of the backbone is disconnected in several places and there is no apparent density for several of the sidechains, indicating that this helix becomes even less well ordered than it is in the wild-type structure. The refined model suggests that backbone atoms within the helix move up to 2 Å (Fig. 15). The end of helix G tilts in about 0.7 Å toward mutation site. Helix D, however, which includes Leu 84, essentially remains in place. It was previously suggested that the interaction between Val 111 and Leu 84 is important in localizing helix F (Morton & Matthews, 1995). The observation that the helix is substantially disordered on substituting Leu 84 with alanine supports this assertion. On the other hand, the mutation Val 111 → Ala did not substantially increase the mobility of the helix, suggesting that the role of Leu 84 is more important.

Isoleucine 27 to alanine plus isoleucine 29 to alanine

Ile 27 and Ile 29 are only two residues apart but elicit very different responses when replaced by alanine. The structural adjustments in the double mutant (not shown) are essentially additive, as might be expected in that Ile 27 and Ile 29 interact with different structural elements within the amino-terminal core. As also seen for the single mutant Ile 29 → Ala, the largest conformational changes in the double mutant occur in the loop region, residues 12 to 16, where the C^a atom of Leu 13 shifts about 0.6 Å and the sidechain rotates to largely occupy the space vacated by Ile 29. The backbone atoms within this loop region in the double mutant are virtually identical with the single mutant. The three-residue turn involving residues 28 to 30 adopts a position that is essentially average between that seen in the Ile 27 → Ala and Ile 29 → Ala structures, with up to 0.4 Å shift in mainchain atoms when compared with the WT* structure.

Isoleucine 27 to alanine plus isoleucine 58 to alanine

When both Ile 27 and Ile 58 are simultaneously replaced by alanine, the resultant structure (not shown) is very similar to the combination of two mutant structures with the single alanine replacements. As with I58A, the largest backbone movement occurs

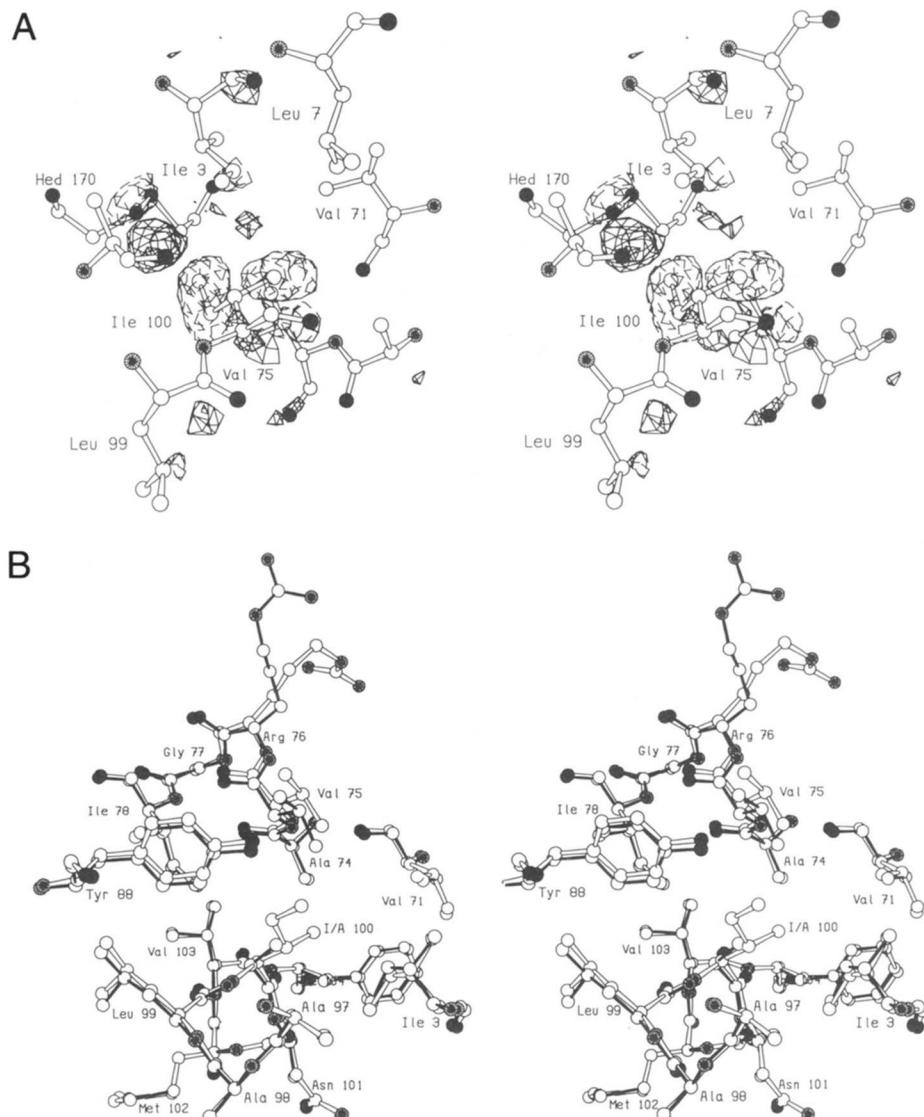


Fig. 8. **A:** Difference density map for mutant I100A. All procedures and conventions as in Figure 3A. “Hed 170” is a molecule of hydroxymethyl disulfide (see text) with the two sulfurs drawn solid. **B:** Superposition of mutant I100A (solid bonds) on WT* (open bonds) based on residues 80–162.

at position 58, where the C^α atom moves 0.6 Å toward the space vacated by Ile 58. The amino-terminus of helix B does, however, move toward the cavity to a greater degree (total shift 0.4 Å) than is the case with either of the single mutants (~0.2 Å). Structural comparison of the double mutant with two single mutants suggests that although Ile 27 and Ile 58 have sidechain atoms within 3.9 Å, they interact weakly explaining the additivity of two single replacements in the double mutant structure. As a result, the cavities created by the single replacements connect in the double mutant to form one large cavity, the size being close to the sum of the cavities in the two single mutants.

Isoleucine 29 to alanine plus isoleucine 58 to alanine

This double mutant (not shown) provides another example of the additivity of the structural adjustments associated with the constit-

uent single replacements. Since I29A displayed larger structural changes whereas I58A showed very little change in structure relative to WT*, the overall structure of the double mutant is essentially identical to I29A, with a root-mean-square (RMS) deviation of 0.11 Å for the mainchain atoms. When the double mutant is compared with the single mutant, Ile 58 → Ala, and with WT*, the RMS deviations are 0.28 Å and 0.31 Å, respectively. The global hinge-bending motion seen in I29A is also seen in the double mutant.

Leucine 121 to alanine plus leucine 133 to alanine

The structures with the single replacements, Leu 121 → Ala and Leu 133 → Ala, have been described (Eriksson et al., 1992; Baldwin et al., 1996). The structural adjustments in response to the simultaneous alanine replacements (Fig. 16) are analogous to the

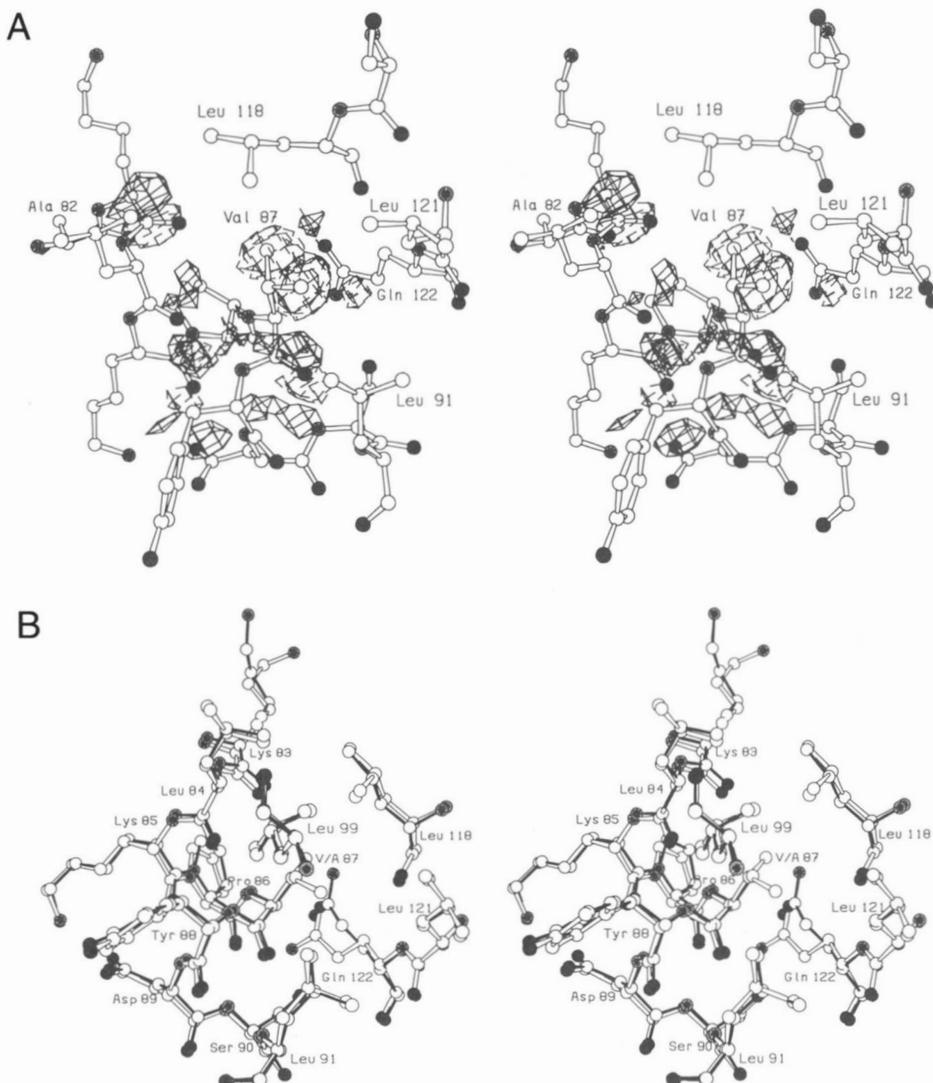


Fig. 9. **A:** Difference density map for mutant V87A. All procedures and conventions as in Figure 3A. **B:** Superposition of mutant V87A (solid bonds) on WT* (open bonds) based on residues 80–162.

sum of the effects of the constituent single replacements. Helix H rotates away slightly, especially the two consecutive alanines Ala 129 and Ala 130, with their C^α atoms moving about 0.6 Å. The whole of helix G shifts ~0.6 Å outwards and toward helix H. As was seen in the L121A mutant structures, the C-terminus of helix J collapses by ~0.8 Å toward the space vacated by Leu 121. The double replacement of the two leucines causes two existing cavities in the carboxy-terminal domain to connect and form a single, larger, hydrophobic cavity which can be described as a flat disk, roughly 9 Å in diameter, and 172 Å³ in volume. In both the initial “F_o – F_c” map as well as in the final “F_o – F_c” map, small positive difference electron density peaks above 4σ were observed within this cavity. The most likely candidates to occupy the cavity are the oxidized and/or reduced forms of β-mercaptopethanol (i.e., HEDS or BME), since both are present in the crystallizing medium. A HEDS molecule is, however, rather large to fit within the cavity with acceptable geometry. BME was therefore modeled to fit the electron density map, although there is no apparent hydrogen bond-

ing partner for hydroxyl group. After refinement, the thermal factors of the atoms associated with this newly modeled BME were all close to 100 Å, indicating that it is highly mobile and/or binds with less than 100% occupancy. It might also be noted that the resolution of the diffraction data for this mutant (2.6 Å; Table 3) is much poorer than all the rest (1.9–1.7 Å), so that the structure is less reliably determined.

Discussion

Changes in conformation

In common with the leucine-to-alanine substitutions in T4 lysozyme studied previously (Eriksson et al., 1992), some of the present “large-to-small” substitutions within the core are associated with larger conformational changes than others. Nevertheless, in terms of the overall protein structure the changes are relatively modest and are localized to the immediate vicinity of the substitution. The

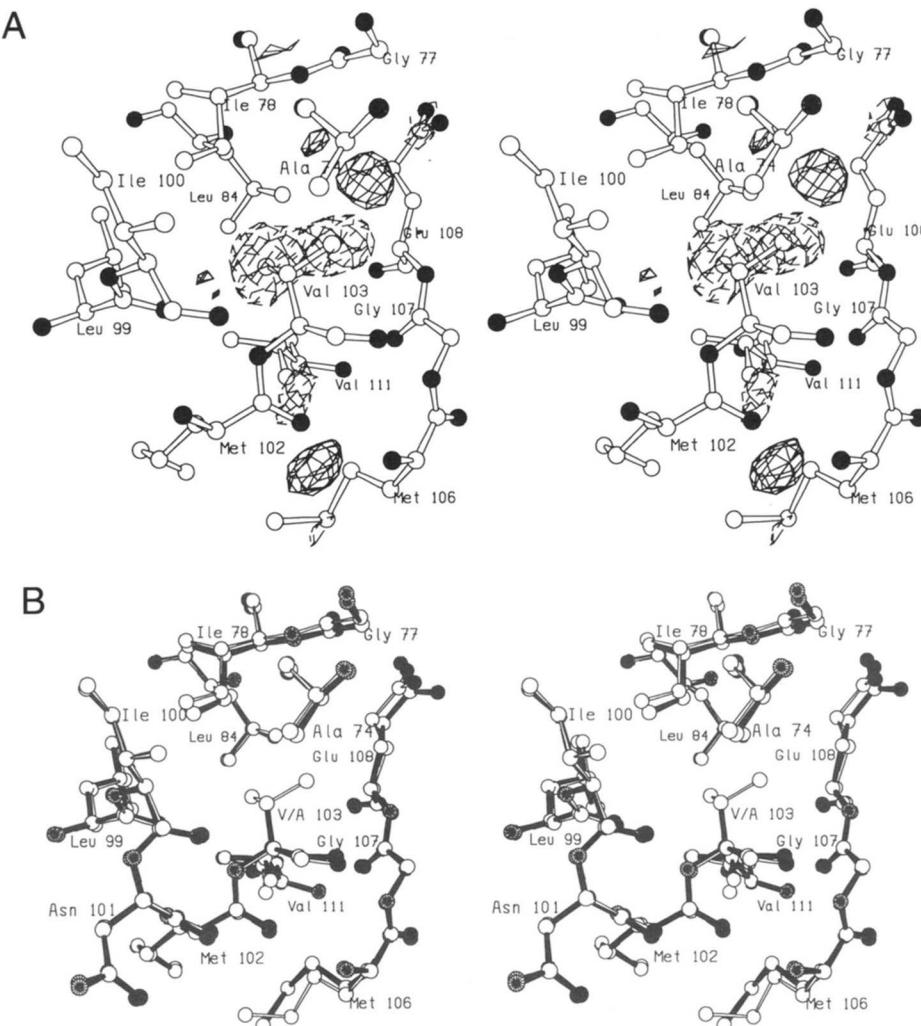


Fig. 10. **A:** Difference density map for mutant V103A. All procedures and conventions as in Figure 3A. **B:** Superposition of mutant V103A (solid bonds) on WT* (open bonds) based on residues 80–162.

RMS change in backbone atoms within the domain that includes the site of mutation is 0.14–0.30 Å. In the other (i.e., mutant-free) domain the discrepancy is 0.09–0.11 Å which can be taken as a measure of the accuracy of the coordinates.

Of the single variants that are entirely internal, mutant I29A has the largest structural changes, both within the domain that includes the mutation (RMS shift in backbone atoms of 0.30 Å) and in terms of overall hinge-bending displacement (Fig. 5B). In this case the mutant structure collapses to such a degree that no cavity is detected. Mutant L84A, which leads to an internal channel, has the largest changes of all the single mutants. The amount of movement that occurs in a given instance seems to depend on the rigidity of the amino acid being substituted, as well as the rigidity of the atoms that surround it. Ile 100, for example, is in the center of helix E which is, in turn, surrounded by four other helices. The sidechain of Ile 100 is well buried and the atoms surrounding it are, themselves, rigidly held and well inside the protein surface. Thus, when Ile 100 is replaced by alanine, the surrounding atoms retain multiple interactions with other neighboring atoms, and move very little (Fig. 5D). A similar situation occurred with L99A (Eriksson

et al., 1992). In contrast, Leu 84 is very close to the surface of the protein and, in addition, is contacted by Val 111 within helix F, which is one of the most mobile helices of the lysozyme structure. Thus, when Leu 84 is truncated to alanine, one of the major interactions stabilizing the position of helix F is removed, and the helix moves toward the position vacated by the leucine sidechain (Fig. 15).

The energetics of cavity formation

The calculation of cavity volume and changes in volume within proteins is decidedly non-trivial (Richards, 1977; Eriksson et al., 1992; Varadarajan & Richards, 1992). With this caveat we show in Figures 17A and B the destabilization of each mutant structure plotted as a function of the increase in cavity volume that is associated with the mutation. The prior data on Leu → Ala substitutions plus one Phe → Ala replacement are also included.

The best-fit line through the data for the Leu → Ala replacements, assuming equal weight for all observations, has an intercept of −2.1 kcal/mol and a slope of 22 cal mol^{−1} Å^{−3}. This differs

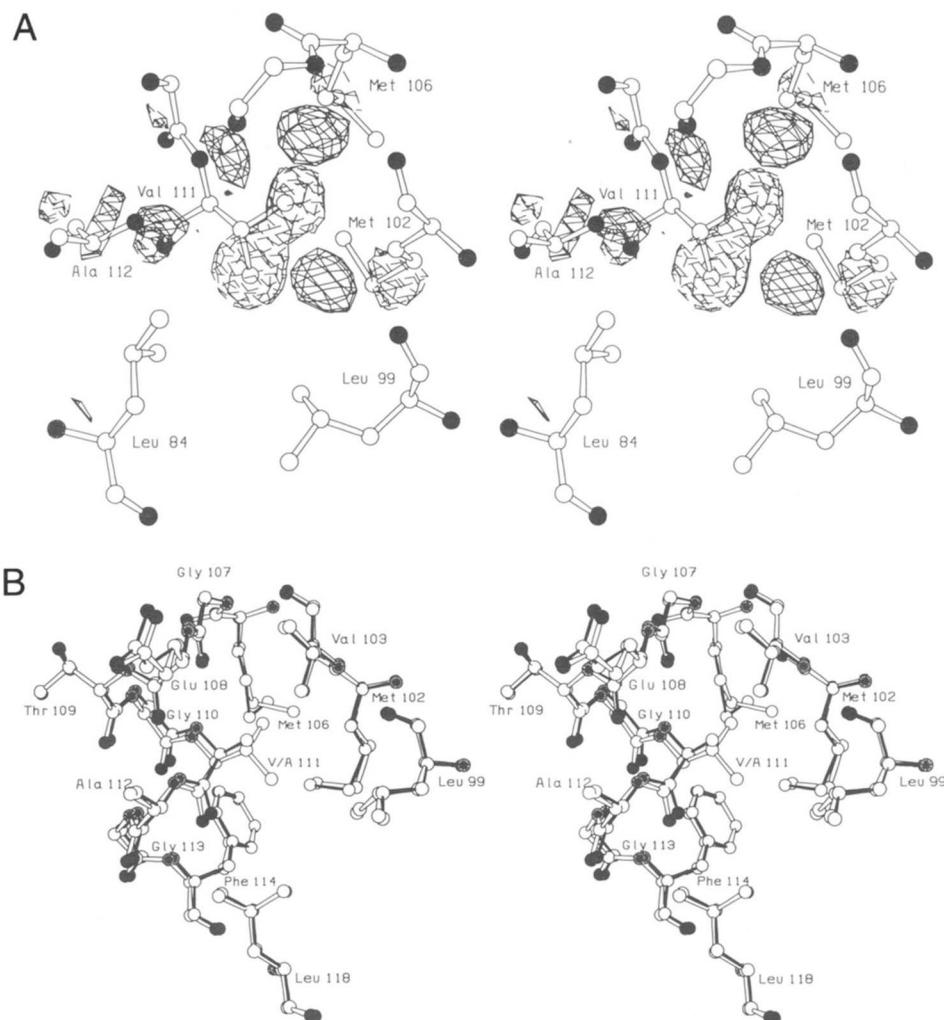


Fig. 11. **A:** Difference density map for mutant V111A. All procedures and conventions as in Figure 3A. **B:** Superposition of mutant V111A (solid bonds) on WT* (open bonds) based on residues 80–162

slightly from the previously reported values (-1.9 kcal/mol , $24 \text{ cal mol}^{-1} \text{ \AA}^{-3}$; Eriksson et al., 1992) in part because the structure of L121A has been redetermined, leading to a decrease from 43 \AA^3 to 21 \AA^3 in cavity volume (Baldwin et al., 1996) and also because the Phe 153 \rightarrow Ala mutant was previously considered together with the Leu \rightarrow Ala substitutions (Eriksson et al., 1992). With the exception of the one variant Val 149 \rightarrow Ala, the observations for all of the other mutants can also be approximately fitted by a series of straight lines of slope $22 \text{ cal mol}^{-1} \text{ \AA}^{-3}$ (Fig. 17A, B). In Figure 17C the data are brought to a common reference scale by subtracting the respective intercept free energies, $\Delta\Delta G_0$ (Table 6). The apparent outlier, Val 149 \rightarrow Ala, is one of the two variants that bind a new internal solvent molecule. Its anomalous behavior has been probed by making further substitutions at site 149 and will be discussed elsewhere. It should be noted, however, that M6A, the other mutant that binds solvent, has “normal” energetics (Fig. 17B).

It thus appears that the present, extended, set of “large-to-small” substitutions within the core of T4 lysozyme substantiate the analysis of Eriksson et al. (1992) based primarily on Leu \rightarrow Ala re-

placements. According to that analysis, a non-polar residue such as leucine within the core of a protein contributes to stability in two ways. First there is a term which depends only on the nature of the sidechain. This is mainly the well-known “hydrophobic effect” (Bernal, 1939; Kauzmann, 1959). Second, there is a term which derives from the van der Waals interactions between the leucine sidechain and the atoms that surround it in the native structure. If the leucine is replaced by an alanine and the wild-type structure is essentially maintained in the mutant, a cavity is formed and the van der Waals interactions are lost. This results in an additional loss of stability above and beyond that due to the reduction in hydrophobicity. To the extent that the protein structure collapses to avoid the formation of a cavity, additional van der Waals interactions are generated which restore some of the otherwise lost stabilization. In the limiting case in which the protein structure collapses to avoid, entirely, the formation of any cavity, the energy loss due to a large-to-small replacement is expected to correspond to the hydrophobic term alone. This is the situation for the mutant Ile 29 \rightarrow Ala (Fig. 17A). In this case the increase in cavity volume is essentially zero and the destabilization is the least of the Ile \rightarrow Ala replacements.

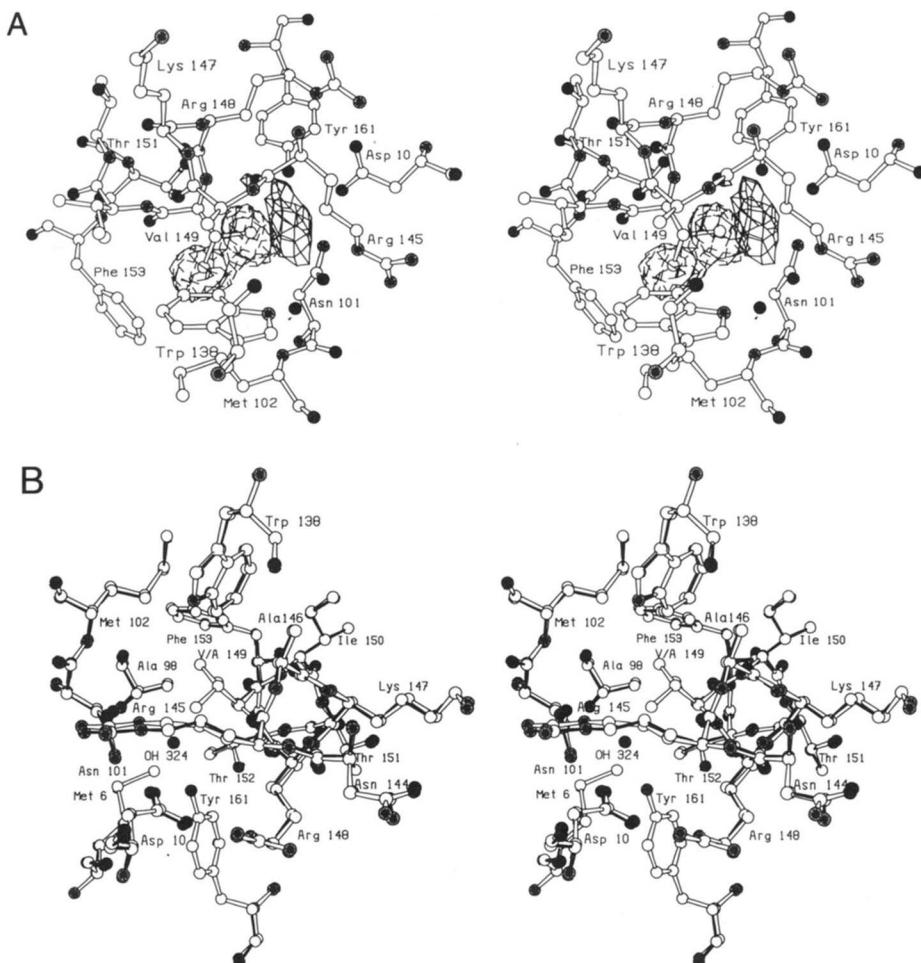


Fig. 12. **A:** Difference density map for mutant V149A. All procedures and conventions as in Figure 3A. **B:** Superposition of mutant V149A (solid bonds) on WT* (open bonds) based on residues 80–162. The ordered solvent molecule bound within the space vacated by the Val 149 sidechain is labeled OH324.

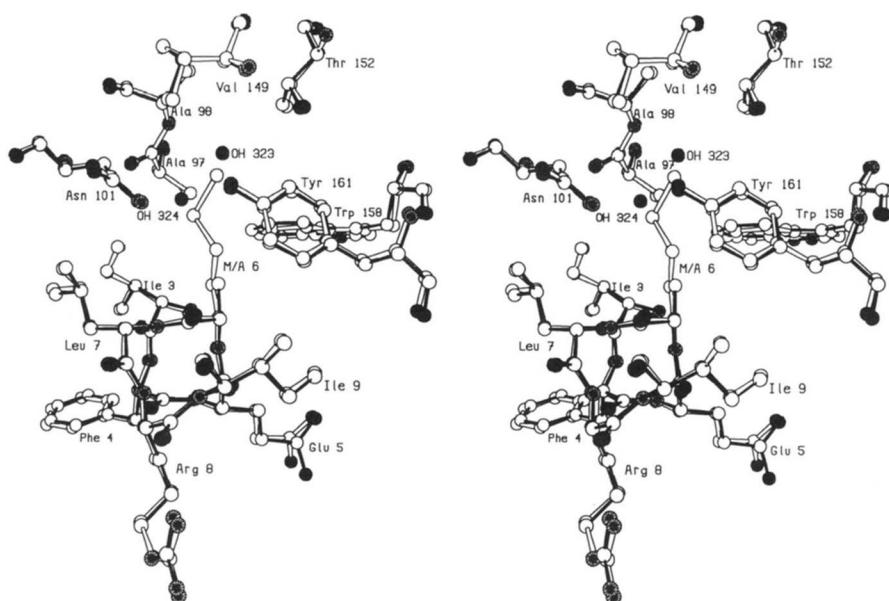


Fig. 13. Superposition of mutant M6A (solid bonds) on WT* (open bonds) based on residues 13–60. The two solvent molecules bound within the space vacated by the methionine sidechain are labeled OH323 and OH324.

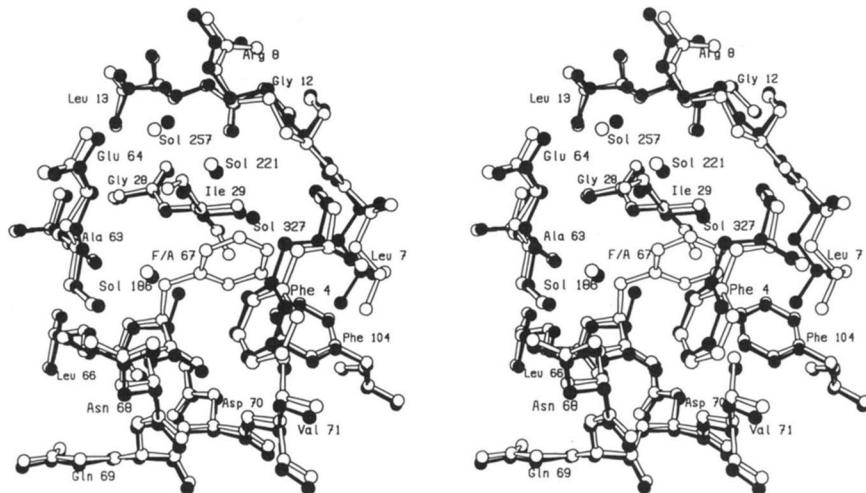


Fig. 14. Superposition of mutant F67A (solid bonds) on WT* (open bonds) based on residues 60–162.

According to the above rationalization, the extrapolated, intercept, values in the energy versus cavity plots (Fig. 17A and B) should provide a direct way to estimate the free energy of transfer of a given sidechain from solvent to core within the context of a folded protein. For Leu \rightarrow Ala, for example, the intercept value ($\Delta\Delta G_o$) is -2.1 kcal/mol. This corresponds quite well with the difference between the water-to-octanol transfer free energies of leucine and alanine (1.9 kcal/mol, Fauchére & Pliška, 1983). Ignoring the outlier, Val 149 \rightarrow Ala, the same is true for the valine-to-alanine substitutions (Table 6). For the methionine and isoleucine-to-alanine substitutions, there are, however, discrepancies ranging up to 0.7 kcal/mol (Table 6). We do not know the reason for this divergence. Note, however, that the number of data points for Phe and Met is very limited. Note also that we have, for

illustrative purposes, used the scale of Fauchére & Pliška (1983), but different solvent transfer and hydrophobicity scales show substantial variation among themselves (Cornette et al., 1987). Another major concern is the reliability of the cavity calculations, especially when the volumes are relatively small (see below). In addition, both entropic and enthalpic terms may contribute to the discrepancies seen in Table 6. It might be expected that the entropy cost of localizing a given sidechain within the folded protein would numerically reduce $\Delta\Delta G_o$ relative to ΔG_t . On average, however, $\Delta\Delta G_o$ is numerically *larger* than ΔG_t . This suggests that some other effect, such as the relaxation of strain, may play a significant role. Strain energy in a given sidechain is expected to range up to about 0.7 kcal/mol in some cases (Karpusas et al., 1989). Also a case was reported in which lysozyme stability was increased by

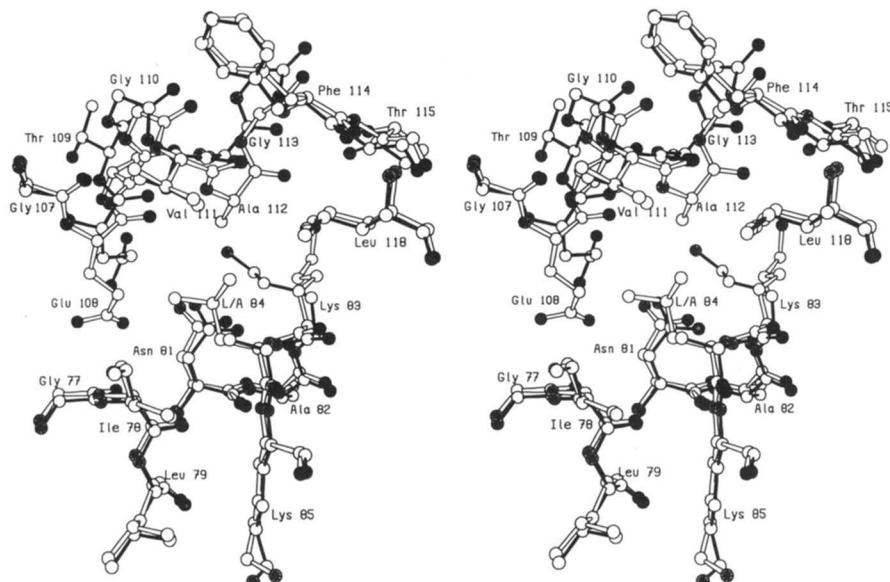


Fig. 15. Superposition of mutant L84A (solid bonds) on WT* (open bonds) based on residues 80–162.

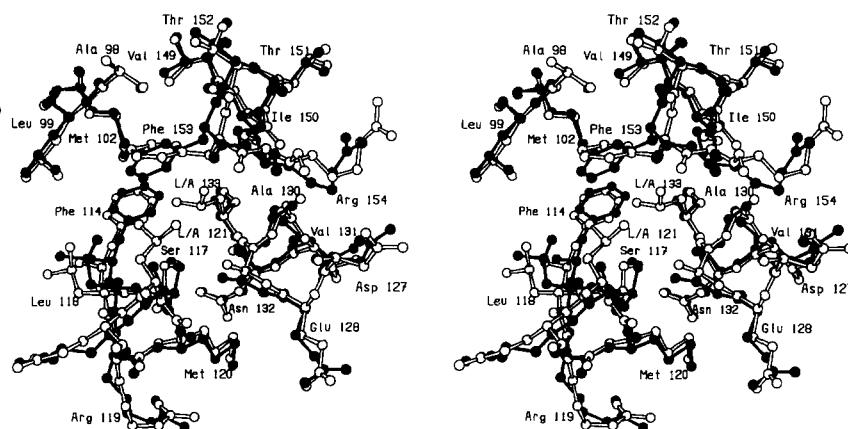


Fig. 16. Superposition of mutant L121A/L133A (solid bonds) on WT* (open bonds) based on residues 80–162.

1.27 kcal/mol due to the apparent removal of strain from the wild-type protein (Blaber et al., 1995). Thus, the consequences of a Leu → Ala or Ile → Ala substitution are not as straightforward as one might at first envisage. Isoleucine and leucine sidechains have essentially equal volumes and equal solvent transfer free energies. They do, however, have different shapes. For this reason, the β -branched sidechain of isoleucine will restrict the flexibility of the local backbone more than leucine. For the same reason, the sidechain of an isoleucine will be more restricted than is the case with leucine (e.g., see Nemethy et al., 1966). Taken together, the lower backbone and sidechain flexibility of isoleucine should diminish entropy loss upon protein folding relative to leucine. Approximately equal numbers of enthalpic contacts would be expected for these two isomeric sidechains. The number of non-local contacts, however, will be smaller for isoleucine due to the same interactions that lower its local flexibility. On balance, it is not straightforward to say *a priori* whether enthalpic or entropic considerations would dominate in either case. It might be noted that the discrepancy between $\Delta\Delta G_o$ and the transfer free energy, ΔG_t , is in the same sense for isoleucine and for methionine (−0.7 and −0.6 kcal/mol, respectively; Table 6) even though the sidechain of one is relatively rigid and the other flexible. This suggests that the

discrepancies do not arise predominantly from differences in side-chain shape or flexibility.

Certainly, the correlation suggested in Figure 17 between destabilization and cavity volume is not expected to apply to non-truncation mutants such as Leu → Val and Met → Val which are likely to introduce strain (see Richards & Lim, 1994).

It would be very helpful if the data included Ile → Ala and other replacements that resulted in larger cavities than are observed in the present sample (Fig. 17). The double mutants I27A/I58A, I29A/I58A and L121A/L133A were constructed with this goal in mind. Although these variants could be crystallized (Table 3) they gradually unfolded over a wide range of temperature in an apparently non-two-state manner.

Relatively few related data are available in the literature. Table 5 gives the loss in stability associated with cavity-forming mutants in barnase (Buckle et al., 1996) and ribonuclease-S (Varadarajan & Richards, 1992). If the cavity volumes quoted by Buckle et al. (1996) are used, there is good agreement between the barnase data and the behavior of the Ile → Ala and Leu → Ala lysozyme mutants shown in Table 5. It should be noted, however, that Buckle et al. (1996) used a different method to calculate cavity volumes and, in addition, used a probe radius of 0.8 Å. If we repeat the calculation for barnase using the procedure of Connolly (1983) with a probe radius of 1.2 Å, much smaller cavity volumes are obtained (Table 5). This points to one of the most difficult issues in analyses of this sort, namely how best to estimate cavity volumes in proteins. We simply have to leave this as an unresolved question.

Table 6. Comparison of extrapolated changes in stability for cavity-creating mutants with solvent transfer values

Amino acid	Free energy of transfer from water to octanol relative to alanine, (ΔG_t), from Fauchére and Pliška (1983) (kcal/mol)	Extrapolated change in free energy, from alanine substitutions ($\Delta\Delta G_o$, Fig. 17) (kcal/mol)	Discrepancy [$\Delta\Delta G_o - \Delta G_t$] (kcal/mol)
Valine	−1.2	−1.2	0
Methionine	−1.2	−1.8	−0.6
Leucine	−1.9	−2.1	−0.2
Phenylalanine	−2.0	−1.8	0.2
Isoleucine	−2.0	−2.7	−0.7

Packing density and thermal motion

Based on four mutations in chymotrypsin inhibitor 2, Fersht and coworkers have suggested that there is an excellent correlation ($r = 0.995$) between the free energy of unfolding relative to wild-type ($\Delta\Delta G$) and the packing density n_c (defined as the number of methyl and methylene groups within 6 Å of the removed atoms) (Otzen et al., 1995). As shown in Figure 18, the lysozyme data are not supportive of this hypothesis. For internal or largely internal sites, n_c depends not so much on the density of atoms in the vicinity, as on the overall number of atoms being removed (the more atoms that are removed, the more surrounding atoms there

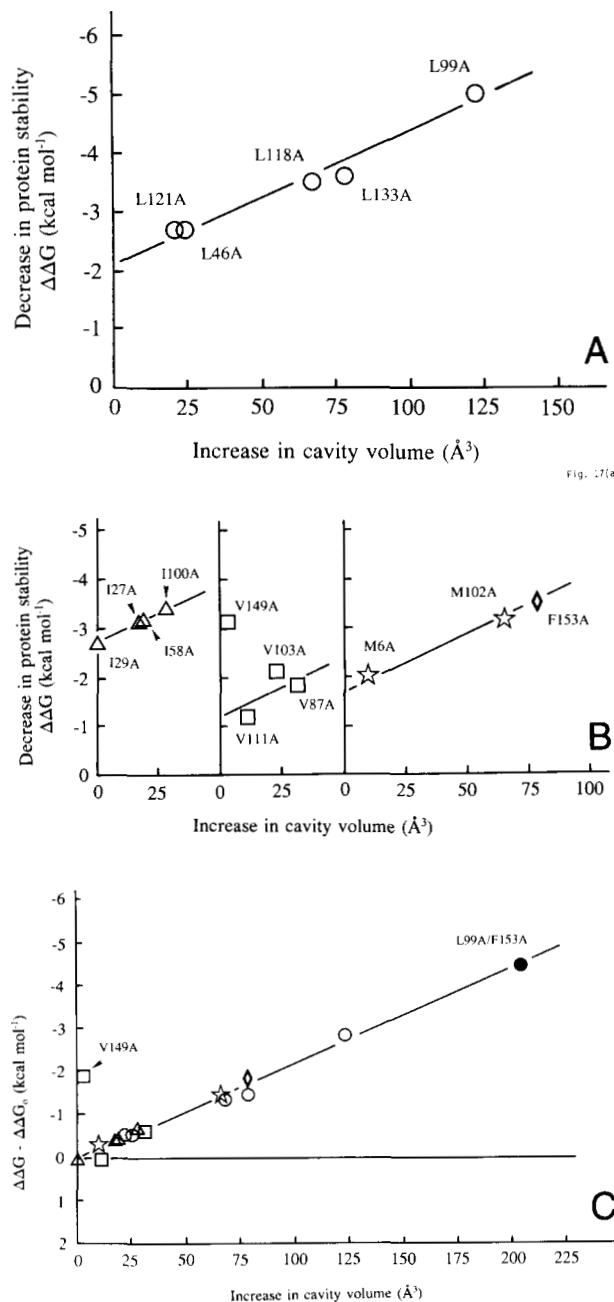


Fig. 17. Plots showing the relationship between $\Delta\Delta G$, the destabilization resulting from cavity-containing point mutants in T4 lysozyme (Table 2) and ΔV_{cav} , the increase in cavity volume [i.e., $V_c(\text{mut}) - V_c(\text{WT}^*)$] from Tables 1 and 4] (after Eriksson et al., 1992). Because the Leu \rightarrow Ala set includes the largest number of observations, and also spans the greatest range of $\Delta\Delta G$ and ΔV_{cav} we have used these observations to determine the slope as $22 \text{ cal mol}^{-1} \text{ \AA}^{-3}$. The other lines are drawn parallel. The intercept values, $\Delta\Delta G_0$, for the different lines, are given in Table 6. **A:** Leucine to alanine substitutions. The data for these mutants are from Eriksson et al. (1992) except for L121A, in which case the structure has been refined with higher resolution data (Baldwin et al., 1996). **B:** Isoleucine to alanine, valine to alanine, methionine to alanine, and phenylalanine to alanine substitutions. Mutant V149A is anomalous in that it binds an internal solvent molecule and is not included in drawing the line through the Val \rightarrow Ala substitutions. **C:** The mutants shown in panels A and B are brought to a common reference scale by subtracting from $\Delta\Delta G$ the intercept value, $\Delta\Delta G_0$. The straight line is drawn through the origin with a slope of $22 \text{ cal mol}^{-1} \text{ \AA}^{-3}$. Data for the double mutant L99A/F153A are from Eriksson et al. (1992).

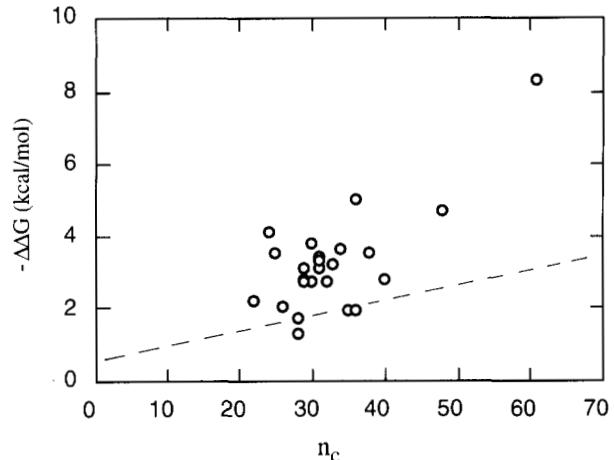


Fig. 18. The relationship between the "packing density," n_c , defined as the number of methyl or methylene groups within 6 \AA of atoms removed in a mutant and the change in free energy ($\Delta\Delta G$) (after Otzen et al., 1995). The straight line shows the correlation between n_c and $\Delta\Delta G$ proposed by Otzen et al. (1995). Lysozyme mutants included in the plot are V87A, V103A, V111A, V149A, I17A, I27A, I29A, I50A, I58A, I100A, M6A, M102A, L46A, L66A, L84A, L91A, L99A, L118A, L121A, L133A, F67A, F153A, L121A/L133A, and F153A/L99A.

will be within 6 \AA). Otzen et al. (1995) derived the proposed dependence of $\Delta\Delta G$ on n_c by including a number of double, triple and quadruple mutations. In cases such as this, n_c increases not because there is increased packing density, but because more and more atoms are being included in its calculation. In the case of T4 lysozyme the plotted data correspond primarily to single substitutions.

Mutations that substantially destabilize T4 lysozyme tend to be restricted to those amino acids that have the lowest thermal motion (Alber et al., 1987; Matthews, 1993). Takano et al. (1995) have also suggested that there might be an inverse correlation between mobility and destabilization for Ile \rightarrow Val substitutions in human lysozyme. The Ile \rightarrow Ala, Leu \rightarrow Ala, and Phe \rightarrow Ala replacements described here, however, do not show any obvious correlation (Fig. 19). It is to be expected that substitutions such as Leu \rightarrow Ala in highly mobile, solvent-exposed, residues will be less destabilizing than equivalent substitutions at rigid, internal sites. The present substitutions, however, were deliberately restricted to buried or largely buried sites, and therefore do not include a wide variation in thermal factors (Fig. 19).

Materials and methods

All mutants were constructed in the cysteine-free pseudo-wild-type lysozyme (WT*) (Matsumura & Matthews, 1989). Methods for generation and purification of the mutant variants as well as thermodynamic analysis were as described previously (Poteete et al., 1991; Eriksson et al., 1993). Crystallographic data were measured on a San Diego Multiwire Systems area detector (Hamlin, 1985) except for mutants I17A and M106A, for which oscillation photography was used (Schmid et al., 1981). The program STRAT (Zhang et al., 1993) was used to optimize data-collection efficiency. Refinements were carried out with TNT (Tronrud et al., 1987) using the procedure described by Baldwin et al. (1996).

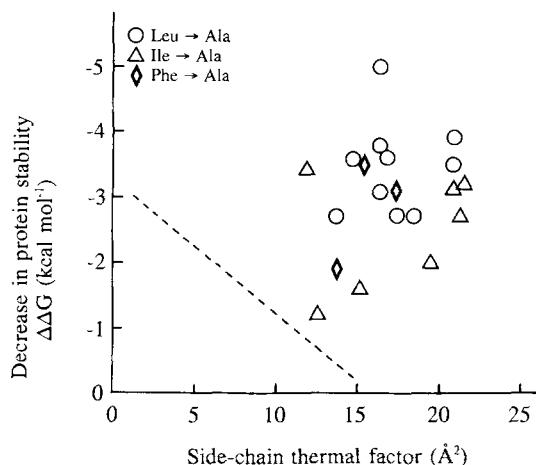


Fig. 19. Destabilization for all buried Leu → Ala, Ile → Ala and Phe → Ala mutants in T4 lysozyme plotted as a function of mobility at the site of substitution. For the purpose of this figure residues were classified as "buried" if their sidechain in the wild-type structure was less than 20% accessible to solvent (Table 1). The thermal factor plotted is the average of the crystallographic B-factors of the sidechain atoms including C^α. The plot was restricted to leucine, isoleucine and phenylalanine since these amino acids have essentially equal solvent transfer free energies and occupy essentially equal volumes. The dashed line shows the apparent correlation for four Ile → Val substitutions in human lysozyme (Takano et al., 1995).

Cavity volumes were calculated by the method of Connolly (1983) in which a sphere is rolled over the surface of the cavity. The total volume that can be occupied by the sphere, V_c , is defined as the volume of the cavity. These values are given in Table 4. Consistent with Eriksson et al. (1992) the radius of the probe sphere was set at 1.2 Å. A probe of larger radius (e.g., 1.4 Å) may be considered as an appropriate mimic of a water molecule, but may not detect smaller cavities. A smaller probe (e.g., 1.0 Å radius or less) may better delineate the shape of smaller, irregular, cavities, but may also escape to the surface of the protein. In general, as the radius of the probe is decreased the calculated cavity volume systematically increases. The choice of the most appropriate probe size is one of the greatest sources of uncertainty in the estimation of cavity volumes. As a representative example, the cavity calculation for mutant V87A with a probe radius of 1.2 Å reveals four cavities of volumes 13 Å³, 10 Å³, 39 Å³, and 23 Å³. If the probe radius is decreased to 1.1 Å, these four cavity volumes increase, respectively, to 14 Å³, 15 Å³, 45 Å³, and 28 Å³ and two new cavities of volumes ~6 Å³ are detected. If the probe radius is further reduced to 1.0 Å, the first three of the original cavities coalesce into a single cavity of volume 84 Å³ and the fourth cavity increases in volume to 36 Å³. An additional five cavities of volumes 5–7 Å³ are also detected. The uncertainty due to errors in the coordinates can be estimated from the variation in the volumes of Cavity I and Cavity II in Table 4. These two cavities occur in WT* plus six mutants, each of which was refined independently, and have root-mean-square volumes of 37 ± 3 Å³ and 22 ± 3 Å³, respectively. (Mutant M102A is excluded because the substitution is close to Cavities I and II.) The observed variation of ± 3 Å³ gives the likely error due to coordinate uncertainty.

As described by Eriksson et al. (1992) the increase in cavity volume due to the mutation was estimated by subtracting the volume of any cavity (or cavities) that might pre-exist in the wild-type

structure ($V_c(\text{WT})$) from the cavity (or cavities) that were observed in the crystal structure of the mutant ($V_c(\text{mut})$). This calculation was carried out within the domain of the protein that included the substitution.

In order to estimate the size of the cavity that would be generated by a given mutation in the absence of any structural rearrangement a model calculation was performed. Starting with the WT* coordinates (Brookhaven Protein Data Bank #1L63) the sidechain in question was truncated to the β-carbon and the size of the resultant cavity was calculated. These values are quoted in Table 4 as V_{model} . There are slight discrepancies with related values quoted by Eriksson et al. (1992) because the radius of the β-carbon used previously was 2.02 Å whereas the present calculation assumes 2.10 Å which is more appropriate for a β-methyl group using a hard sphere model.

Acknowledgments

We thank Dr. Cai Zhang for extensive discussions and help, Sheila Snow and Joel Lindstrom for technical assistance and Drs. Ryota Kuroki and Robert DuBose for the data for mutants I17A and I50A. This work was supported in part by NIH grant GM21967 to B.W.M.

References

- Alber T, Dao-pin S, Nye JA, Muchmore DC, Matthews BW. 1987. Temperature-sensitive mutations of bacteriophage T4 lysozyme occur at sites with low mobility and low solvent accessibility in the folded protein. *Biochemistry* 26:3754–3758.
- Baldwin E, Xu J, Hajiseyedjavadi O, Baase WA, Matthews BW. 1996. Thermodynamic and structural compensation in "size-switch" core repacking variants of T4 lysozyme. *J Mol Biol* 259:542–559.
- Becktel WJ, Scheiman JA. 1987. Protein stability curves. *Biopolymers* 26:1859–1877.
- Bell JA, Wilson KP, Zhang X-J, Faber HR, Nicholson H, Matthews BW. 1991. Comparison of the crystal structure of bacteriophage T4 lysozyme at low, medium and high ionic strengths. *Proteins* 10:10–21.
- Bernal JD. 1939. Structure of proteins. *Nature* 143:663–667.
- Blaber M, Baase WA, Gassner N, Matthews BW. 1995. Alanine scanning mutagenesis of the α-helix 115–123 of phage T4 lysozyme: Effects on structure, stability and the binding of solvent. *J Mol Biol* 246:317–330.
- Brandis JF, Hunt L. 1967. The thermodynamics of protein denaturation. III. The denaturation of ribonuclease in water and in aqueous urea and aqueous ethanol mixtures. *J Am Chem Soc* 89:4826–4838.
- Buckle AM, Cramer P, Fersht AR. 1996. 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. *Biochemistry* 35:4298–4305.
- Connolly ML. 1983. Solvent-accessible surfaces of proteins and nucleic acids. *Science* 221:709–713.
- Cornette JL, Cease KB, Margalit H, Spouge JL, Berzofsky JA, DeLisi C. 1987. Hydrophobicity scales and computational techniques for detecting amphiphilic structures in proteins. *J Mol Biol* 195:659–685.
- Dao-pin S, Baase WA, Matthews BW. 1990. A mutant T4 lysozyme (Val 131 Ala) designed to increase thermostability by the reduction of strain within an α-helix. *Proteins* 7:198–204.
- Dill KA. 1990. Dominant forces in protein folding. *Biochemistry* 29:7133–7155.
- Eriksson AE, Baase WA, Zhang X-J, Heinz DW, Blaber M, Baldwin EP, Matthews BW. 1992. The response of a protein structure to cavity-creating mutations and its relationship to the hydrophobic effect. *Science* 255:178–183.
- Eriksson AE, Baase WA, Matthews BW. 1993. Similar hydrophobic replacements of Leu 99 and Phe 153 within the core of T4 lysozyme have different structural and thermodynamic consequences. *J Mol Biol* 229:747–769.
- Faber HR, Matthews BW. 1990. A mutant T4 lysozyme displays five different crystal conformations. *Nature* 348:263–266.
- Fauchére J-L, Pliska V. 1983. Hydrophobic parameters π of amino acid side-chains form the partitioning of N-acetyl-amino-acid amides. *Eur J Med Chem* 18:369–375.
- Hamlin R. 1985. Multiwire area X-ray diffractometers. *Meth Enzymol* 114:416–452.

- Heinz DW, Baase WA, Zhang X-J, Blaber M, Dahlquist FW, Matthews BW. 1994. Accommodation of amino acid insertions in an α -helix of T4 lysozyme. Structural and thermodynamic analysis. *J Mol Biol* 236:869–886.
- Karpusas M, Baase WA, Matsumura M, Matthews BW. 1989. Hydrophobic packing in T4 lysozyme probed by cavity-filling mutants. *Proc Natl Acad Sci USA* 86:8237–8241.
- Kauzmann W. 1959. Some factors in the interpretation of protein denaturation. *Adv Protein Chem* 16:1–63.
- Kellis JT Jr, Nyberg K, Fersht AR. 1989. Energetics of complementary side-chain packing in a protein hydrophobic core. *Biochemistry* 28:4914–4922.
- Kleywegt GJ, Jones TA. 1994. Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Cryst D50*:178–185.
- Matsumura M, Becktel WJ, Matthews BW. 1988. Hydrophobic stabilization in T4 lysozyme determined directly by multiple substitutions of Ile 3. *Nature* 334:406–410.
- Matsumura M, Matthews BW. 1989. Control of enzyme activity by an engineered disulfide bond. *Science* 243:792–794.
- Matsumura M, Wozniak JA, Dao-pin S, Matthews BW. 1989. Structural studies of mutants of T4 lysozyme that alter hydrophobic stabilization. *J Biol Chem* 264:16059–16066.
- Matthews BW. 1993. Structural and genetic analysis of protein stability. *Ann Rev Biochem* 62:139–160.
- Morton A, Matthews BW. 1995. Specificity of ligand binding in a buried non-polar cavity of T4 lysozyme: Linkage of dynamics and structural plasticity. *Biochemistry* 34:8576–8588.
- Nemethy G, Leach SJ, Scheraga HA. 1966. The influence of amino acid side-chains on the free energy of helix-coil transitions. *J Phys Chem* 70:998–1004.
- Nicholson H, Anderson DE, Dao-pin S, Matthews BW. 1991. Analysis of the interaction between charged sidechains and the α -helix dipole using designed thermostable mutants of phage T4 lysozyme. *Biochemistry* 30:9816–9828.
- Otzen DE, Rheinheimer M, Fersht AR. 1995. Structural factors contributing to the hydrophobic effect: The partly exposed hydrophobic minicore in chymotrypsin inhibitor 2. *Biochemistry* 34:13051–13058.
- Poteete AR, Dao-pin S, Nicholson H, Matthews BW. 1991. Second-site revertants of an inactive T4 lysozyme mutant restore activity structuring the active site cleft. *Biochemistry* 30:1425–1432.
- Richards FM. 1977. Areas, volumes, packing and protein structure. *Ann Rev Biophys Biophys Eng* 6:151–176.
- Richards FM, Lim WA. 1994. An analysis of packing in the protein folding problem. *Quart Rev Biophys* 26:423–498.
- Schmid MF, Weaver LH, Holmes MA, Grüter MG, Ohlendorf DH, Reynolds RA, Remington SJ, Matthews BW. 1981. An oscillation data collection system for high-resolution protein crystallography. *Acta Cryst A37*:701–710.
- Sharp KA, Nicholls A, Fine RF, Honig B. 1991. Reconciling the magnitude of the microscopic and macroscopic hydrophobic effects. *Science* 252:106–109.
- Shortle D, Stites WE, Meeker AK. 1990. Contributions of the large hydrophobic amino acids to the stability of staphylococcal nuclease. *Biochemistry* 29:8033–8041.
- Takano K, Ogasahara K, Kaneda H, Yamagata Y, Fujii S, Kanaya E, Kikuchi M, Oobatake M, Yutani K. 1995. Contribution of hydrophobic residues to the stability of human lysozyme: Calorimetric studies and X-ray structural analysis of the five isoleucine to valine mutants. *J Mol Biol* 254:62–76.
- Takano K, Yamagata Y, Fujii S, Yutani K. 1997. Contribution of the hydrophobic effect to the stability of human lysozyme: Calorimetric studies and X-ray structural analyses of the nine valine to alanine mutants. *Biochemistry* 36:688–698.
- Tronrud DE, Ten Eyck LF, Matthews BW. 1987. An efficient general-purpose least-squares refinement program for macromolecular structures. *Acta Cryst A43*:489–503.
- Varadarajan R, Richards FM. 1992. Crystallographic structures of ribonuclease S variants with nonpolar substitution at position 13: Packing and cavities. *Biochemistry* 31:12315–12327.
- Yutani K, Ogasahara K, Tsujita T, Sugino Y. 1987. Dependence of conformation stability on hydrophobicity of the amino acid residue in a series of variant proteins substituted at a unique position of tryptophan synthase α -subunit. *Proc Natl Acad Sci USA* 84:4441–4444.
- Zhang X-J, Matthews BW. 1993. STRAT: A program to optimize X-ray data collection on an area detector system. *J Appl Cryst* 26:457–462.
- Zhang X-J, Wozniak JA, Matthews BW. 1995. Protein flexibility and adaptability seen in 25 crystals forms of T4 lysozyme. *J Mol Biol* 250:527–552.