Protein Science (1996), 5:2399–2415. Cambridge University Press. Printed in the USA.
Copyright © 1996 The Protein Society

Protein structural plasticity exemplified by insertion and deletion mutants in T4 lysozyme

INGRID R. VETTER,[1] WALTER A. BAASE, DIRK W. HEINZ,[2] JIAN-PING XIONG, SHEILA SNOW, AND BRIAN W. MATTHEWS

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

(RECEIVED August 13, 1996; ACCEPTED September 17, 1996)

Abstract

To further investigate the ways in which proteins respond to changes in the length of the polypeptide chain, a series of 32 insertions and five deletions were made within nine different α-helices of T4 lysozyme. In most cases, the inserted amino acid was a single alanine, although in some instances up to four residues, not necessarily alanine, were used. Different insertions destabilized the protein by different amounts, ranging from approximately 1 to 6 kcal/mol. In one case, no protein could be obtained. An "extension" mutant in which the carboxy terminus of the molecule was extended by four alanines increased stability by 0.3 kcal/mol. For the deletions, the loss in stability ranged from approximately 3 to 5 kcal/mol. The structures of six insertion mutants, as well as one deletion mutant and the extension mutant, were determined, three in crystal forms nonisomorphous with wild type. In all cases, including previously described insertion mutants within a single α-helix, there appears to be a strong tendency to preserve the helix by translocating residues so that the effects of the insertion are propagated into a bend or loop at one end or the other of the helix. In three mutants, even the hydrophobic core was disrupted so as to permit the preservation of the α-helix containing the insertion. Translocation (or "register shift") was also observed for the deletion mutant, in this case a loop at the end of the helix being shortened. In general, when translocation occurs, the reduction in stability is only moderate, averaging 2.5 kcal/mol. Only in the most extreme cases does "bulging" or "looping-out" occur within the body of an α-helix, in which case the destabilization is substantial, averaging 4.9 kcal/mol. Looping-out can occur for insertions close to the end of a helix, in which case the destabilization is less severe, averaging 2.6 kcal/mol. Mutant A73-[AAA], as well as mutants R119-[A] and V131-[A], include shifts in the backbone of 3–6 Å, extending over 20 residues or more. As a result, residues 114–142, which form a "cap" on the carboxy-terminal domain, undergo substantial reorganizations such that the interface between this "cap" and the rest of the protein is altered substantially. In the case of mutant A73-[AAA], two nearby α-helices, which form a bend of approximately 105 degrees in the wild-type structure, reorganize in the mutant structure to form a single, essentially straight helix. These structural responses to mutation demonstrate the plasticity of protein structures and illustrate ways in which their three-dimensional structures might change during evolution.

Keywords: deletions; insertions; lysozyme; structural plasticity; translocation

It is generally agreed that insertions within α-helices and β-sheets of proteins are tolerated less well than insertions in loops or near the ends of secondary structure elements (e.g., Pascarella & Argos, 1992). In an analysis of 1,400 examples of insertions and deletions found in the Protein Data Bank, only 24 were observed within α-helices and 20 within β-strands. For the helices, all the insertions occurred within four residues of the helix termini and typically resulted in a helical extension of a few residues (Pascarella & Argos, 1992). From these data, it might appear obvious that insertions and deletions within secondary structure elements are unfavorable for the protein and thus are eliminated during evolution. On the other hand, it has been shown in a number of cases that proteins will tolerate genetically engineered insertions and deletions, even within α-helices (Sondek & Shortle, 1990, 1992; Marti et al., 1992; Heinz et al., 1993, 1994; Kavanaugh et al., 1993; Keefe et al., 1993, 1994; Sharff et al., 1995). In many cases, the resultant loss of stability is comparable with that resulting from a destabilizing point mutation (Sondek & Shortle, 1990; Heinz et al., 1994). Thus, it may be that insertions within secondary structure elements are not as deleterious as might be inferred from an analysis of known structures (Heinz et al., 1994).

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

[1]Present address: Max-Planck Institute for Molecular Physiology, Postfach 102664, 44026 Dortmund, Germany.

[2]Present address: Institut fuer Organische Chemie und Biochemie, Universitaet Freiburg, Albertstrasse 21, D-79104 Freiburg i. Breisgau, Germany.

A previous analysis of a series of insertions into a representative α-helix on the surface of T4 lysozyme (Heinz et al., 1993, 1994) suggested that the protein could accommodate insertions in different ways. One response was a "looping-out" of the inserted residues. This disrupted and destabilized the α-helix and was observed only rarely. The more common response consisted of a register shift, or translocation, of the adjacent residues such that the helix was preserved and the structural perturbation propagated into a loop or bend at one end of the helix. The latter finding has led to a different way of looking at the apparent infrequency of insertions (or deletions) in secondary structure elements in the database; specifically, a mutation that occurs within secondary structure may not appear as such, but can be shifted into the loop and turn regions where such mutations are commonly found.

In general, deletions are expected to be more destabilizing than insertions because there are fewer possible ways to compensate for a deletion; an insertion within an α-helix might induce a register shift into a neighboring loop. This could, however, be more difficult for a deletion because of the length constraint of the loop. Moreover, a deletion cannot loop-out like an insertion.

To date, relatively few structures of mutants containing an insertion or deletion in a secondary structure element have been described (Heinz et al., 1993, 1994; Kavanaugh et al., 1993; Keefe et al., 1993, 1994; Sharff et al., 1995). To obtain a better understanding of this phenomenon, a series of mutants was constructed in which a single alanine was inserted at one or more sites within each of the α-helices of T4 lysozyme (Fig. 1). The helices are as defined by Nicholson et al. (1991), although it should be emphasized that there is often an uncertainty of one or two amino acids in specifying the beginning and end of a helix. (Fig. 2 has additional information on helix location in cases where the structures of mutant proteins were obtained.)

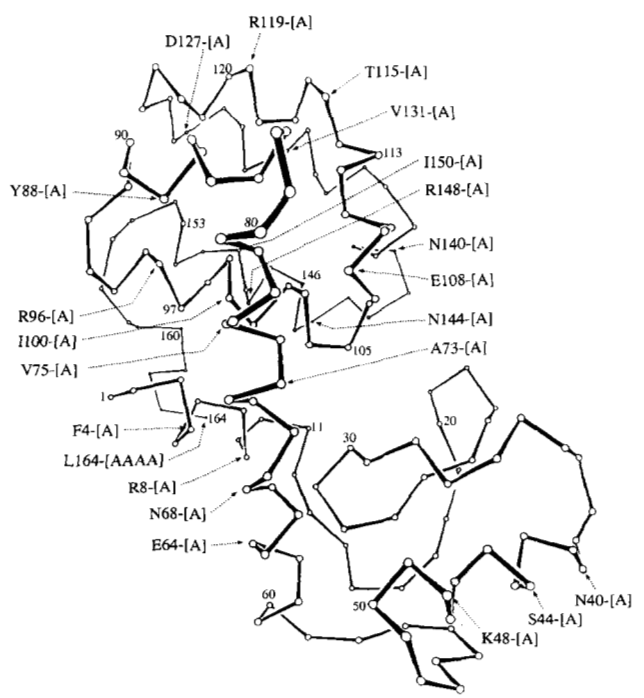

Fig. 1. Schematic drawing of the C^{α} backbone of T4 lysozyme showing the locations of insertion and deletion mutants. The figure shows a representative variant at each site (i.e., it does not include all variants).

To investigate α-helix 143–155 in more detail, a series of insertions of up to four alanines was made at position 148, as well as a series of five "revertants," in which the alanines inserted after Arg 148 were allowed to revert randomly back to other amino acids in an attempt to regain stability (cf. Heinz et al., 1993, 1994). The long α-helix, 60–79, which bridges the two domains of the lysozyme molecule (Fig. 1), was also subject to a variety of insertions following Ala 73. A series of five mutants with the deletion of a single amino acid was also constructed. Also, an extension of four alanines was made at the carboxy terminus of the molecule. Mutants are identified as follows: 148-[D], for example, indicates the mutant in which an aspartic acid is inserted following residue 148; mutant 148-[DS] has Asp-Ser inserted following residue 148; R8Δ is a deletion mutant with Arg 8 deleted; mutant L164-[AAAA] has a four-alanine extension beyond Leu 164, i.e., at the carboxy terminus of the molecule. All mutants were constructed in a pseudo-wild-type cysteine-free lysozyme designated WT* (Matsumura & Matthews, 1989).

Results

Design of mutants

Systematic analysis of a single α-helix

In previous studies, Heinz et al. (1993, 1994) analyzed the insertion mutants at various sites within α-helix 39–50 of T4 lysozyme (Fig. 1). To test the generality of these results, a series of single, double, triple, and quadruple substitutions was made between residues Arg 148 and Val 149 of the α-helix that includes residues 143–155 (Figs. 1, 2). Analogous to helix 39–50 of T4 lysozyme, and to most helices in proteins, helix 143–155 is amphipathic, with one side buried and the other exposed to solvent. It is also not involved in crystal contacts which, it was hoped, would simplify crystallization.

A series of different insertions was also made between Ala 73 and Ala 74 in the long α-helix that connects the two domains of the lysozyme molecule. These insertions present a severe challenge for the integrity of the molecule because they will tend to change substantially the alignment of one domain relative to the other (Fig. 1).

Analysis of alanine insertion within different α-helices

Because it is not possible to obtain a statistically meaningful result from only two α-helices, a second series of insertions was designed. Here, a single alanine was inserted at one or more positions in every remaining helix of the protein, the sites chosen being on the solvent-accessible rather than the buried face of the helix (Fig. 2). Helix 93–106 is largely buried and in this case the site of insertion followed Ile 100, which has only limited accessibility to solvent (Table 1).

An "extension" mutant

An extension of T4 lysozyme was designed to determine the feasibility of extending or attaching another secondary structure element to the end of the protein. The wild-type protein terminates at Leu 164. Residues 159–161 are in the conformation of a 3_{10}-helix, whereas the last two residues are disordered in the wild-type crystal structure (Weaver & Matthews, 1987). Four alanines were added after Leu 164 to see if they might stabilize the two disordered residues and/or form a continuation of the helix.

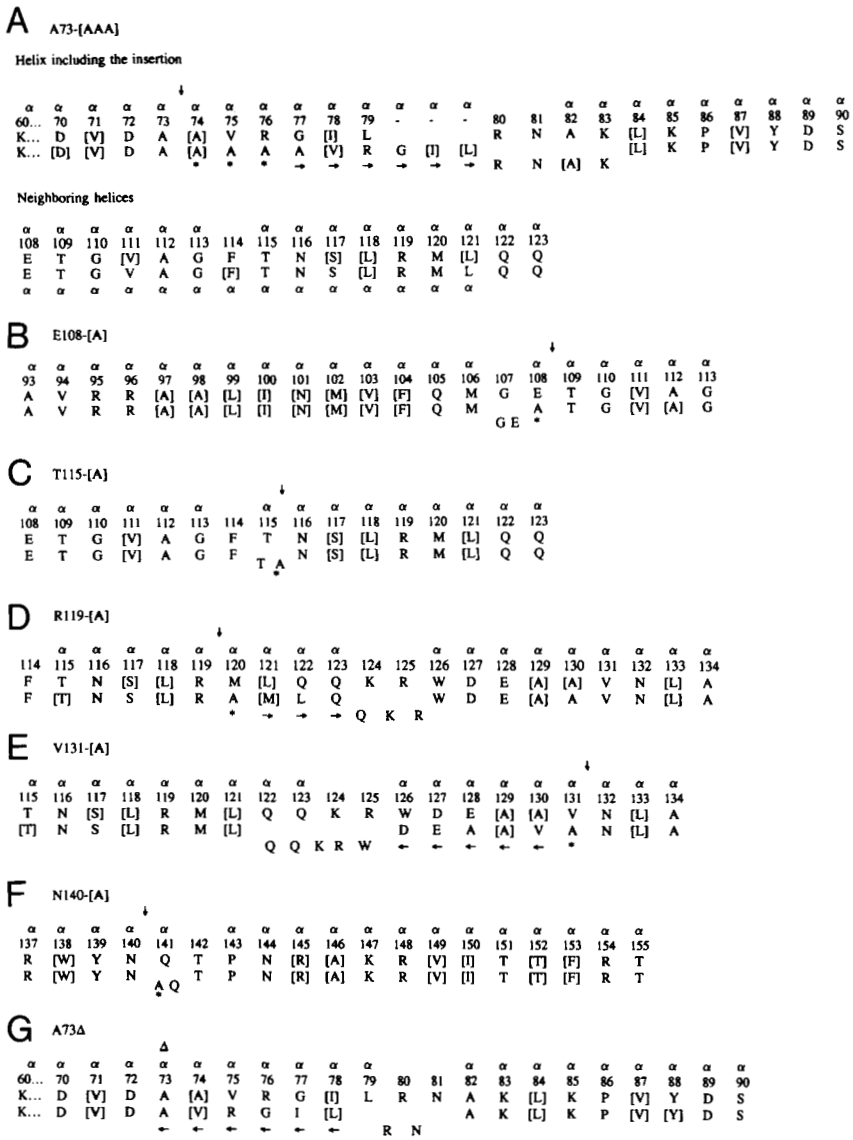

Fig. 2. Schematic drawing summarizing the response of T4 lysozyme to insertions and deletions in cases where crystal structures were determined. The first three lines give the numbered amino acid sequence of the wild-type protein with the residues within α-helices denoted "α." The fourth line gives the sequence of the mutant aligned in accordance with the structural correspondence between mutant and wild type. For example, in the case of A73-[AAA], Val 75 in the mutant structure is translocated three positions to occupy the site previously occupied by Ile 78 in wild type. In the case of A73-[AAA], the mutation causes two neighboring helices to meld into one (see text), as indicated at the bottom of the relevant panel. Insertion and deletion sites are indicated by arrowheads. Inserted residues are indicated by stars. Residues that are translocated are indicated by small arrowheads in the bottom line. Residues that "loop-out" in the three-dimensional structure are "looped-out" in the figure (e.g., Gly 107–Glu 108 in mutant E108-[A]). Residues that are largely buried (i.e., less than 12% surface area accessible to solvent calculated using the program "NACCESS," courtesy of S.J. Hubbard and J.M. Thornton, University College, London) are enclosed by brackets.

Deletion mutants

Five mutants with a single-residue deletion were constructed. They were made at positions corresponding to five of the insertions, namely Arg 8, Ser 44, Ala 73, Arg 119, and Asp 127. Four of these sites are in the middle of an α-helix, whereas one (Asp 127) is closer to an end.

Thermal stability and activity

All mutants with insertions after position Arg 148 formed inclusion bodies when expressed in *Escherichia coli*, indicating substantial destabilization. This loss of stability is also apparent in Table 2. The single alanine insertions reduced the free energy of unfolding by 0.7–6.6 kcal/mol at pH 5.4, relative to WT*. In contrast, the extension mutant L164-[AAAA] was 0.3 kcal/mol more stable than WT* (Table 1). The deletion mutants had free energies of unfolding 2.8–5.4 kcal/mol less than WT* (Table 1).

Because we were interested primarily in the stabilities and structures of the mutant proteins, we did not attempt to measure their catalytic activities. Their behavior in the halo assay, however, which provides an overall indication of stability and activity (Streisinger et al., 1961), was as follows: mutants S44Δ, V75-[A], T115-[A],

Table 1. *Thermodynamic data for single amino acid insertions and deletions within different helices in T4 lysozyme*

Mutant	Solvent accessibility (%)[a]	Helix[b]	T_m (°C)	pH 5.4 ΔH (kcal/mol)	pH 5.4 $\Delta\Delta G$ (kcal/mol)	pH 3.0 $\Delta\Delta G$ (kcal/mol)
Insertions						
F4-[A]	52	3–9	56.8	102	−3.1	−3.7
R8-[A]	61	3–9	46.1	56	−5.6	N.D.
N40-[A]	108	39–50	57.3[c]	97[c]	−2.8[c]	−2.6
S44-[A]	90	39–50	53.4[c]	55[c]	−3.8[c]	−3.2
K48-[A]	80	39–50	53.8[c]	65[c]	−3.7[c]	−2.9
E64-[A]	59	60–79	54.7	92	−3.7	−4.0
N68-[A]	81	60–79	52.1	82	−4.4	−4.7
A73-[A]	73	60–79	52.8	103	−4.2	−5.3
V75-[A]	34	60–79	57.4	115	−2.7	−3.5
Y88-[A]	20	82–90	53.9	105	−3.9	−4.2
R96-[A]	41	93–106	36.4	23	−6.6	N.D.
I100-[A]	6	93–106		Protein not obtained		
E108-[A]	42	93–106	58.1	105	−2.7	−2.8
T115-[A]	88	115–123	60.5	117	−1.8	−2.0
R119-[A]	63	115–123	58.6	97	−2.6	−2.6
D127-[A]	107	126–134	55.1	97	−3.5	−3.3
V131-[A]	86	126–134	57.3	104	−2.9	−3.0
N140-[A]	94	137–141	59.6	115	−2.1	−2.2
N144-[A]	102	143–155	63.2	127	−0.7	−0.8
K147-[A]	68	143–155	59.3	115	−2.1	−1.9
R148-[A]	27	143–155	51.5	88	−4.6	−4.0
I150-[A]	13	143–155	40.9	34	−6.0	−4.0
Deletions						
R8Δ	61	3–9	45.8	45	−5.4	N.D.
S44Δ	90	39–50	58.1	93	−2.8	−3.0
A73Δ	73	60–79	53.8	98	−3.9	−4.6
R119Δ	63	115–123	55.1	95	−3.6	−3.8
D127Δ	107	126–134	56.8	100	−3.1	−2.7
Extension						
L164-[AAAA]	50	159–162	65.8	131	0.2	0.3

[a]Solvent accessibility of the side chain preceding the insertion (e.g., Phe 4) relative to the same side chain in a fully extended polypeptide chain (Alber et al., 1987).
[b]Nicholson et al. (1991).
[c]Data from Heinz et al. (1993, 1994).

N140[A], and L164-[AAAA] had clear halos at both 37 °C and at room temperature; mutants F4-[A], E64-[A], N68-[A], A73Δ, Y88-[A], E108-[A], R119-[A], D127[A], D127Δ, V131-[A], and N144-[A] had halos at room temperature, but not at 37 °C; mutants R8-[A], R8Δ, R96-[A], I100-[A], R119Δ, I150-[A], A73-[AAA], and R148-[AAA] had no apparent halos.

Crystal structure determinations

Crystals isomorphous to WT* were obtained for mutants E108-[A], T115-[A], N140-[A], L164-[AAAA], and A73Δ. Using the same conditions, mutant A73-[AAA] crystallized in space group $P6_222$. L164-[AAAA] was crystallized initially using 2.2 M phosphate, 0.25 M NaCl, 50 mM reduced β-mercaptoethanol, pH 7.1, and yielded twinned crystals. Addition of a small amount of dioxane (0.25%, v/v) to the crystallization buffer allowed the growth of crystals isomorphous with WT* suitable for data collection. Crystallization using different conditions (18% PEG 8K, 0.1 M Hepes, 50 mM β-mercaptoethanol, pH 7.5) also yielded crystals isomorphous with WT* (cf. Bell et al., 1991). Mutant V131-[A] was crystallized in space group $P2_12_12_1$ using 20% PEG 4K, 0.1 M phosphate, 0.1 M sodium acetate, pH 6.6, at 15 °C. R119-[A] crystallized in the hexagonal space group $P6_5$ using 2.6 M phosphate, 0.25 M NaCl, 50 mM reduced β-mercaptoethanol, pH 7.1. Addition of dioxane (0.25%, v/v) was necessary to give crystals suitable for data collection. Crystals were obtained of R148-[D], but were too small for data collection. No crystals have been obtained to date for the other mutants. Coordinates of all structures determined have been submitted to the Brookhaven Data Bank for immediate release.

Insertion and extension mutants

To anticipate the more detailed structural results that will follow, Figure 2 summarizes the locations of the key α-helices and the overall structural responses that were observed in those cases where crystal structures were obtained (Table 3).

Table 2. *Thermodynamic data for insertion mutants following Ala 73 and Arg 148*

Mutant	Unfolding at pH 5.4			Unfolding at pH 3.0		
	ΔT_m (°C)	ΔH (kcal/mol)	$\Delta\Delta G$ (kcal/mol)	ΔT_m (°C)	ΔH (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
WT*	65.22[a]	131	0	51.66[a]	114	0
A73-[A]	−12.5	103	−4.2	−19.7	57	−5.3[b]
A73-[AA]	−18.0	75	−5.7[b]	−29.2	25	−6.2[b]
A73-[AAA]	−15.9	69	−5.0[b]	−21.6	31.5	−4.8[b]
A73-[L]	−9.0	109	−3.1	−14.2	74	−4.3
A73-[R]	−8.2	110	−2.9	−17.6	68	−5.1[b]
A73-[VL]	−5.2	123.5	−1.7	−5.8	98	−2.0
R148-[A]	−13.7	88	−4.6	−14.4	54	−4.0
R148-[AA]	−17.5	67	−5.4[b]	−19.5	34.5	−4.5
R148-[AAA]	−15.9	72	−5.0[b]	−20.0	31	−4.5
R148-[AAAA]	−23.2	31	−5.6[b]		Not two-state	
R148-[D]	−14.7	77	−4.8[b]	−9.1	70	−3.0
R148-[S]	−15.7	78	−5.0	−17.0	43	−4.3
R148-[DS]	−19.0	60	−5.6[b]	−20.7	24	−4.3
R148-[TT]	−20.5	55	−5.9[b]	−19.9	21	−4.1
R148-[VP]	−22.0	55	−6.3[b]	−18.0	18.5	−3.8

[a]The number quoted is the T_m of WT*.
[b]Estimated error in excess of ±0.5 kcal/mol (see text).

Mutant A73-[AAA] is especially interesting in that it has an insertion of three residues in the middle of the long α-helix that connects the two domains of T4 lysozyme (Fig. 1). The insertion caused a translocation of residues 74–79 toward the C terminus, thereby extending the helix by one turn (Fig. 2). Associated with this lengthening of the interdomain helix, the amino terminus of the helix that follows (residues 82–90) moves substantially away from the protein and the helix itself is tilted by approximately 25 degrees relative to its alignment in wild type (Fig. 3A). The helix is also shortened by two amino acids. The space that is created by the movement of helix 82–90 is occupied by the side chains of Leu 79 and Tyr 88. In the wild-type structure, the neighboring helices 108–113 and 115–123 form a bend with an angle of approximately 105 degrees. In the mutant structure, these residues reorganize to form a single, essentially straight helix extending from residue 108 to 121 (Fig. 2A, 3B). Helix 126–134 also undergoes a substantial change in alignment (Fig. 3B) so that, in total, essentially every backbone atom between residues 105 and 140 undergoes a shift of 3 Å to 8 Å (Fig. 4G). In association with these movements, a deep depression is formed on the surface of the protein (Fig. 3C) and part of the hydrophobic core including Val 111 and Leu 121 is exposed to solvent (Fig. 2A). As a result,

Table 3. *X-ray data collection and refinement statistics*

Mutant	E108-[A]	A73-[AAA]	T115-[A]	R119-[A]	N140-[A]	V131-[A]	A73Δ	L164-[AAAA] (PEG)	L164-[AAAA] (phosphate)
Data collection									
Space group	$P3_221$	$P6_222$	$P3_221$	$P6_5$	$P3_221$	$P2_12_12_1$	$P3_221$	$P3_221$	$P3_221$
Cell dimensions									
a (Å)	61.0	95.9	60.9	84.7	61.4	36.6	60.8	60.9	60.6
b (Å)	61.0	95.9	60.9	84.7	61.4	61.4	60.8	60.9	60.6
c (Å)	97.1	117.9	97.3	48.0	97.5	88.7	97.0	96.7	97.5
Resolution (Å)	1.7	2.7	1.9	1.9	2.0	2.1	2.0	1.8	1.7
Unique reflections	22,695	7,329	14,180	13,196	11,720	11,106	16,587	20,392	21,597
Completeness (%)	95	80	93	83	74	86	96	95	88
R_{merge} (%)[a]	4.5	6.9	4.6	6.6	7.5	5.2	7.4	6.2	6.8
Refinement									
R (%)[b]	17.3	17.9	19.7	18.5	18.2	19.2	17.3	15.6	17.5
Bond discrepancy (Å)	0.014	0.0015	0.015	0.016	0.019	0.019	0.017	0.013	0.014
Angle discrepancy (°)	2.0	2.9	2.2	2.4	2.7	2.8	2.4	2.0	2.1

[a]R_{merge} is the average agreement between repeated intensity measurements.
[b]*R* is the crystallographic residual following refinement, including all measured data. Bond and angle discrepancies are the average discrepancies of the bond lengths and angles from expected stereochemical values.

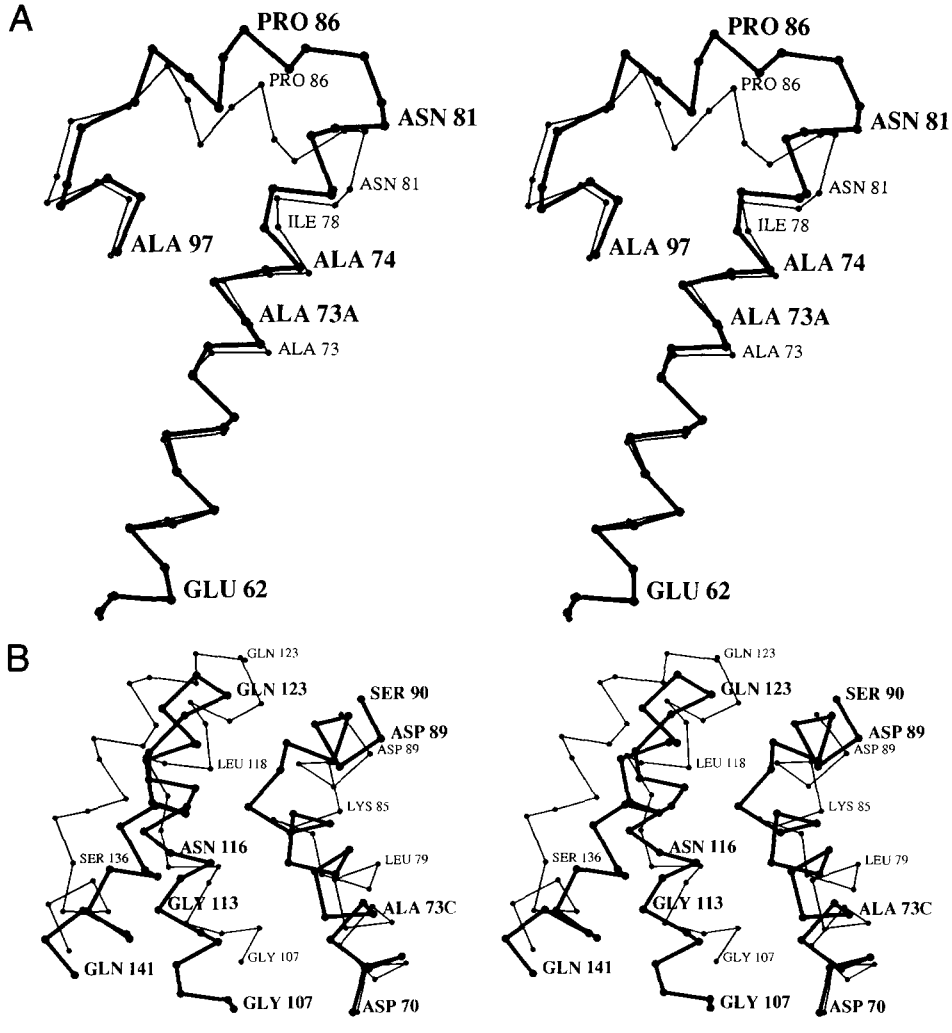

Fig. 3. A: Stereo drawing comparing the C^{α}-backbone of the wild-type structure (thin lines and smaller labels) and the insertion mutant A73-[AAA] (thick lines and larger labels) in the immediate vicinity of the mutation (residues 60–97). The superposition is based on residues 1–70 and 144–162. **B:** Comparison of the structure of mutant A73-[AAA] (thick bonds) with wild-type lysozyme (thin bonds) showing the associated movement of α-helices 115–123 and 126–134. **C:** Left panel compares the surfaces of WT* (green) and A73-[AAA] (cyan). The superposition is based on residues 1–70 and 144–162; residues from 70 to 162 are included in the figure. Four side chains are labeled and shown in red to help orient the viewer. Right panel shows the superposition of the α-carbon backbones of WT* (green) and A73-[AAA] (yellow) in the same orientation. The depression that occurs in the mutant structure is at the top left, in the vicinity of Lys 124 and Trp 126. (*Figure continues on facing page.*)

the surface area of A73-[AAA] that is made up of nonpolar atoms increased by 10% relative to wild type, whereas the polar surface area remains essentially unchanged. Solvent accessibility calculations show that Asp 70 is essentially buried in mutant A73-[AAA]. This is because Gln 105 moves to largely cover the side chain of Asp 70, and presumably hydrogen bonds to it. The apparent inaccessibility to solvent of Asp 70 does not, therefore, imply that it is buried within the hydrophobic core.

The insertion E108-[A] is located close to a site where one helix ends and another begins (Fig. 2). In the WT* structure, the helical-type hydrogen bonding continues from one helix to the other so that there is no loop between the two helices. This hydrogen bonding pattern is disrupted by the inserted alanine, and a short loop is created. The structural changes are shown in more detail in Figure 5. The helices that precede and follow the inserted alanine are essentially unaffected.

The T115-[A] insertion is located at a site where two helices come together at an obtuse angle of approximately 105 degrees (Fig. 6). The turn connecting the helices is slightly widened, but otherwise, no other residues are influenced significantly. The modest changes that occur when this insertion is made between helices 108 and 113 and helices 115 and 123 can be contrasted with the complete reorganization that occurs in this region in association with mutant A73-[AAA] (see above).

Mutant R119-[A] crystallized in a space group different from wild type. In common with A73-[AAA] and V131-[A] (described below), it resulted in pronounced structural changes in the carboxy-terminal domain of the molecule, especially in α-helices 115–123 and 126–134 (Figs. 4, 7). The mutation is located in the middle of the first of these helices. The inserted amino acid causes a translocation of residues 120–123 toward the C terminus of the helix. Met 120, which points outward in the wild-type structure, moves

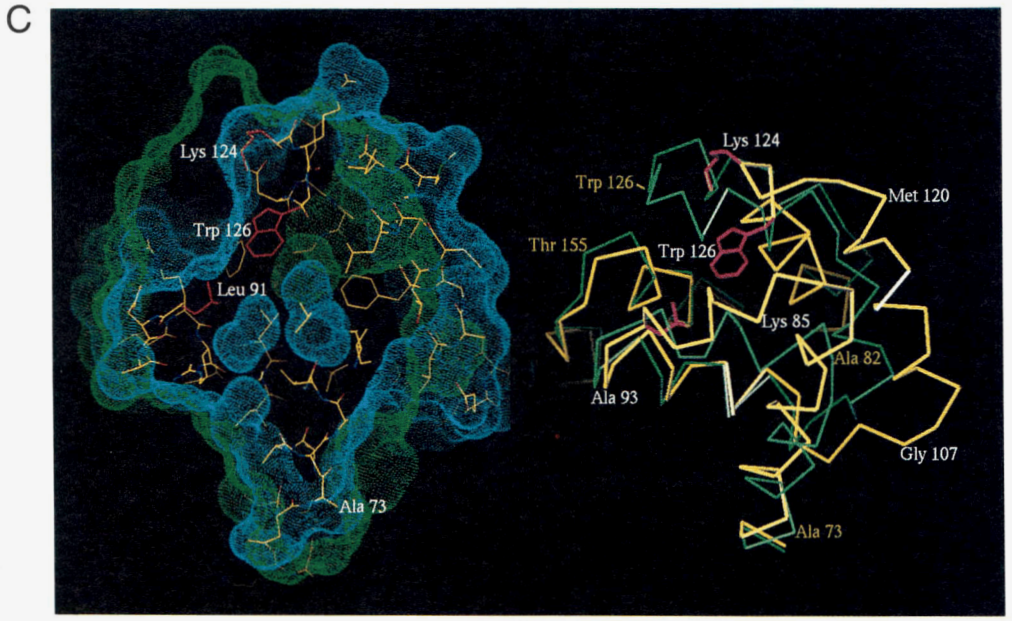

Fig. 3. *Continued.*

toward the core of the protein and fills the space between helices 115–123 and 126–134, the latter helix in turn moving away from the core. The remaining space separating helix 126–134 from the core is occupied by the side chain of Leu 133, which moves toward Met 120, and by the side chain of Phe 153. The latter fills the space occupied by Leu 121 in the wild type. Similar to V131-[A], the side chain of Trp 126 switches to the outside of the molecule and packs against a symmetry-related molecule. The loop connecting helices 115–123 and 126–134 also participates in contacts between two symmetry-related molecules.

Mutant V131-[A] crystallized in space group $P2_12_12_1$ and, as noted above, displayed substantial structural rearrangements in the C-terminal domain (Figs. 4, 8A). Although the site of the insertion is closer to the C terminus than to the N terminus of the helix, it caused the translocation of residues 126–131 toward the N terminus (Fig. 2). The helix containing the insertion is displaced by up to 4 Å and the adjacent loop undergoes a major rearrangement (Fig. 8A). Trp 126 switches to the outside of the molecule (Fig. 8B) and Leu 121 moves toward the core. The core of the protein is disrupted and a deep depression is formed. A hydrophobic patch on the surface of the molecule that is created by the mutation interacts with the hydrophobic patch of a symmetry-related molecule in the crystal lattice.

The N140-[A] insertion is positioned close to a bend between helix 126–134 and the short helix 137–141. As depicted in Figure 9, the mutation results in a widening of the bend connecting the two helices. This in turn leads to a rearrangement of side chains: Gln 141 points inside the active site cleft, whereas it is pointing to the outside in WT*, and Glu 22, on the opposite side of the active site cleft (Fig. 1), moves away by approximately 0.8 Å (Fig. 4).

The "extension" mutant, L164-[AAAA], in which four alanines were added to the C terminus, was crystallized from phosphate and from PEG, in both cases isomorphous with the wild type. In the electron density map for the crystals grown from phosphate, there was moderately strong electron density for the last two residues of the wild-type chain, Asn 163 and Leu 164, which are disordered in the wild-type structure. However, no density was apparent for the four terminal alanines. In contrast, the crystals grown from PEG did show density, albeit weak, for these four residues. The alanines do not form an α-helix, but are arranged as a sort of cap, protecting the side chain of Ile 9. This shielding of a hydrophobic surface might account for the slightly increased stability of this mutant as compared with WT*.

Deletion mutants

The only deletion mutant that formed crystals was A73Δ. This mutant is especially interesting because the deletion is in the middle of the long α-helix connecting the two domains of the protein (Fig. 1). The deletion causes a shift in register of the six C-terminal residues of the helix (Fig. 2). As a consequence, the loop at the C terminus of the helix is also shortened. The arrangement of side chains is such that the register shift can be achieved without substantial disruption of the hydrophobic core. The largest change is due to the replacement of Gly 77 with Ile 78, which causes the neighboring helix to be shifted up to 2.7 Å (Figs. 4, 10A,B). Repositioning of the isoleucine side chain creates a cleft between helix 60–80 and helix 107–113, where electron density apparently corresponding to a hydroxyethyldisulfide molecule (i.e., oxidized β-mercaptoethanol), present in the crystallizing medium, could be seen (not included in Fig. 10A).

Discussion

Structural plasticity

One of the most interesting aspects of the mutants described here is the variability and the extent of the changes in structure that they incur. In general, the structures of mutant proteins have been found to be very similar to their parents, with changes localized to the

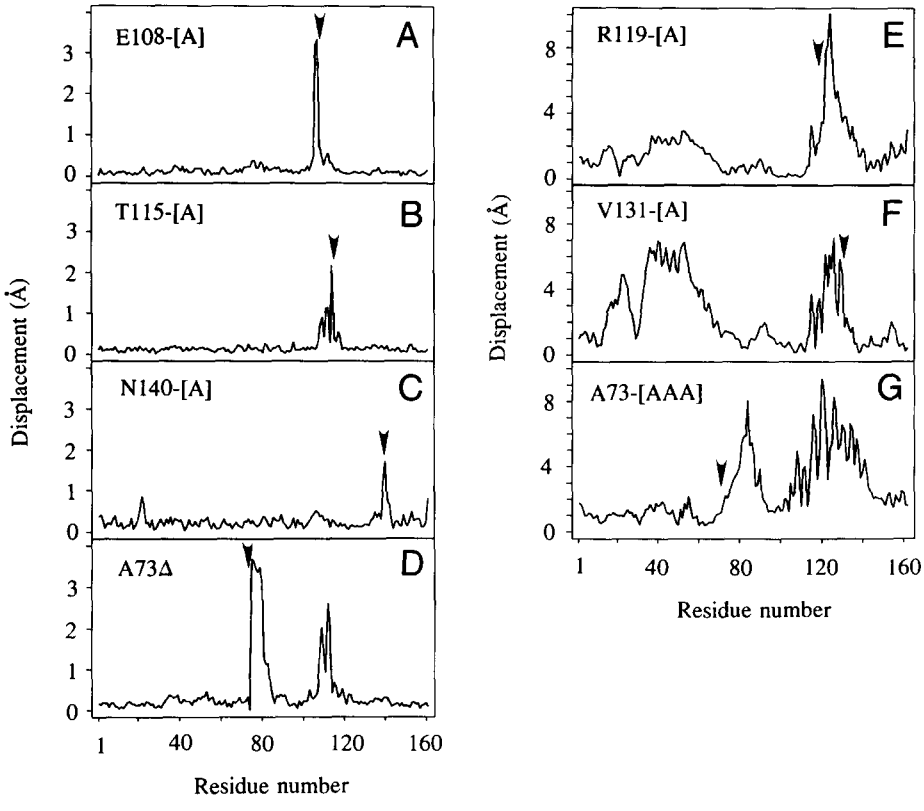

Fig. 4. "Shift plots" showing the changes in the backbone of insertion and deletion mutants relative to wild type. Arrowheads indicate the location of the insertion or deletion. **A:** E108-[A]. **B:** T115-[A]. **C:** N140-[A]. **D:** A73Δ. **E:** R119-[A]. **F:** V131-[A]. **G:** A73-[AAA]. In the first four panels and the last, the backbone of the mutant lysozyme was superimposed on that of wild type and the figure shows the discrepancy for each residue. In the case of R119-[A] and V131-[A], the superpositions of the mutant structure on wild type were based, respectively, on the backbone atoms of residues 80–123 and 80–162, i.e., on the general region that includes the insertion. In these two cases, the apparent changes in the amino-terminal part of the protein are due to a change in the hinge-bending angle between the amino- and the carboxy-terminal domains.

immediate vicinity of the substitution or insertion (e.g., Matthews, 1995). In three of the mutants described here, however, namely A73-[AAA], R119-[A], and V131-[A], the structural reorganization is decidedly nonlocal (Figs. 4, 11). Although these three insertions are separated in the amino acid sequence, they all cause a reorganization of residues 114–142, which form a "cap" over the carboxy-terminal domain (Fig. 1).

To visualize the consequences of this reorganization, a computer program was written to display the interactions that occur at the interface between the "cap" and the rest of the protein (I.R. Vetter, unpubl.). The results for the wild-type protein are shown in Figure 12. Although the interface itself is not flat, the results are projected onto a plane that is parallel to the overall interface region. Figure 12A summarizes the interaction surface contributed by the "cap" of the interface. The backbone shown in blue in the figure corresponds to residues 114–142. Figure 12B summarizes the interaction surface contributed by the opposite side of the interface, which is comprised of residues 83–112 and 144–154 (backbone drawn in blue). The green contours show the van der Waals separation across the interface. Higher contours correspond

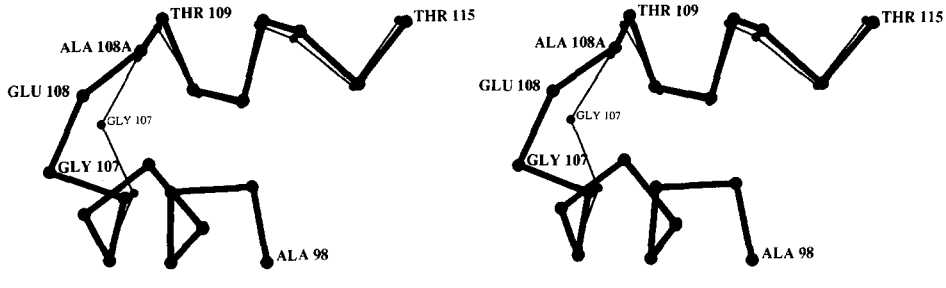

Fig. 5. Comparison of the backbone structure of mutant E108-[A] (thick bonds and larger labels) with wild-type lysozyme (thin bonds and smaller labels).

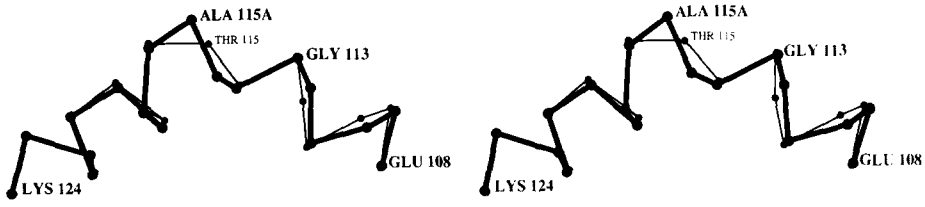

Fig. 6. Comparison of the backbone structure of mutant T115-[A] (thick bonds and larger labels) with wild-type lysozyme (thin bonds and smaller labels).

to closer approaches. Each contact between a pair of atoms that are within 4 Å is indicated by a black dot. The surrounding circle is drawn larger as the contact distance decreases. The numbers identify the residues that participate in these contacts. Hydrogen bonds are indicated by red spheres, with the size inversely related to the distance. Other contacts between polar atoms are indicated by green spheres. Salt bridges are indicated by purple spheres.

Figures 13, 14, and 15 show the interface analyses for the mutants A73-[AAA], R119-[A], and V131-[A], respectively. It is immediately apparent that the part of the interface that is well within the protein includes only nonpolar van der Waals contacts. On the other hand, the shape of the van der Waals interaction surface varies substantially from mutant to mutant. The polar interactions that do occur across the interface are primarily hydrogen bonds. These all occur at or close to the solvent-accessible boundary of the interface and, again, vary substantially from mutant to mutant. The overall impression is that the three insertion mutants A73-[AAA], R119-[A], and V131-[A] not only cause substantial reorganization within the carboxy-terminal domain, but also result in wholesale changes in the contact surface between the residue 114–142 "cap" and the rest of the protein.

As a control, Figure 16 shows the "cap" interface for mutant E108-[A]. In this case, the structural changes are relatively localized (Fig. 4A) and the "cap" interface interactions are seen to be quite similar to those of wild-type lysozyme.

Structural response to insertions

In terms of their stabilities, the insertion mutants constructed here are fairly similar to the lysozyme mutants reported previously (Heinz et al., 1993, 1994). In terms of structure, however, a more diverse set of responses was observed. With the possible exception of mutant I100-[A], for which no protein was obtained, the single alanine insertions in a variety of different helical locations destabilized the protein by 0.7–6.6 kcal/mol at pH 5.4. In agreement with results for staphylococcal nuclease (Sondek & Shortle, 1990, 1992), this corresponds to the range of destabilization observed for typical substitutions within the hydrophobic core and elsewhere (Eriksson et al., 1993; Heinz et al., 1993, 1994). Somewhat unexpectedly, the single amino acid deletions were no more destabilizing (3.1–5.4 kcal/mol) than the single amino acid insertions at pH 5.4 (Table 1). At pH 5.4, all of the mutants for which protein was obtained unfolded and refolded in an apparently two-state manner, indicative of well-defined three-dimensional structures for these proteins.

In three of the insertion mutants for which structures were obtained (E108-[A], T115-[A], and N140-[A]), the site of the mutation is either at the end, or very close to the end, of an α-helix (Fig. 2). In addition, the insertion site is in a context where one helix ends and another begins almost immediately. This is a common occurrence in the carboxy-terminal domain of T4 lysozyme because of the 94 C-terminal residues, only 16 are not in α-helices. As is apparent in Figure 4, in these three cases, the structural changes are quite localized. [In N140-[A], the residues around Glu 22 are influenced slightly as well (<1 Å).] In each case, the close proximity of the insertion site to the end of the helix readily permits translocation of the peptide chain into the loop at the end of the helix with modest structural change and modest loss of stability (1.8–2.7 kcal/mol).

In three additional mutants for which structures were obtained (A73-[AAA], R119-[A] and V131-[A]), the site of insertion is three or more residues from the end of the helix. Here again, translocation was observed in each case (Fig. 2), but associated with substantial changes in the backbone structure (4–7 Å) in the C-terminal domain, especially within α-helices 115–123 and 126–134 (Fig. 4). The deletion mutant A73Δ also resulted in a translocation (or shift in register) (Fig. 2), with more modest, but still substantial (2–3 Å) changes in backbone conformation. In three of

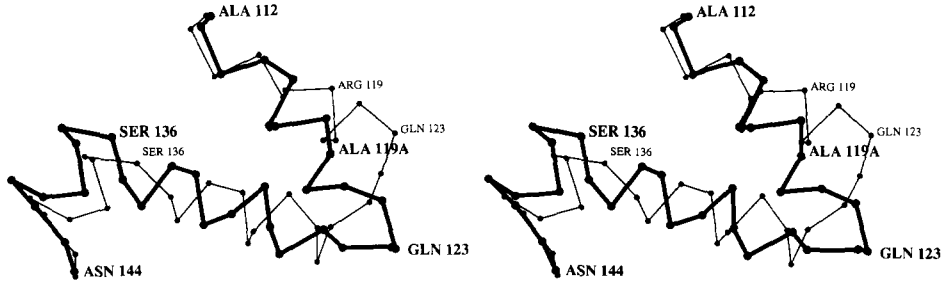

Fig. 7. Comparison of the backbone structure of mutant R119-[A] (thick bonds and larger labels) with wild-type lysozyme (thin bonds and smaller labels).

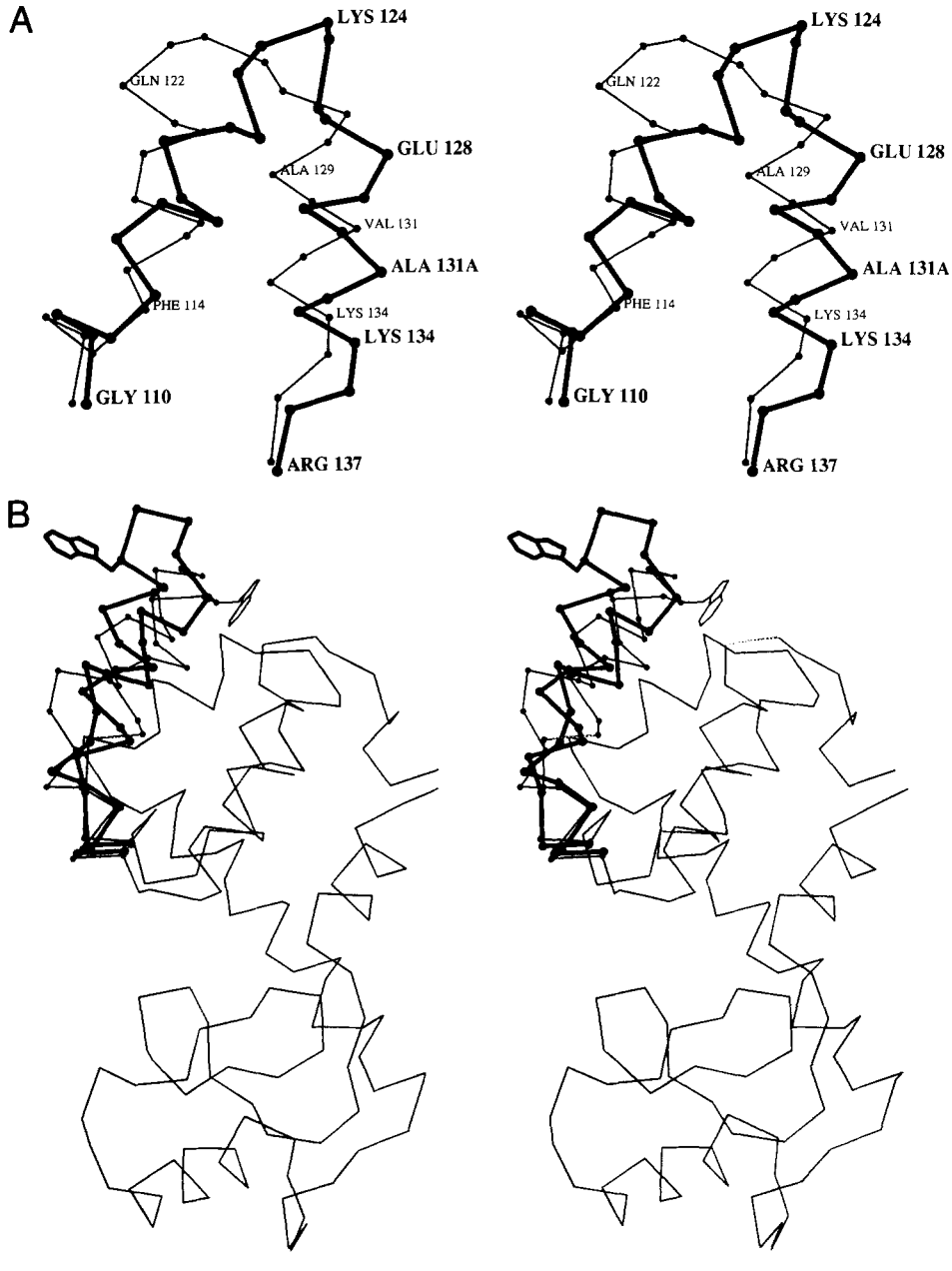

Fig. 8. A: Stereo drawing comparing the backbone of wild-type lysozyme (thin lines and smaller labels) with that of the insertion mutant V131-[A] (thick lines and larger labels). **B:** Comparison of the overall structure of wild-type lysozyme (thin lines) with V131-[A] (thick bonds for residues 108–138) in the vicinity of the insertion. There is a translocation of five amino acids including the movement of Trp 126 from a partially buried pocket on the surface of the protein to a solvent-exposed site, where it participates in a crystal contact.

these four cases, the translocation resulted in alterations of the amino acids at five sites within the helix in which the insertion or deletion occurred (Fig. 2). Notwithstanding these unusually large structural perturbations, the destabilization resulting from these translocating mutants ranged from moderate (2.6 kcal/mol) to large (5.0 kcal/mol), with an average value of 3.6 kcal/mol.

In the prior study of insertion mutations within α-helix 39–50, the length of the helix remained invariant (Heinz et al., 1993, 1994). This suggested that the beginning and end of the helix were determined not only by the amino acid sequence, but also by interactions between the residues within the helix and the rest of the protein. In contrast, the present data include several examples where the length of the helix changes (Fig. 2). Most strikingly, in mutant A73-[AAA], the insertion of three alanines within the interdomain helix increases the length by three amino acids and causes the neighboring helices 108–113 and 115–123 to reorganize into a single helix, 108–121 (Fig. 2).

In terms of the overall energetic and structural response of the protein to insertions within helices, it appears that the mutants can be considered in three classes: (1) translocation mutants; (2) mutants with looping-out at the ends of helices; and (3) mutants with looping-out in the body of the helix. We briefly discuss these categories below.

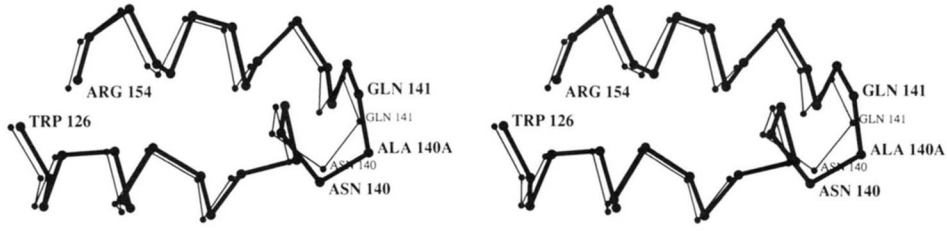

Fig. 9. Comparison of the backbone structure of mutant N140-[A] (thick bonds and larger labels) with wild-type lysozyme (thin bonds and smaller labels).

Translocation mutants

"Translocation," or "register shift," mutants comprise by far the largest class of insertion mutants for which we have structural data. As summarized in Table 4, this is also the class for which the destabilization is, on average, least. It strongly suggests that this is the response of choice. If the amino acid sequence allows translocation either toward the amino terminus or toward the carboxy terminus of the helix without introducing highly unfavorable interactions, then the polypeptide chain will, in fact, undergo translocation. The critical determinants are the buried or largely buried sites. Translocations that would result in the placement of highly polar or charged residues at buried hydrophobic sites are not observed. This is consistent with the observation that such substitutions are very destabilizing (e.g., the substitutions Met 102 → Lys and Leu 133 → Asp within the core of T4 lysozyme destabilize the protein by ~7 and ~6 kcal/mol, respectively, at pH 5–6 (Dao-pin et al., 1991). Translocations that involve the substitution of smaller

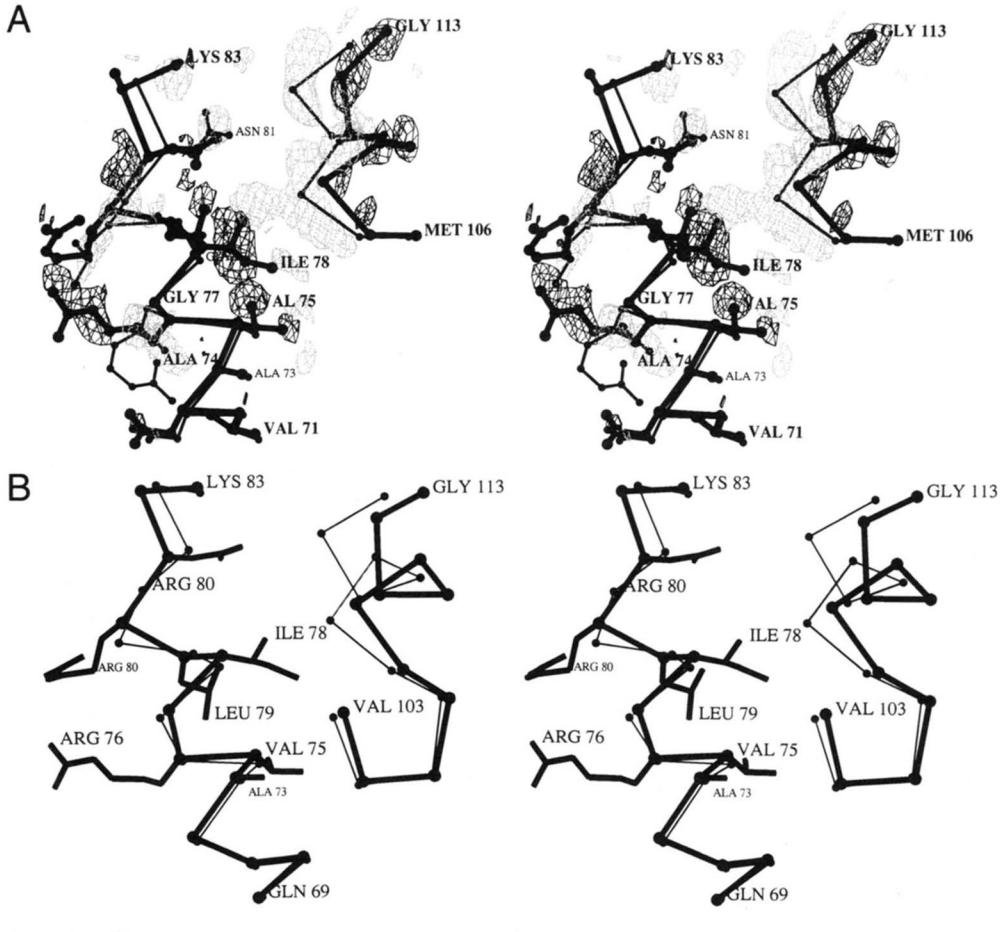

Fig. 10. A: Initial difference electron density map for mutant A73Δ. ($F_{mut} - F_{WT*}$) where F_{mut} and F_{WT*} are the observed structure amplitudes for the mutant and WT* crystals. Phases from the refined structure of WT*. Coefficients are at $+3.5\sigma$ (darker lines) and -3.5σ (lighter lines) where σ is the RMS density throughout the unit cell. The wild-type structure is shown in thin lines with smaller labels and the refined mutant structure in thicker lines and larger labels. **B:** Stereo drawing comparing the backbone structure of the deletion mutant A73Δ (thick bonds) with that of wild type (thin bonds). The deletion of Ala 73 causes a translocation of four residues within the helix.

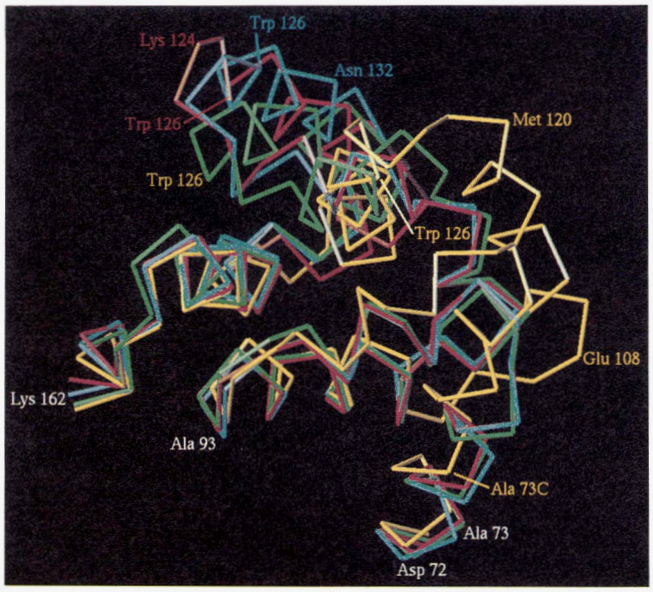

Fig. 11. Color figure comparing the backbone structures within the carboxy-terminal domain of WT* (green), R119-[A] (red), V131-[A] (cyan), and A73-[AAA] (yellow). For mutants R119-[A] and V131-[A], the superposition is based on residues 70–112 and 144–162; for A73-[AAA], residues 1–70 and 144–162 were used. In each case, the residues chosen are those that are best conserved in the respective structures (cf. Fig. 4). Residues numbered in white occupy essentially the same positions in all structures. Residues such as Trp 126, however, move substantially.

or less polar side chains at internal sites (e.g., Leu → Ala, Ala → Ser) are, however, observed (Heinz et al., 1994). This is consistent with the observations that such replacements need only be modestly destabilizing [Leu → Ala replacements average 3.5 kcal/mol for proteins in general (Pace, 1992) and Ala → Ser replacements at internal sites in T4 lysozyme average 1.9 kcal/mol (Blaber et al., 1993)]. The nature of the residues that are translocated to solvent-exposed sites on the surface of the helix seems to be of little importance, consistent with the finding that substitutions at solvent-exposed sites in general have little effect on protein stability (Alber et al., 1987).

Mutants with looping-out at the ends of helices

In general, "looping-out," or "bulging," is destabilizing and, at least in terms of the structures that we have been able to determine, occurs relatively infrequently. In two cases, however, namely N40-[A] and K48-[HP] (Heinz et al., 1993, 1994), an insertion close to the end of an α-helix resulted in looping-out at the end of the helix such that the terminal residue retained essentially the same position as in the wild type. In such cases the destabilization was modest (2.8 and 2.4 kcal/mol; Table 4) and we choose to consider this as a distinct class.

Consistent with this classification, Kavanaugh et al. (1993) found that an insertion in the first turn of an α-helix of hemoglobin generated a $3_{10} \rightarrow \alpha$ bulge. Also, Keefe et al. (1993) observed that an insertion within the first turn of an α-helix in staphylococcal nuclease generated a bulge or "α-aneurism."

Mutants with looping-out in the body of a helix

We have only one structurally characterized example of this type, namely S44-[A], which is the most destabilizing single-residue insertion within α-helix 39–50 (Heinz et al., 1993, 1994). The looping-out of the polypeptide chain within an existing α-helix would be expected to be substantially destabilizing. Also, such structures are rarely observed in known protein structures (Pascarella & Argos, 1992).

Although we do not have structural information, the thermodynamic data suggest that looping-out may occur for the mutants with insertions following Arg 148 (Table 1). If translocation were to occur, it would require the displacement of seven residues toward the carboxy terminus or six residues toward the amino terminus (Fig. 2). Translocations involving different displacements, i.e., by one, two, three, or four residues, if they were to occur, would be expected to result in different changes in protein stability. In contrast, however, the four mutants R148-[A], R148-[AA], R148-[AAA], and R148-[AAAA] all destabilize the protein to about the same degree. Also, the destabilization is large, suggestive of looping-out rather than translocation (cf. Table 4). Furthermore, the replacement of the inserted alanines with different types of residues, as in R148-[D], R148-[DS], and R148-[TT], leaves the stability largely the same (Table 1). This is consistent with these residues being looped-out and exposed to solvent. It is not consistent with these residues being integral with the helix. For example, if mutant R148-[D] were to cause translocation toward the carboxy terminus, it would require that the inserted aspartate be located in place of the buried Val 149 (Fig. 2); if translocation were to occur toward the amino terminus, it would require that Lys 147 be put in place of the buried Ala 146. Either seems highly unlikely.

The above considerations give some insight into the likely response of a protein to insertions and, to some degree, deletions as well. Wherever possible, a protein will accommodate an insertion or deletion within a helix by a translocation. The sites that are buried are critical in determining whether a translocation will, in fact, occur. If a putative translocation preserves the general nature and size of the amino acids at the buried sites within the helix, then it can be expected to occur. The translocation can include substitutions such as Ala → Ser and Leu → Thr, as well as more conservative replacements. Only in the case of severe mismatches, such as the substitution of a charged or highly polar residue for a nonpolar residue, will translocation not occur. Within the body of an α-helix, looping-out is very rare and, when it occurs, is very destabilizing. Close to the ends of helices, however, the polypeptide backbone has more freedom to adopt alternative changes in conformation, and in such cases looping-out can occur with relatively modest loss of stability.

Energetic consequences of insertions

The stabilities of the variants were determined at pH 5.4 and pH 3.0. The former is close to the pH of maximum stability for most mutant lysozymes, but is technically demanding because of a loss of reversibility for the least stable variants. The lower pH has been used for a number of prior studies (e.g., Eriksson et al., 1993). It has the advantages that the unfolded protein is exposed to reduced temperatures, and that unfolding shows better reversibility than at higher pH values, presumably because of an increased net charge on the molecule. The $\Delta\Delta G$ values of the mutant lysozymes at the two values of pH are generally comparable (Table 2). For the most part, no contribution to $\Delta\Delta G$ from electrostatic effects is seen for these mutants in these solvents. Some departures from this rule are apparent for mutants such as A73-[R], which introduces a new charged residue, and for some of the insertions after Arg 148.

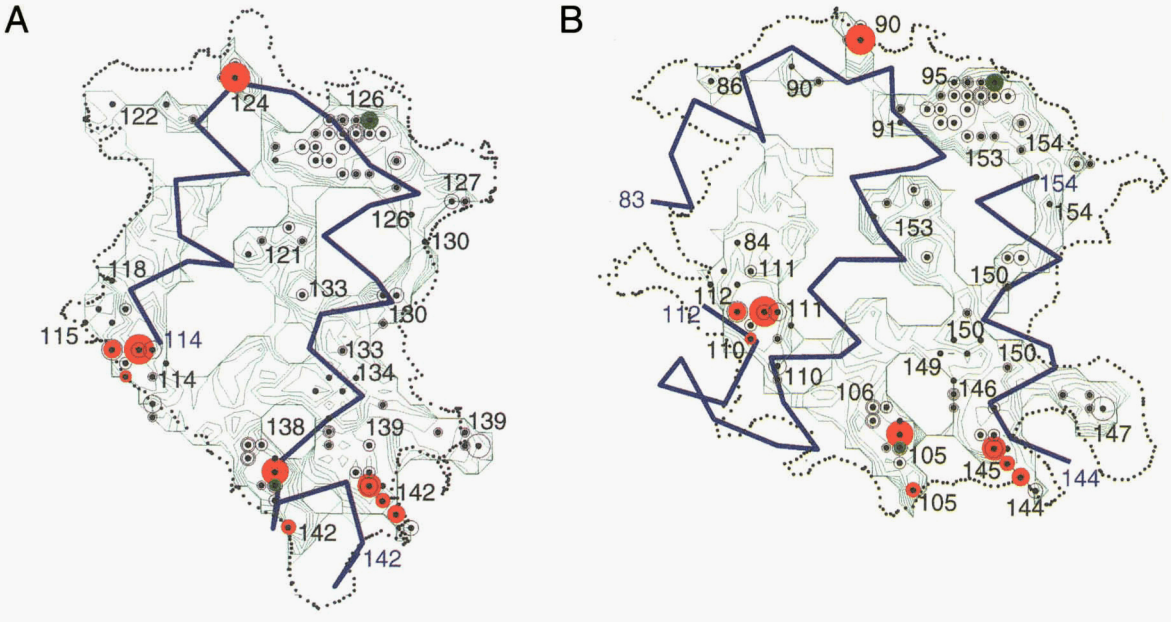

Fig. 12. Diagrammatic analysis of the interface between the "cap" formed by residues 114–142 and the remainder of the carboxy-terminal domain in WT* lysozyme. The representation is made up of two panels. **A:** Summarizes all the interactions that are generated by atoms that form one side of the interface (residues 114–142). **B:** Shows the interactions that are formed by all the atoms on the other side of the interface. Black dots show van der Waals contacts between pairs of atoms whose centers are within 4.0 Å. Residues that participate in these contacts are numbered. Atoms that participate in hydrogen bonds across the interface are indicated by red spheres, salt bridges are indicated by purple spheres, and other nonpolar contacts are shown by green spheres. Green contour lines show the separation across the interface between the van der Waals surfaces with contours corresponding to decreasing separations in steps of 0.1 Å. In general, the black dots (nonpolar contacts) and colored spheres (polar contacts) tend to occur at or near the minima of the green contours because these correspond to the closest approaches across the interface. The dotted black line defines the region of the interface from which solvent is excluded (the line is defined as the path of contact that is traced out by a water molecule at the surface of the protein that just touches both sides of the interface). The information is projected onto a plane that is essentially parallel to the contact interface, the same plane being used for Figures 12–16. **A:** Interaction surface formed by the "cap" residues 114–142. The backbone trace of these residues is shown in blue. **B:** Interaction surface in WT* lysozyme that opposes the surface shown in A, composed of residues 83–112 plus 144–154 (backbone drawn in blue).

However, because these cases occur for proteins that might depart from two-state melting in at least one of the solvents, the uncertainty associated with their ΔG values in that solvent is much greater. Except for these cases, enthalpies of unfolding are roughly comparable with wild type, consistent with the maintenance of a well-defined three-dimensional structure.

It has been noted previously that regions of a protein that are more mobile tend to be less susceptible to destabilization by point mutations than are regions that are more rigid (Alber et al., 1987). The same trend is seen with the insertion mutants. Figure 17 shows the destabilization for the 22 single-alanine insertion mutations in T4 lysozyme plotted as a function of the crystallographic thermal factor of the main-chain atoms in the wild-type structure at the site of the insertion. The plot shows a very broad scatter, which is not surprising because the consequences of a given insertion can vary substantially depending on the compatibility of the surrounding amino acid sequence with translocation or other types of structural response. Nevertheless, it does appear that the mobility of the backbone is inversely related to the maximum destabilization that can occur at a given site. Consistent with this hypothesis, the insertion mutant for which protein was not obtained, suggesting substantial destabilization, occurred in a region of low mobility (Fig. 17).

Figure 18 shows the destabilization resulting from the single alanine insertions in T4 lysozyme and six additional single alanine insertions in staphylococcal nuclease (Sondek & Shortle, 1990) plotted as a function of the distance from the end of the helix. The figure suggests that the most destabilizing insertions tend to occur well within the body of the helix, although there is no clear correlation with the distance from the end. Pascarella and Argos (1992) found that the vast majority of recognizable insertions in known protein structures appeared to occur outside helices, and, in the rare cases that they occurred within helices, were within four residues of one of the helix termini.

It was noted above that the part of the carboxy-terminal domain that spans residues 114–142 acts, to some degree, as a flexible "cap" and undergoes large structural changes in the insertion mutants R119-[A], V131-[A], and A73-[AAA] (Figs. 4, 11). In the WT* structure, the backbone atoms within this region have an average value 20.8 $Å^2$, essentially identical with wild type (21.2 $Å^2$). In each of the three mutant structures, the backbone thermal factors for the atoms increase by 11–13 $Å^2$ relative to the rest of the protein, indicating a "loosening" or increased mobility of this region of the molecule. This could also explain why the enthalpies of unfolding of these three molecules (69–104 kcal/mol at pH 5.4) are lower than that of wild type (131 kcal/mol) (Tables 1, 2).

We assume that the types of structural changes seen here in response to insertions and deletions illustrate changes that can occur during protein evolution. At the same time, however, it should be noted that some of the mutations described here are quite destabilizing and so might be selected against via either reversion or

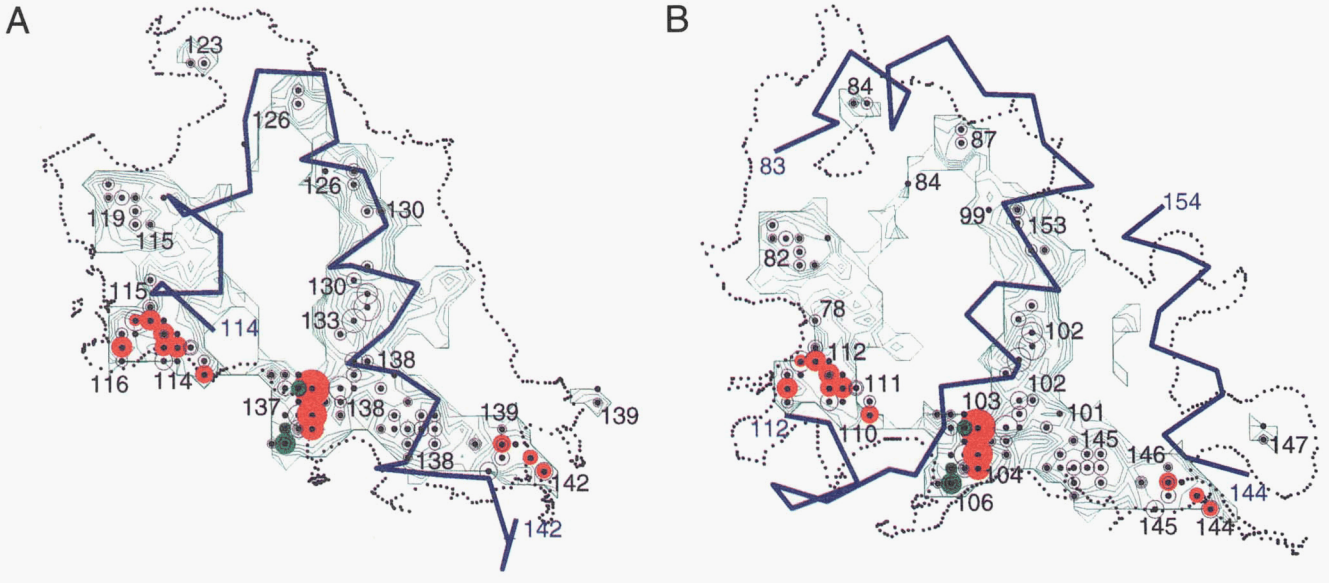

Fig. 13. A: "Cap" interaction surface in mutant A73-[AAA]. **B:** Opposing interface. All conventions are as in Figure 12.

compensating substitutions. Lesk and Chothia (1986) have noted that protein structures tend to respond to changes in the sequence of core residues by rigid-body movement of α-helices. For proteins of 20% sequence homology, these shifts amount to 3–5 Å or so. This is the same order of magnitude of the changes seen here, although, in this case, the adjustments are triggered by a highly localized alteration, rather than distributed changes. In contrast to the idea of rigid-body shifts of existing secondary structure elements, the present study includes an example (in A73-[AAA]) where two α-helices that are inclined at 105 degrees meld to form a single, continuous helix. Presumably changes such as this occur during evolution as well, although they may be difficult to identify in contemporary protein structures because the amino acid sequence might also have changed to such an extent that the lineage of a given α-helix is uncertain.

Materials and methods

Mutagenesis and protein purification

Mutants were constructed using the method of Kunkel (1985). The alanine mutants with insertions after position 148 were mutated further using oligonucleotides with degenerate codons (NNN or NNG/C) at the site of the alanine. The revertants were screened for increased stability using an in vivo halo assay (Streisinger et al., 1961). Mutants were cloned and the sequences were confirmed by DNA sequence analysis of the mutant lysozyme gene. Where possible, the mutant lysozymes were purified as described previously (Poteete et al., 1991). Mutants R96-[A], R148-[A], I150-[A], R148-[A], R148-[AA], R148-[AAA], R148-[AAAA], R148-[D], R148-[DS], R148-[VP], and R8Δ gave inclusion bodies, in which case the standard purification was adapted as follows.

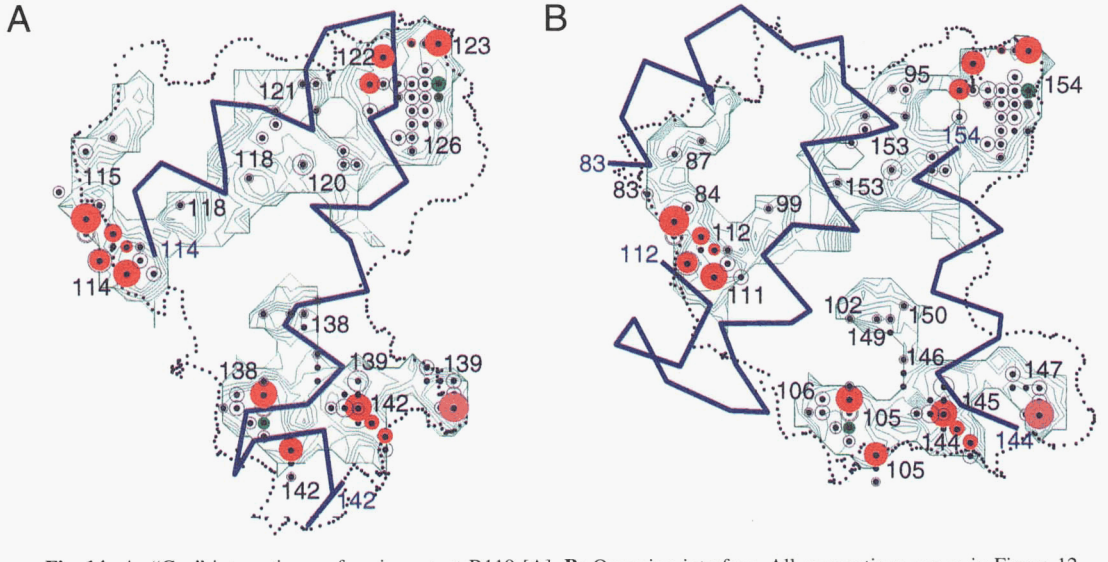

Fig. 14. A: "Cap" interaction surface in mutant R119-[A]. **B:** Opposing interface. All conventions are as in Figure 12.

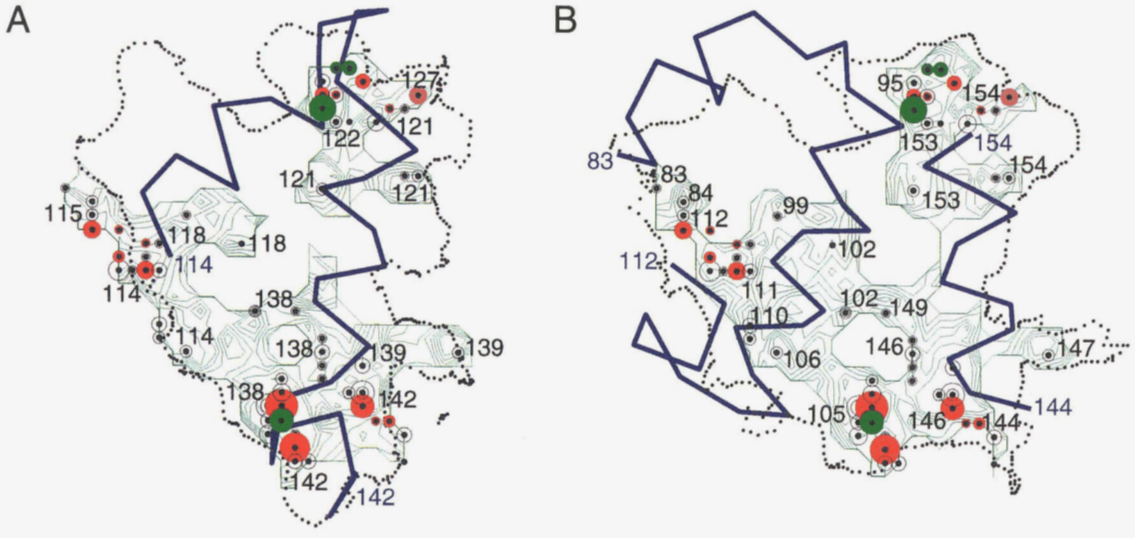

Fig. 15. A: "Cap" interaction surface in mutant V131-[A]. **B:** Opposing interface. All conventions are as in Figure 12.

The cells obtained from a 3-L culture are passed two or three times through a French press and then spun for 35 min at 18,000 rpm. Any protein that may be in the supernatant is purified in the standard way. The pellet is resuspended in 50 mM Tris-HCl, 10 mM sodium EDTA, 50 mM NaCl, 1 mM phenyl methane sulfonyl fluoride, 2.5 mM benzamide, 0.1 mM DTT, pH 8.0, to which is also added an amount equal to one-tenth the volume of 10% Triton X-100 in 50 mM Tris, 10 mM sodium EDTA, pH 8.0. The suspension is stirred at 4 °C for 2 h and centrifuged at 15,000 rpm for 20 min. Ten volumes of 2.5% octyl-β-D-glucopyranoside in 50 mM Tris, 10 mM sodium EDTA, pH 8.0, per apparent volume of pellet is added and the suspension stirred at 4 °C for 2 h. Following centrifugation at 15,000 rpm for 20 min, the pellet is washed twice with double-deionized water. The pellet is resuspended in freshly made 4 M urea solution (unbuffered, pH 6–7), to which is added sufficient 20 mM glycine, 20 mM phosphoric acid to lower the pH to 3–3.5, then mixed well for a few minutes and centrifuged for 20 min at 18,000 rpm. The supernatant, containing the protein, is immediately dialyzed against 50 mM citric acid:NaOH buffer, 10% glycerol, pH 5.5, for 1.5 h and then against 50 mM citrate, 10% glycerol, pH 3.0. The solution is centrifuged at 15,000 rpm for 15 min. This procedure provided at least 40 mg of protein, estimated by SDS-PAGE to be approximately ~95% pure. If necessary, gel filtration FPLC was used for further purification.

Thermal stability

The thermal stabilities were determined in 0.10 M NaCl, 1.4 mM acetic acid, 8.6 mM sodium acetate, pH 5.42 (at 22 °C), and in 0.025 M KCl, 2.95 mM H_3PO_4, 17.0 mM KH_2PO_4, pH 3.01 (at

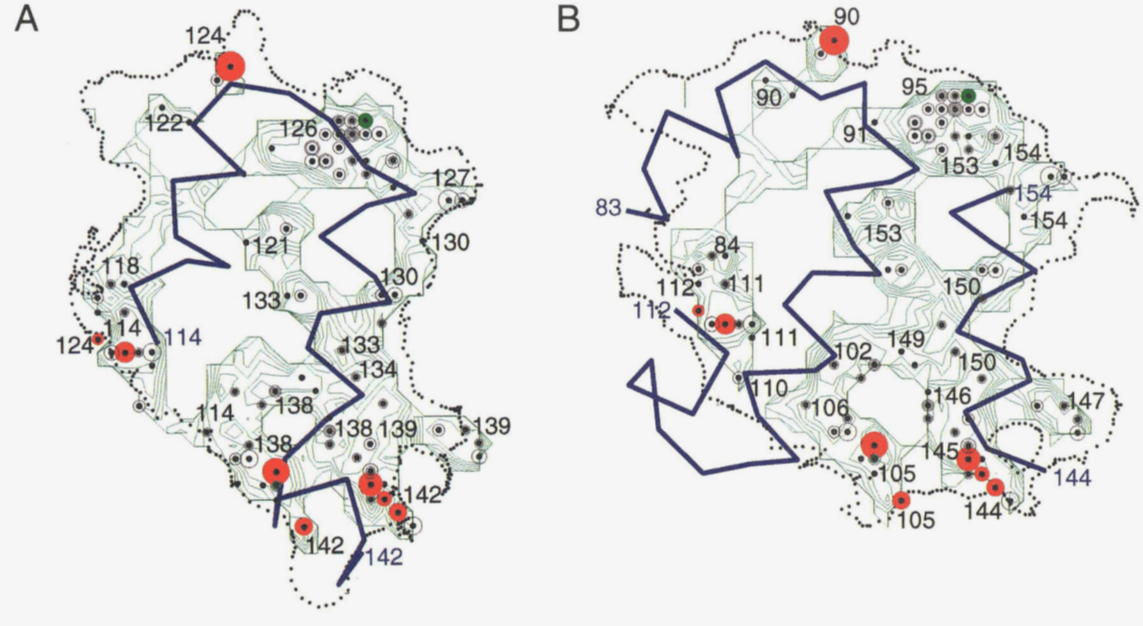

Fig. 16. A: "Cap" interaction surface in mutant E108-[A]. **B:** Opposing interface. All conventions are as in Figure 12.

Table 4. *Stabilities of "translocation" and "looping-out" mutants in T4 lysozyme at pH 5.4*

Translocation mutants		Mutants with looping-out at the ends of α-helices		Mutants with looping-out within an α-helix	
Mutant	ΔΔG (kcal/mol)	Mutant	ΔΔG (kcal/mol)	Mutant	ΔΔG (kcal/mol)
N40-[AA]	−1.1[a]	N40-[A]	−2.8[a]	S44-[A]	−3.8[a]
N40-[AAA]	−1.7[a]	K48-[HP]	−2.4[a]	R148-[A][c]	−4.6
S44-[AA]	−3.4[a]	Average	−2.6	R148-[AA][c]	−5.4
S44-[AAA]	−2.2[a]			R148-[AAA][c]	−5.0
A73-[AAA]	−5.0			R148-[AAAA][c]	−5.6
E108-[A]	−2.7			Average	−4.9 ± 0.7
T115-[A][b]	−1.8				
R119-[A]	−2.6				
V131-[A][b]	−2.9				
N140-[A][b]	−2.1				
Average	−2.5 ± 1.0				

[a]From Heinz et al. (1993, 1994).
[b]These mutants are close to the ends of helices. They are classified as "translocation" mutants in that the effect of the insertion is propagated into a loop at the end of the helix. The end of the helix itself, however, is also perturbed.
[c]We do not have structural data for these mutants. The evidence that looping-out occurs is indirect (see text).

22 °C). The CD signal at 223 nm was monitored as a function of temperature (Eriksson et al., 1993) and converted to a melting temperature, T_m, and an enthalpy of unfolding at the T_m, $\Delta H(T_m)$, via two-state van't Hoff analysis (Zhang et al., 1995a). ΔG values were computed using the constant ΔC_p model at 53 °C (where ΔC_p was taken as 2.5 kcal/mol-deg) for pH 5.4 buffer and at 42 °C (where ΔC_p was taken as 1.8 kcal/mol-deg) for pH 3.0 buffer (Zhang et al., 1995a). $\Delta\Delta G$ values were computed as ΔG(mutant) − ΔG(WT*), as previously, such that a negative value indicates a protein that is less stable than the reference protein.

The pH 5.4 data listed in Tables 1 and 2 result from the averaging of at least three independent melts, usually more. The pH 3 data were usually taken three times, but for several proteins only twice. The standard deviation of the mean in T_m values near the T_m of WT* was ±0.15 °C. However, this number increased to approximately ±0.35 °C as the value of ΔH decreased. The standard deviation of ΔH was approximately ±4.5 kcal/mol, which is ±3% at the ΔH of WT*, but increases to 18% for a ΔH of 25 kcal/mol, approximately the lowest ΔH that can be considered two-state in these data sets. For these reasons, and because of the constant ΔC_p approximation, we estimate the uncertainty in $\Delta\Delta G$ to be ±0.15 kcal/mol, but to rise to as much as ±1.0 kcal/mol for the least stable constructs.

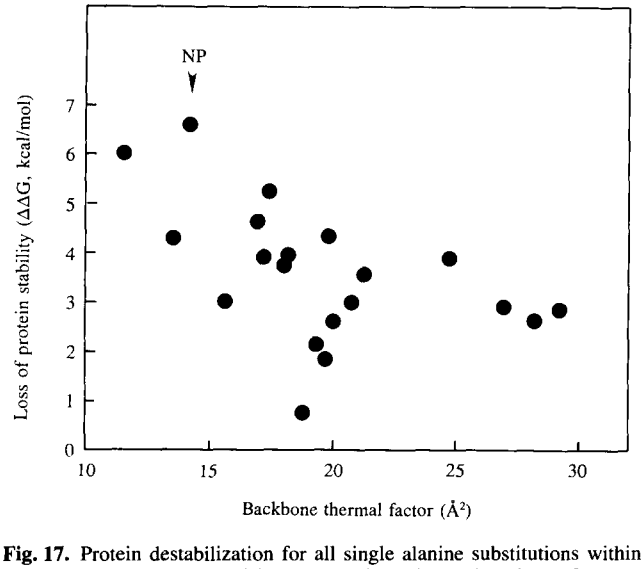

Fig. 17. Protein destabilization for all single alanine substitutions within α-helices in T4 lysozyme ($\Delta\Delta G$ at pH 5.4; Table 2) plotted as a function of the backbone thermal factor at the site of insertion. The thermal factor plotted is the average value for main-chain atoms in the wild-type structure (Weaver & Matthews, 1987) for the residue preceding and the residue following the insertion site. One mutant, indicated NP, gave no protein (Table 1). For this mutant, the value of $\Delta\Delta G$ is unknown, although this is, presumably, one of the least stable proteins.

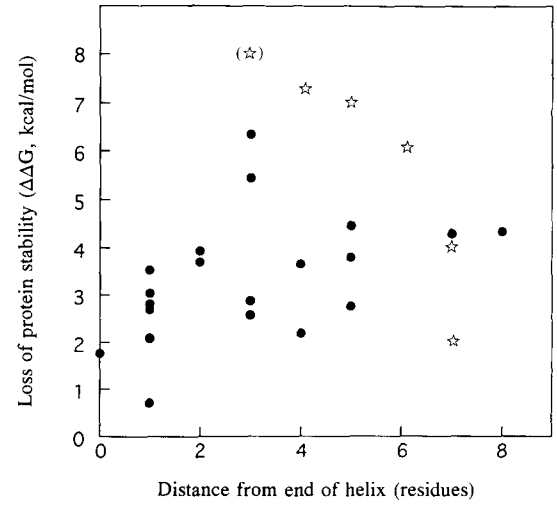

Fig. 18. Solid circles show the destabilization associated with all single-alanine insertions within α-helices in T4 lysozyme ($\Delta\Delta G$ at pH 5.4; Table 1) plotted as a function of the distance of the site of insertion, in residues, from the closest end of the helix. Stars correspond to single alanine insertion mutants in staphylococcal nuclease (Sondek & Shortle, 1990).

Reversibility of thermal unfolding was judged by restoration of CD signal on cooling after completion of data collection. Reversibility was highest at pH 3, usually in excess of 95%. Reversibility was in excess of 85% in pH 5.4 buffers, except for the least stable proteins, where it dropped below 20%.

Crystallographic analysis

Crystallization using the hanging drop method was attempted under "standard" conditions used for WT* lysozyme, namely 2–2.6 M phosphate, 0.25 M NaCl, and ~50 mM hydroxyethyldisulfide (HEDS) (i.e., oxidized β-mercaptoethanol), pH 6.7–7.1 (Eriksson et al., 1993). A variety of other crystallization conditions was explored as well (cf. Zhang et al., 1995b).

X-ray diffraction data were collected using a San Diego Multiwire Systems area detector (Hamlin, 1985; Zhang & Matthews, 1993). Structures of mutants with crystals isomorphous to WT* space group ($P3_221$) were determined and refined by standard methods (Jones, 1982; Tronrud et al., 1987; Tronrud, 1992). Residues that appeared to undergo significant change in structure following an insertion or deletion were deleted from the protein model. Following initial refinement, difference maps were calculated and used to place these parts of the structure. Mutant V131-[A], which crystallized in space group $P2_12_12_1$, was solved by molecular replacement using the programs POLARRFN (Crowther, 1972), MERLOT (Fitzgerald, 1988), and TFFC (Tickle, 1985), and refined as described above. Mutants R119-[A] and A73-[AAA] crystallized, respectively, in space groups $P6_5$ and $P6_222$ and were solved using AMoRe (Navaza, 1994).

Acknowledgments

We are most grateful to Joan Wozniak for help in purifying and crystallizing mutant lysozymes; Eric Anderson and Nadine Gassner for advice on purification from inclusion bodies; Joel Lindstrom for assistance with thermodynamic analyses; Larry Weaver, Nadine Gassner, and X.-J. Zhang for help with the crystallographic analyses; and Lynn Markham for help in constructing the insertion mutants in the interdomain helix. We thank S.J. Hubbard and J.M. Thornton for the computer program "NACCESS." This work was supported in part by a Deutsche Forschungsgemeinschaft fellowship to I.R.V. and an 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.

Bell JA, Wilson KP, Zhang XJ, Faber HR, Nicholson H, Matthews BW. 1991. Comparison of the crystal structure of bacteriophage T4 lysozyme at low, medium and high ionic strengths. *Proteins Struct Funct Genet 10*:10–21.

Blaber M, Lindstrom JD, Gassner N, Xu J, Heinz DW, Matthews BW. 1993. Energetic cost and structural consequences of burying a hydroxyl group within the core of a protein determined from Ala → Ser and Val → Thr substitutions in T4 lysozyme. *Biochemistry 32*:11363–11373.

Crowther RA. 1972. The fast rotation function. In: Rossmann MG, ed. *The molecular replacement method*. New York: Gordon and Breach. pp 173–185.

Dao-pin S, Anderson DE, Baase WA, Dahlquist FW, Matthews BW. 1991. Structural and thermodynamic consequences of burying a charged residue within the hydrophobic core of T4 lysozyme. *Biochemistry 30*:11521–11529.

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.

Fitzgerald PMD. 1988. MERLOT, an integrated package of computer programs for the determination of crystal structures by molecular replacement. *J Appl Crystallogr A 21*:273–278.

Hamlin R. 1985. Multiwire area X-ray diffractometers. *Methods Enzymol 114*:416–452.

Heinz DW, Baase WA, Dahlquist FW, Matthews BW. 1993. How amino-acid insertions are allowed in an α-helix of T4 lysozyme. *Nature 361*:561–564.

Heinz DW, Baase WA, Zhang XJ, 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.

Jones TA. 1982. FRODO: A graphics fitting program for macromolecules. In: Sayre D, ed. *Crystallographic computing*. Oxford: Oxford University Press. pp 303–317.

Kavanaugh JS, Moo-Penn WF, Arnone A. 1993. Accommodation of insertions in helices: The mutation in hemoglobin catonsville (Pro 37α-Glu-Thr 38α) generates a $3_{10} \rightarrow \alpha$ bulge. *Biochemistry 32*:2509–2513.

Keefe LJ, Quirk S, Gittis A, Sondek J, Lattman EE. 1994. Accommodation of insertion mutations on the surface and in the interior of staphylococcal nuclease. *Protein Sci 3*:391–401.

Keefe LJ, Sondek J, Shortle D, Lattman EE. 1993. The α aneurism: A structural motif revealed in an insertion mutant of staphylococcal nuclease. *Proc Natl Acad Sci USA 90*:3275–3279.

Kunkel TA. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc Natl Acad Sci USA 82*:488–492.

Lesk AM, Chothia C. 1986. The response of protein structures to amino-acid sequence changes. *Phil Trans R Soc (London) A317*:345–356.

Marti T, Otto H, Rösselet SJ, Heyn MP, Khorana HG. 1992. Consequences of amino acid insertions and/or deletions in transmembrane helix C of bacteriorhodopsin. *Proc Natl Acad Sci USA 89*:1219–1223.

Matsumura M, Matthews BW. 1989. Control of enzyme activity by an engineered disulfide bond. *Science 243*:792–794.

Matthews BW. 1995. Studies on protein stability with T4 lysozyme. In: Eisenberg D, Richards FM, Edsall JT, Anfinsen CB, eds. *Advances in protein chemistry on "protein stability."* New York: Academic Press. pp 249–278.

Navaza J. 1994. AMoRe: An automated package for molecular replacement. *Acta Crystallogr A 50*:157–163.

Nicholson H, Anderson DE, Dao-pin S, Matthews BW. 1991. Analysis of the interaction between charged side-chains and the α-helix dipole using designed thermostable mutants of phage T4 lysozyme. *Biochemistry 30*:9816–9828.

Pace CN. 1992. Contribution of the hydrophobic effect to globular protein stability. *J Mol Biol 226*:29–35.

Pascarella S, Argos P. 1992. Analysis of insertions/deletions in protein structures. *J Mol Biol 224*:461–471.

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.

Sharff AJ, Rodseth LE, Szmelcman S, Hofnung M, Quiocho FA. 1995. Refined structures of two insertion/deletion mutants probe function of the maltodextrin binding protein. *J Mol Biol 246*:8–13.

Sondek J, Shortle D. 1990. Accommodation of single amino acid insertions by the native state of staphylococcal nuclease. *Proteins Struct Funct Genet 7*:299–305.

Sondek J, Shortle D. 1992. Structural and energetic differences between insertions and substitutions in staphylococcal nuclease. *Proteins Struct Funct Genet 13*:132–140.

Streisinger G, Mukai F, Dreyer WJ, Miller B, Horiuchi S. 1961. Mutations affecting the lysozyme of phage T4. *Cold Spring Harbor Symp Quant Biol 26*:25–30.

Tickle IJ. 1985. In: Machin PA, ed. *Molecular replacement. Proceedings of the Daresbury Study Weekend February 1985*. SERC Daresbury Laboratory. pp 22–26.

Tronrud DE. 1992. Conjugate-direction minimization: An improved method for the refinement of macromolecules. *Acta Crystallogr A 48*:912–916.

Tronrud DE, Ten Eyck LF, Matthews BW. 1987. An efficient general-purpose least-squares refinement program for macromolecular structures. *Acta Crystallogr A 43*:489–503.

Weaver LH, Matthews BW. 1987. Structure of bacteriophage T4 lysozyme refined at 1.7 Å resolution. *J Mol Biol 193*:189–199.

Zhang XJ, Baase WA, Shoichet BK, Wilson KP, Matthews BW. 1995a. Enhancement of protein stability by the combination of point mutations in T4 lysozyme is additive. *Protein Eng 8*:1017–1022.

Zhang XJ, Matthews BW. 1993. STRAT: A program to optimize X-ray data collection on an area detector system. *J Appl Crystallogr 26*:457–462.

Zhang XJ, Wozniak JA, Matthews BW. 1995b. Protein flexibility and adaptability seen in 25 crystals forms of T4 lysozyme. *J Mol Biol 250*:527–552.