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

Permalink
https://escholarship.org/uc/item/56z4m5zm

Journal
Protein Science, 7(1)

ISSN
0961-8368

Authors
Xu, J
Baase, WA
Baldwin, E
et al.

Publication Date
1998

DOI
10.1002/pro.5560070117

Peer reviewed
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 \( \leq 0.3 \) Å); in other cases mainchain atoms moved up to 2 Å. In the case of Ile 29 → 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 → Ala (V149A) and Met 6 → Ala (M6A), however, are exceptions and have, respectively, one and two well-ordered water molecules within the cavity. The Val 149 → 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 → 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 → 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 cal mol\(^{-1}\) Å\(^{-2}\). 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 → 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 sidechains were replaced with smaller ones have shown that the loss in stability resulting from the same type of the replacement, e.g., Leu → 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 \( \Delta G \), the change of free energy of unfolding between mutant and wild-type protein, and the cavity size created, for a series of Leu → 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 (Fauçhère & Piliška, 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@xray.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

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Fractional solvent accessibility</th>
<th>Mutant</th>
<th>ΔΔG (kcal/mol)</th>
<th>Increase in cavity volume (Å³)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leu 7</td>
<td>0.0</td>
<td>L7A</td>
<td>−2.7</td>
<td>Non-isomorphous crystals</td>
<td></td>
</tr>
<tr>
<td>Leu 13</td>
<td>0.56</td>
<td>None</td>
<td>−2.7</td>
<td>Partially solvent exposed</td>
<td></td>
</tr>
<tr>
<td>Leu 15</td>
<td>0.19</td>
<td>None</td>
<td>−2.7</td>
<td>Partially solvent exposed</td>
<td></td>
</tr>
<tr>
<td>Leu 52</td>
<td>0.58</td>
<td>None</td>
<td>−2.7</td>
<td>Partially solvent exposed</td>
<td></td>
</tr>
<tr>
<td>Leu 33</td>
<td>0.64</td>
<td>L33A</td>
<td>−3.6</td>
<td>Crystals not obtained</td>
<td></td>
</tr>
<tr>
<td>Leu 39</td>
<td>0.15</td>
<td>L39A</td>
<td>−0.9</td>
<td>Partially solvent exposed</td>
<td></td>
</tr>
<tr>
<td>Leu 46</td>
<td>0.0</td>
<td>L46A</td>
<td>−2.7</td>
<td>24</td>
<td>Cavity</td>
</tr>
<tr>
<td>Leu 66</td>
<td>0.0</td>
<td>L66A</td>
<td>−2.7</td>
<td>Crystalization not attempted</td>
<td></td>
</tr>
<tr>
<td>Leu 79</td>
<td>0.27</td>
<td>None</td>
<td>−2.7</td>
<td>Solvent exposed</td>
<td></td>
</tr>
<tr>
<td>Leu 84</td>
<td>0.00</td>
<td>L84A</td>
<td>−3.8</td>
<td>Large cavity with channel to surface</td>
<td></td>
</tr>
<tr>
<td>Leu 91</td>
<td>0.16</td>
<td>L91A</td>
<td>−3.1</td>
<td>Small non-isomorphous crystals</td>
<td></td>
</tr>
<tr>
<td>Leu 99</td>
<td>0.00</td>
<td>L99A</td>
<td>−5.0</td>
<td>123</td>
<td>Cavity</td>
</tr>
<tr>
<td>Leu 118</td>
<td>0.01</td>
<td>L118A</td>
<td>−3.5</td>
<td>67</td>
<td>Cavity</td>
</tr>
<tr>
<td>Leu 121</td>
<td>0.00</td>
<td>L121A</td>
<td>−2.7</td>
<td>21</td>
<td>Cavity</td>
</tr>
<tr>
<td>Leu 133</td>
<td>0.00</td>
<td>L133A</td>
<td>−3.6</td>
<td>78</td>
<td>Cavity</td>
</tr>
<tr>
<td>Leu 164</td>
<td>0.50</td>
<td>None</td>
<td>−2.7</td>
<td>Partially solvent exposed</td>
<td></td>
</tr>
<tr>
<td>Ile 3</td>
<td>0.11</td>
<td>L3A</td>
<td>−1.1</td>
<td>Crystals not obtained</td>
<td></td>
</tr>
<tr>
<td>Ile 9</td>
<td>0.81</td>
<td>None</td>
<td>−2.7</td>
<td>Solvent exposed</td>
<td></td>
</tr>
<tr>
<td>Ile 17</td>
<td>0.53</td>
<td>L17A</td>
<td>−2.7</td>
<td>Crevise formed</td>
<td></td>
</tr>
<tr>
<td>Ile 27</td>
<td>0.0</td>
<td>L27A</td>
<td>−2.7</td>
<td>17</td>
<td>Crevise formed</td>
</tr>
<tr>
<td>Ile 29</td>
<td>0.05</td>
<td>L29A</td>
<td>−2.7</td>
<td>No detectable cavity</td>
<td></td>
</tr>
<tr>
<td>Ile 50</td>
<td>0.29</td>
<td>L50A</td>
<td>−2.0</td>
<td>Crevise formed</td>
<td></td>
</tr>
<tr>
<td>Ile 58</td>
<td>0.11</td>
<td>L58A</td>
<td>−3.2</td>
<td>18</td>
<td>Crevise formed</td>
</tr>
<tr>
<td>Ile 78</td>
<td>0.00</td>
<td>L78A</td>
<td>−2.7</td>
<td>Small non-isomorphous crystals</td>
<td></td>
</tr>
<tr>
<td>Ile 100</td>
<td>0.00</td>
<td>L100A</td>
<td>−3.4</td>
<td>28</td>
<td>Crevise formed</td>
</tr>
<tr>
<td>Ile 150</td>
<td>0.00</td>
<td>None</td>
<td>−2.7</td>
<td>Partially solvent exposed</td>
<td></td>
</tr>
<tr>
<td>Val 57</td>
<td>0.03</td>
<td>None</td>
<td>−2.7</td>
<td>Partially solvent exposed</td>
<td></td>
</tr>
<tr>
<td>Val 71</td>
<td>0.01</td>
<td>V71A</td>
<td>−1.5</td>
<td>Small non-isomorphous crystals</td>
<td></td>
</tr>
<tr>
<td>Val 75</td>
<td>0.00</td>
<td>None</td>
<td>−2.7</td>
<td>Partially solvent exposed</td>
<td></td>
</tr>
<tr>
<td>Val 87</td>
<td>0.00</td>
<td>V87A</td>
<td>−1.7</td>
<td>31</td>
<td>Crevise formed</td>
</tr>
<tr>
<td>Val 94</td>
<td>0.00</td>
<td>V94A</td>
<td>−1.8</td>
<td>Surface indentation</td>
<td></td>
</tr>
<tr>
<td>Val 103</td>
<td>0.12</td>
<td>V103A</td>
<td>−2.2</td>
<td>23</td>
<td>Crevise formed</td>
</tr>
<tr>
<td>Val 111</td>
<td>0.00</td>
<td>V111A</td>
<td>−2.7</td>
<td>11</td>
<td>Crevise formed</td>
</tr>
<tr>
<td>Val 131</td>
<td>0.12</td>
<td>V131A</td>
<td>−0.23</td>
<td>Solvent exposed</td>
<td></td>
</tr>
<tr>
<td>Val 149</td>
<td>0.00</td>
<td>V149A</td>
<td>−3.2</td>
<td>3</td>
<td>Solvent bound within cavity</td>
</tr>
<tr>
<td>Met 1</td>
<td>0.61</td>
<td>None</td>
<td>−2.7</td>
<td>Partially solvent exposed</td>
<td></td>
</tr>
<tr>
<td>Met 6</td>
<td>0.00</td>
<td>M6A</td>
<td>−1.9</td>
<td>10</td>
<td>Solvent bound within cavity</td>
</tr>
<tr>
<td>Met 102</td>
<td>0.01</td>
<td>M102A</td>
<td>−3.3</td>
<td>65</td>
<td>Crevise formed</td>
</tr>
<tr>
<td>Met 106</td>
<td>0.64</td>
<td>M106A</td>
<td>−2.3</td>
<td>Surface indentation</td>
<td></td>
</tr>
<tr>
<td>Met 120</td>
<td>0.00</td>
<td>M120A</td>
<td>−0.20</td>
<td>Surface indentation</td>
<td></td>
</tr>
<tr>
<td>Phe 4</td>
<td>0.03</td>
<td>None</td>
<td>−2.7</td>
<td>Partially solvent exposed</td>
<td></td>
</tr>
<tr>
<td>Phe 67</td>
<td>0.10</td>
<td>F67A</td>
<td>−1.9</td>
<td>Crevise formed</td>
<td></td>
</tr>
<tr>
<td>Phe 104</td>
<td>0.78</td>
<td>F104A</td>
<td>−3.1</td>
<td>Crevise formed</td>
<td></td>
</tr>
<tr>
<td>Phe 117</td>
<td>0.17</td>
<td>None</td>
<td>−2.7</td>
<td>Partially solvent exposed</td>
<td></td>
</tr>
<tr>
<td>Phe 153</td>
<td>0.01</td>
<td>F153A</td>
<td>−3.5</td>
<td>79</td>
<td>Crevise formed</td>
</tr>
</tbody>
</table>

*The table includes all leucines, isoleucines, valines, methionins, 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.

1Heinz et al. (1994).
2Eriksson et al. (1992).
3Baldwin et al. (1996).
4Data from Matsumura et al. (1988) at pH 2.0. Mutant made in the wild-type protein.
5Unpublished results of R. DuBose.
6Unpublished results of R. Kuruki.
7Dao-pin et al. (1990).
8Volume calculated in the presence of bound solvent.
9E.B. and B.W.M., unpublished.
10Blaber 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 deep cavities 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 (ΔΔG) were measured at pH 3.0. Additional details are given in Table 2.

Table 2. Thermostabilities of mutant lysozymes

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

*All 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). ΔTm is the change in melting temperature of the mutant relative to that of the wild-type (51.65°C). ΔH is the enthalpy of unfolding at the Tm of the mutant. ΔΔG, 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, ΔCp, estimated in this case to be 1.8 kcal/mol deg. The estimated error in ΔTm and ΔΔG is about ±0.2°C and 0.15 kcal/mol, respectively. The uncertainty in ΔΔG 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 Å² and 61 Å²,

![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).](image-url)
Cavity-creating mutations in T4 lysozyme

Table 3. Crystallographic data collection and refinement statistics

<table>
<thead>
<tr>
<th>Mutant lysozyme</th>
<th>Cell dimensions</th>
<th>Data collection</th>
<th>Refined model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$a, b$ (Å)</td>
<td>$c$ (Å)</td>
<td>Resolution (Å)</td>
</tr>
<tr>
<td>I17A</td>
<td>61.1</td>
<td>96.3</td>
<td>1.8</td>
</tr>
<tr>
<td>I27A</td>
<td>60.8</td>
<td>97.0</td>
<td>1.7</td>
</tr>
<tr>
<td>I29A</td>
<td>61.0</td>
<td>96.6</td>
<td>1.75</td>
</tr>
<tr>
<td>I50A</td>
<td>60.9</td>
<td>97.5</td>
<td>1.8</td>
</tr>
<tr>
<td>I58A</td>
<td>60.8</td>
<td>97.0</td>
<td>1.7</td>
</tr>
<tr>
<td>I100A</td>
<td>60.8</td>
<td>97.2</td>
<td>1.7</td>
</tr>
<tr>
<td>V87A</td>
<td>60.7</td>
<td>96.7</td>
<td>1.9</td>
</tr>
<tr>
<td>V103A</td>
<td>61.0</td>
<td>97.1</td>
<td>1.9</td>
</tr>
<tr>
<td>V111A</td>
<td>61.0</td>
<td>96.7</td>
<td>1.9</td>
</tr>
<tr>
<td>V149A</td>
<td>60.8</td>
<td>97.0</td>
<td>1.7</td>
</tr>
<tr>
<td>M4A</td>
<td>60.9</td>
<td>96.6</td>
<td>1.9</td>
</tr>
<tr>
<td>M106A</td>
<td>60.9</td>
<td>96.5</td>
<td>1.9</td>
</tr>
<tr>
<td>F67A</td>
<td>60.8</td>
<td>97.0</td>
<td>1.9</td>
</tr>
<tr>
<td>L84A</td>
<td>60.9</td>
<td>97.1</td>
<td>1.75</td>
</tr>
<tr>
<td>I27A/I29A</td>
<td>61.0</td>
<td>97.0</td>
<td>1.9</td>
</tr>
<tr>
<td>I27A/I58A</td>
<td>60.9</td>
<td>96.9</td>
<td>1.9</td>
</tr>
<tr>
<td>I29A/I58A</td>
<td>61.0</td>
<td>96.8</td>
<td>1.8</td>
</tr>
<tr>
<td>L121A/L133A</td>
<td>61.1</td>
<td>96.4</td>
<td>2.6</td>
</tr>
</tbody>
</table>

*The cell dimensions of wild-type lysozyme are $a = b = 61.2$ Å, $c = 96.8$ Å. 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 $\Delta_{bond}$ and $\Delta_{angle}$ 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 antiparallel $\beta$-strands and is completely inaccessible to solvent. The electron density map (Fig. 3A) and the refined structure of the Ile 27 → 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.
of residues 14–17 move toward the space vacated by the deleted atoms, the largest shift (≈0.4 Å) 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 Å). Other backbone movements (≈0.3 Å) 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 Å 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 Å 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 Å closer to Leu 13. The backbone atoms of residues 28 and 29 also move about 0.5 Å 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.
**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). Cβ of Ile 27 moves by 0.9 Å toward the space vacated by the isoleucyl sidechain and Cγ moves 0.7 Å. In doing so, Cα 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) shows that one turn of helix C, comprising residues from Ala 74 to Ile 78, rotates essentially about the axis of the helix such that Cα of Val 75 moves 0.4 Å and Cα 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. 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.
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 $\chi_1$, $\chi_2$, and $\chi_4$ of Met 106 are $-79^\circ$, $-164^\circ$, and $113^\circ$, respectively. In the mutant they change to $66^\circ$, $-176^\circ$, and $74^\circ$. 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, C4 of Met 102 rotates away to avoid a potential stearic 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.
Cavity-creating mutations in T4 lysozyme

<table>
<thead>
<tr>
<th>Protein</th>
<th>Center of cavity</th>
<th>Observed cavity size</th>
<th>Increase in cavity volume</th>
<th>Model cavity volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>X (Å)</td>
<td>Y (Å)</td>
<td>Z (Å)</td>
<td>V&lt;sub&gt;mut&lt;/sub&gt; (Å³)</td>
</tr>
<tr>
<td>I27A</td>
<td>28.0</td>
<td>8.9</td>
<td>0.6</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>32.1</td>
<td>4.0</td>
<td>-4.9</td>
<td>20</td>
</tr>
<tr>
<td>I29A</td>
<td>43.9</td>
<td>20.0</td>
<td>23.5</td>
<td>17</td>
</tr>
<tr>
<td>I58A</td>
<td>41.3</td>
<td>15.5</td>
<td>22.9</td>
<td>18</td>
</tr>
<tr>
<td>I100A</td>
<td>28.0</td>
<td>8.8</td>
<td>0.3</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>32.1</td>
<td>3.8</td>
<td>-4.9</td>
<td>20</td>
</tr>
<tr>
<td>V87A</td>
<td>28.0</td>
<td>8.6</td>
<td>0.6</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>32.2</td>
<td>4.1</td>
<td>-4.9</td>
<td>23</td>
</tr>
<tr>
<td>V103A</td>
<td>28.3</td>
<td>8.6</td>
<td>0.6</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>32.1</td>
<td>4.0</td>
<td>-4.9</td>
<td>19</td>
</tr>
<tr>
<td>V111A</td>
<td>27.8</td>
<td>9.1</td>
<td>1.2</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>32.3</td>
<td>4.1</td>
<td>-4.9</td>
<td>27</td>
</tr>
<tr>
<td>V149A</td>
<td>28.1</td>
<td>8.9</td>
<td>0.6</td>
<td>35b</td>
</tr>
<tr>
<td>(without water)</td>
<td>32.2</td>
<td>4.0</td>
<td>-4.9</td>
<td>21</td>
</tr>
<tr>
<td>V149A</td>
<td>28.1</td>
<td>8.9</td>
<td>0.6</td>
<td>35b</td>
</tr>
<tr>
<td>(with water)</td>
<td>32.2</td>
<td>4.0</td>
<td>-4.9</td>
<td>21</td>
</tr>
<tr>
<td>M6A</td>
<td>27.9</td>
<td>8.7</td>
<td>0.5</td>
<td>39b</td>
</tr>
<tr>
<td>(without water)</td>
<td>32.2</td>
<td>3.9</td>
<td>-4.8</td>
<td>25</td>
</tr>
<tr>
<td>M6A</td>
<td>27.9</td>
<td>8.7</td>
<td>0.5</td>
<td>39b</td>
</tr>
<tr>
<td>(with water)</td>
<td>32.2</td>
<td>3.9</td>
<td>-4.8</td>
<td>25</td>
</tr>
<tr>
<td>M102A</td>
<td>29.5</td>
<td>9.5</td>
<td>1.0</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>32.2</td>
<td>4.5</td>
<td>-4.7</td>
<td>35</td>
</tr>
<tr>
<td>I27A/I29A</td>
<td>43.8</td>
<td>19.8</td>
<td>23.7</td>
<td>21</td>
</tr>
<tr>
<td>I27A/I58A</td>
<td>41.4</td>
<td>16.2</td>
<td>22.9</td>
<td>27</td>
</tr>
<tr>
<td>I29A/I58A</td>
<td>41.4</td>
<td>15.6</td>
<td>23.2</td>
<td>12</td>
</tr>
<tr>
<td>L121A/L133A</td>
<td>27.2</td>
<td>4.1</td>
<td>0.5</td>
<td>8</td>
</tr>
</tbody>
</table>

| I & II     | 30.5  | 6.5  | -2.6 | 8                 | 126               | 234               |                     |

* | *Cavity volumes (V<sub>cavity</sub>) 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 Å³. No cavities are detected in the amino-terminal domain. ΔV<sub>model</sub> 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. b | *Mutants 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. c | *Model cavity for Ile 29 connects to the surface.*
Table 5. Stability and cavity volume data for alanine substitutions in barnase and ribonuclease-S

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Cavity volume (Å³)</th>
<th>Cavity volume (Å³)</th>
<th>ΔDG (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barnase Leu 14 → Ala</td>
<td>-4.3</td>
<td>89</td>
<td>24</td>
</tr>
<tr>
<td>Barnase Ile 76 → Ala</td>
<td>-3.2</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>Barnase Ile 88 → Ala</td>
<td>-4.0</td>
<td>91</td>
<td>21</td>
</tr>
<tr>
<td>Barnase Ile 96 → Ala</td>
<td>-1.9</td>
<td>83</td>
<td>44</td>
</tr>
<tr>
<td>R-Nase-S Met 13 → Ala</td>
<td>-4.3</td>
<td>24</td>
<td>-</td>
</tr>
</tbody>
</table>

*The 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.

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 Cα 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.
**Cavity-creating mutations in T4 lysozyme**

![Fig. 7. A: Difference density map for mutant 158A. All procedures and conventions as in Figure 3A. B: Superposition of mutant 158A (solid bonds) on WT* (open bonds) based on residues 1–60.](image)

**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 Cα 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
at position 58, where the Cα 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 constituent 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
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 Cα 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 “Fo-Fc” map as well as in the final “Fo-Fc” 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-mercaptoethanol (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 bonding 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
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 L29A 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. 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.
slightly from the previously reported values (−1.9 kcal/mol, 24 cal mol\(^{-1}\) Å\(^{-3}\); Eriksson et al., 1992) in part because the structure of L121A has been redetermined, leading to a decrease from 43 Å\(^{3}\) to 21 Å\(^{3}\) in cavity volume (Baldwin et al., 1996) and also because the Phe 153→Ala mutant was previously considered together with the Leu→Ala substitutions (Eriksson et al., 1992). With the exception of the one variant Val 149→Ala, the observations for all of the other mutants can also be approximately fitted by a series of straight lines of slope 22 cal mol\(^{-1}\) Å\(^{-3}\) (Fig. 17A, B). In Figure 17C the data are brought to a common reference scale by subtracting the respective intercept free energies, \(\Delta G_0\) (Table 6). The apparent outlier, Val 149→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→Ala replacements. 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→Ala (Fig. 17A). In this case the increase in cavity volume is essentially zero and the destabilization is the least of the Ile→Ala replacements.

*Fig. 11. A: Difference density map for mutant VI11A. All procedures and conventions as in Figure 3A. B: Superposition of mutant VI11A (solid bonds) on WT* (open bonds) based on residues 80–162*
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.
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 → Ala, for example, the intercept value ($\Delta G_r$) 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, Fauchère & Pliška, 1983).

Ignoring the outlier, Val 149 → 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 Fauchère & 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 $\Delta G_r$ relative to $\Delta G$. On average, however, $\Delta G_r$ is numerically larger than $\Delta G$. 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...
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 \( \beta \)-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 \textit{a priori} whether enthalpic or entropic considerations would dominate in either case. It might be noted that the discrepancy between \( \Delta \Delta G \), and the transfer free energy, \( \Delta G_t \), 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 sidechain 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 (Buckley et al., 1996) and ribonuclease-S (Varadarajan & Richards, 1992). If the cavity volumes quoted by Buckley 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 Buckley 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

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Free energy of transfer from water to octanol relative to alanine, ( \Delta G_t ), from Fauqué and Pliška (1983) (kcal/mol)</th>
<th>Extrapolated change in free energy, from alanine substitutions ( \Delta \Delta G ), Eq. (17) (kcal/mol)</th>
<th>Discrepancy ( \Delta \Delta G - \Delta G_t ) (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valine</td>
<td>−1.2</td>
<td>−1.2</td>
<td>0</td>
</tr>
<tr>
<td>Methionine</td>
<td>−1.2</td>
<td>−1.8</td>
<td>−0.6</td>
</tr>
<tr>
<td>Leucine</td>
<td>−1.9</td>
<td>−2.1</td>
<td>−0.2</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>−2.0</td>
<td>−1.8</td>
<td>0.2</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>−2.0</td>
<td>−2.7</td>
<td>−0.7</td>
</tr>
</tbody>
</table>

### 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 wildtype \( \Delta G \) and the packing density \( n_s \) (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, \( n_s \) 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...
Cavity-creating mutations in T4 lysozyme

Refinements were carried out with TNT (Tronrud et al., 1987) using the procedure described by Baldwin et al. (1996).

The mutants shown in different lines, are given in Table 1. The other lines are drawn parallel. The straight line is drawn through the origin with a slope of 22 cal mol$^{-1}$ Å$^{-3}$. Data for the double mutant L99A/F153A are from Eriksson et al. (1992).


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).
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, $V_s$, 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 Å$^3$, 10 Å$^3$, 13 Å$^3$, and 23 Å$^3$. If the probe radius is decreased to 1.1 Å, these four cavity volumes increase, respectively, to 14 Å$^3$, 15 Å$^3$, 45 Å$^3$, and 28 Å$^3$ and two new cavities of volumes ~6 Å$^3$ 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 Å$^3$ and the fourth cavity increases in volume to 36 Å$^3$. An additional five cavities of volumes 5–7 Å$^3$ are also detected. The uncertainty due to errors in the coordinates can be estimated from the variation in the volumes of Cavities I and 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 $37 \pm 3$ Å$^3$ and $22 \pm 3$ Å$^3$, respectively. (Mutant M102A is excluded because the substitution is close to Cavities I and II.) The observed variation of $\pm 3$ Å$^3$ 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 ($V_s$(WT)) from the cavity (or cavities) that were observed in the crystal structure of the mutant ($V_s$(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 $\beta$-carbon and the size of the resultant cavity was calculated. These values are quoted in Table 4 as $V_{\text{water}}$. There are slight discrepancies with related values quoted by Eriksson et al. (1992) because the radius of the $\beta$-carbon used previously was 2.02 Å whereas the present calculation assumes 2.10 Å which is more appropriate for a $\beta$-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 Kuraki and Robert Dubose for the data for mutants 117A and 150A. This work was supported in part by NIH grant GM21967 to B.W.M.

References


Das-pin S, Baase WA, Matthews BW. 1990. A mutant T4 lysozyme (Val 131 Ala) designed to increase thermoeositivity by the reduction of strain within an a-helix. Proteins 7:198–204.


Faber HR, Matthews BW. 1990. A mutant T4 lysozyme displays five different crystal conformations. Nature 348:263–266.


Cavity-creating mutations in T4 lysozyme


