ELSEVIER

Biophysical Chemistry 100 (2003) 325–340

Biophysical Chemistry

www.elsevier.com/locate/bpc

Multiple methionine substitutions are tolerated in T4 lysozyme and have coupled effects on folding and stability

Nadine C. Gassner[1], Walter A. Baase, Blaine H.M. Mooers, Robert D. Busam, Larry H. Weaver, Joel D. Lindstrom, Michael L. Quillin, Brian W. Matthews*

Institute of Molecular Biology, Howard Hughes Medical Institute and Department of Physics, 1229 University of Oregon, Eugene, OR 97403-1229, USA

Received 28 April 2002; accepted 23 May 2002

Abstract

In order to further explore the tolerance of proteins to amino acid substitutions within the interior, a series of core residues was replaced by methionine within the C-terminal domain of T4 lysozyme. By replacing leucine, isoleucine, valine and phenylalanine residues a total of 10 methionines could be introduced, which corresponds to a third of the residues that are buried in this domain. As more methionines are incorporated the protein gradually loses stability. This is attributed in part to a reduction in hydrophobic stabilization, in part to the increased entropic cost of localizing the long, flexible methionine sidechains, and in part to steric clashes. The changes in structure of the mutants relative to the wildtype protein are modest but tend to increase in an additive fashion as more methionines are included. In the most extreme case, namely the 10-methionine mutant, much of the C-terminal domain remains quite similar to wildtype (root-mean-square backbone shifts of 0.56 Å), while the F and G helices undergo rotations of approximately 20° and center-of-mass shifts of approximately 1.4 Å. For up to six methionine substitutions the changes in stability are additive. Beyond this point, however, the multiple mutants are somewhat more stable than suggested from the sum of their constituents, especially for those including the replacement Val111→Met. This is interpreted in terms of the larger structural changes associated with this substitution. The substituted sidechains in the mutant structures have somewhat higher crystallographic thermal factors than their counterparts in WT*. Nevertheless, the interiors of the mutant proteins retain a well-defined structure with little suggestion of molten-globule characteristics. Lysozymes in which selenomethionine has been incorporated rather than methionine tend to have increased stability. At the same time they also fold faster. This provides further evidence that, at the rate-limiting step in folding, the structure of the C-terminal domain of T4 lysozyme is similar to that of the fully folded protein.

Keywords: Methionine; Selenomethionine; Lysozyme; Core packing; Protein stability; Protein folding

*Corresponding author. Tel.: +1-541-346-2572; fax: +1-541-346-5870.

[1] Present address: Molecular and Cellular Physiology, Stanford University, 279 Campus Drive, Stanford, CA 94305-5101, USA.

0301-4622/03/$ - see front matter
PII: S0301-4622(02)00290-9

1. Dedication

Joseph Cohn and John Edsall's book *Proteins, Peptides and Amino Acids as Ions and Dipolar Ions* [1] is truly the bible of protein physical chemistry. It was written 59 years ago and even now serves as a useful reference to the basic attributes of the 'machines of life'. Almost one generation later this success was repeated with Edsall and Wyman's *Biophysical Chemistry* [2] putting a name to a then young field of scientific endeavor. It is truly an honour to acknowledge one of the great and also one of the first biophysical chemists and to dedicate this article to the memory of John T. Edsall.

2. Introduction

It is well known that proteins are usually tolerant of amino acid substitutions on their surfaces. This is not surprising in that solvent-exposed sidechains can often be replaced with little if any interference from the rest of the protein. In contrast, the cores of proteins are usually tightly packed, and substitutions at interior sites tend to introduce steric clashes or otherwise reduce the stability of the protein [3,4]. Nevertheless, proteins can tolerate some changes in core residues [5–7]. In order to explore the importance of jigsaw-like packing in defining protein structure we previously introduced a series of methionines into the core of T4 lysozyme [6]. Methionine was chosen because it has approximately the same volume as the common core residues leucine, isoleucine and phenylalanine. At the same time it has a flexible sidechain and, relative to most amino acids, is more able to adapt its conformation to occupy the space vacated by the residue that it replaces.

Here we extend this approach to ask how many methionines can be tolerated within the C-terminal core of T4 lysozyme (this being the largest domain of the protein). In previous work [6] methionine was primarily substituted for leucine. To further increase the number of methionines additional, less isosteric, replacements of isoleucine and valine were included in the largest constructs. We describe the effects on both structure and stability of successively including more methionines in the protein. The effects of these variants on the kinetics of refolding are also described. A comparison of the methionine-containing variants with their selenomethionine counterparts supports the idea [8] that the formation of wildtype-like structure in the C-terminal domain of T4 lysozyme is a rate-limiting step in folding.

3. Methods

3.1. Protein preparation

All mutants (Table 1) were constructed in the cysteine-free pseudo-wildtype lysozyme (WT*) [9]. For all of the multiple and most of the single mutants the modified genes were confirmed by sequencing in both the 3′ and 5′ directions. Protein expression and purification was carried out as described [6,10,11]. Protein expression for the selenomethionine variants (Table 2) was carried out in RR1 [12], a strain of bacteria not auxotrophic for methionine, using an adaptation of the procedure described by Van Duyne et al. [13]. All amino acids used for media preparation, including D,L-selenomethionine, were of analytical grade from Sigma. Selenomethionine incorporation was maximized when the culture was grown for 10 h to an A_{600} of 1.1, before induction to allow depletion of exogenous methionine levels. Thirty minutes before induction, the shaker temperature was lowered from 37 °C to 28 °C to reduce inclusion body formation. Fifteen minutes prior to induction, selenomethionine was added as a dry powder to 0.3 mM. Upon addition of IPTG to 1 mM the shaker speed was reduced to half (150 rev./min) and induction carried out for 4 h or more. To harvest the protein, cells were disrupted by three passages through a French press at 20 000 psi in lysis buffer (0.1 M KCl, 10 mM Tris–HCl, 5 mM $MgCl_2$, pH 8.0) supplemented with 15 mM methionine to reduce oxidation. All subsequent buffers included 15 mM methionine.

Expression of most of the variants, including 9a-M and 10a-M, yielded soluble protein that could be purified by standard methods [10]. The lysozymes 9b-M and 10b-M tended to express as

Table 1

Name of new variant[a]	Substitutions included[a]	Mets	PDB	Reference
I78M	I78M	1	1CU0	[8,19]
L84M	L84M	1	1CU2	[8,19]
V87M	V87M	1	1CU3	[8,19]
L91M	L91M	1	1CU5	[8,19]
L99M	L99M	1	1L93	[15]
I100M	I100M	1	1CUP	[8,19]
V103M	V103M	1	1CUQ	[8,19]
V111M	V111M	1	–	This work
G110R/V111M[b]	G110R/V111M	1	1CV1	[8,19]
L118M	L118M	1	1CV4	[8,19]
L121M	L121M	1	1CV3	[8,19]
L133M	L133M	1	1CV5	[19]
V149M	V149M	1	1CV6	[8,19]
F153M	F153M	1	1L88	[15]
2-M	118/121	2	1KS3	This work
3-M[a]	84/91/99	3	1KW5	This work
4a-M	84/91/99/153	4	1KY0	This work
4b-M	84/91/99/133	4	1KW7	This work
5-M[a]	84/91/99/118/121	5	1D3M(1D3N)	[19]
6a-M	84/91/99/118/121/153	6	1L0J	This work
6b-M	84/91/99/118/121/133	6	1KY1	This work
7a-M[a]	84/91/99/118/121/133/153	7	1CX7(1CX6)	[19]
7b-M[a]	84/91/99/111/118/121/133	7	1L0K	This work
7c-M	78/84/91/99/118/121/133	7	–	This work
8a-M	84/87/91/99/111/118/121/133	8	1LWG	This work
8b-M	84/91/99/100/103/118/121/133	8	–	This work
9a-M[c]	84/87/91/99/G110R/111/118/121/133/153	9	1LWK	This work
9b-M[d]	78/84/91/99/100/103/118/121/133	9	–	This work
9c-M	84/91/99/118/121/133/149/150/153	9	–	This work
10a-sM	84/87/91/99/100/103/G110R/111/118/121/133	10	1LPY	This work
10b-M	78/84/91/99/118/121/133/149/150/153	10	–	This work

[a] The naming convention is intended to indicate the number of methionines that have been engineered into the protein. '2-M', for example, has two engineered methionines. The designation '118/121' indicates that these are located at sites 118 and 121. If a mutation is to a residue other than methionine this is stated explicitly. Variants 4a-M and 4b-M both have four engineered methionines, but some of these are at different sites. In some prior publications a naming convention was used based on the total number of methionines, i.e. including the engineered ones plus the five that are present in the wildtype protein. For example, '3-M' was referred to as '8M', 5-M was referred to as '10M' and 7a-M and 7b-M were referred to as 12Ma and 12Mb [20]. Where the structure has been determined the PDB code is given. Codes given in parentheses refer to the selenomethionine-substituted protein.

[b] This mutant was described by Gassner and Matthews [19] and Gassner et al. [8] as 'V111M' but has subsequently been found to be the double mutant G110R/V111M (see text).

[c] This mutant was described by Gassner et al. [20] as '14M'. It has subsequently been found to include the additional mutation G110R.

[d] This mutant was described by Gassner et al. [6] as '10-Met'. It was subsequently found not to include the F153M substitution as described.

inclusion bodies and were purified as described by Gassner et al. [11]. The selenomethionine proteins were observed to elute from the CM-Sepharose column at a somewhat higher ionic strength than their methionine-containing counterparts.

3.2. Activity

The proteins all gave halos on the standard *E. coli* plate assay [14] showing that they were active. This was subsequently confirmed in vitro for some

Table 2
Crystallographic data collection and refinement statistics

Mutant	PDB code	Temperature (°C)	Resolution (Å)	Cell dimensions		Completeness (%)	R_{merge} (%)	R (%)	Δ_{bond} (Å)	Δ_{angle} (°)
				a,b (Å)	*c* (Å)					
2-M	1KS3	20	2.16	61.1	96.9	93	7.2	15.7	0.018	2.4
3-M	1KW5	20	1.75	61.1	96.9	90	4.1	16.1	0.018	2.7
4a-M	1KW7	20	1.89	61.2	97.1	93	4.5	15.8	0.020	2.7
4b-M	1KY0	20	1.97	61.3	96.5	94	4.1	16.0	0.020	2.8
5-M	1D3M	20	2.12	61.2	96.6	93	5.8	15.8	0.017	2.6
6a-M	1KY1	20	2.05	61.4	96.5	93	4.2	16.5	0.014	2.1
6b-M	1L0J	20	1.98	61.3	96.5	92	4.9	15.9	0.020	2.7
7a-M	1CX7	20	1.94	61.3	96.5	86	4.4	15.2	0.018	3.0
7b-M	1L0K	20	2.00	61.3	96.5	94	5.5	15.7	0.021	2.7
8a-M	1LWG	−173	1.7	60.4	90.4	99	5.2	19.6	0.019	2.9
9a-sM	1LWK	−173	2.1	60.3	91.4	99	5.0	21.7	0.018	2.9
10a-sM	1LPY	−173	1.7	60.2	93.2	96	4.3	20.7	0.017	2.6
WT*	1LWK	−173	1.45	60.2	95.4	99	5.6	19.7	0.018	2.6

All structures are isomorphous with WT* which has space group $P3_221$ and, at room temperature, cell dimensions of $a=b=60.9$ Å, $c=96.8$ Å [15]. The PDB code for the room temperature structure is 1L63.

of the multiple mutants using circular dichroism (CD) at 223 nm to monitor degradation of peptidoglycan fragments [11].

3.3. Structure determination

The proteins were crystallized isomorphously with WT* under similar conditions [15]. X-Ray data collection for the variants with up to seven introduced methionines was conducted at room temperature using a San Diego Multiwire Detector [16]. Difference density maps were calculated with WT* and mutant structure factors and WT* phases (Fig. 3a). Each mutant showed clear positive density features confirming introduction of the methionine sulfur atom at sites of mutation. For the variants with 9 and 10 replacements, the selenomethionine-substituted proteins were used. Also in these two cases as well as 8a-M the crystals were immersed in Paratone (Exxon Chemical) and then flash-frozen prior to data collection. Data for 8a-M and 9a-sM were collected at the Stanford Synchrotron Research Laboratory and for 10a-sM at the Advanced Light Source, Berkeley.

Atomic coordinates were refined with the TNT package [17] using WT* as a starting model [15]. Rigid-body refinement, typically using a 4-Å resolution cutoff, was first used to place the mutant in the unit cell. Next, rigid-body refinement was carried out for the two domains using a 3-Å resolution cutoff. In some cases individual elements of secondary structure were also refined as rigid units. Finally, positional and B-factor refinement were started with a 2.5-Å or 3-Å resolution cutoff and increased by 0.5 Å increments to incorporate the highest resolution data collected. Difference density maps were checked frequently for possible errors, and to conservatively place water molecules.

3.4. Protein stability

Stabilities were determined by monitoring the circular dichroism at 223 nm as a function of temperature [15,18] in a buffer of 0.1 M sodium chloride, 1.4 mM acetic acid, 8.6 mM sodium acetate, pH 5.42.

Standard free energy changes were computed for the process of protein unfolding at 59 °C assuming a change in heat capacity of 2.5 kcal/mol °C except in those cases where the enthalpy change upon unfolding was unusually low (specific instances noted in Table 3). In those cases the isotherm was moved to the apparent melting temperature of the variant. This has the effect of ignoring both the apparent ΔH and ΔC_p values of

Table 3
Thermal stability of methionine- and selenomethionine-substituted lysozyme

Mutant	ΔT_m (°C)	ΔH (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
WT*	–	130	–
WT*-sM	1.0	124	0.2
I27M/L33M	−10.4	61	−3.0
I17M/I127M/L33M	−9.3	55	−3.3*
V111M	−1.4	123	−0.6
G110R/V111M	−2.0	127	−0.7
L133M	−1.2	128	−0.4
2-M	−5.3	121	−1.9
3-M	−8.8	106	−3.1
3-sM	−6.6	103	−2.4
4a-M	−10.5	108	−3.8
4b-M	−9.6	112	−3.5
5-M	−12.1	108	−4.3
5-sM	−8.5	101	−3.0
6a-M	−13.3	101	−4.7
6b-M	−13.2	104	−4.7
7a-M	−14.5	96	−5.0
7a-sM	−8.8	94	−3.0
7b-M	−11.9	107	−4.3
7b-sM	−7.5	107	−2.6
7c-M	−17.6	74	−5.5
8a-M	−15.2	86	−4.9
8b-M	−21.2	62	−6.5*
9a-M	−17.7	74	−5.4
9a-sM	−10.3	103	−3.5
9b-M	−24.9	42	−7.0*
9b-sM	−19.2	45	−7.2*
9c-M	−24.4	46	−7.0*
10a-M	−23.4	48	−7.0*
10a-sM	−17.2	68	−5.2

ΔT_m is the change in melting temperature of the mutant protein relative to WT*, which unfolds at 65.3 °C. ΔH is the enthalpy of unfolding measured at the melting temperature and $\Delta\Delta G$ is the change in the free energy of unfolding of the mutant lysozyme relative to wildtype. '3M', for example, is the mutant with three substituted methionines (plus the five present in the native protein) defined in Table 1. '3-sM', for example, is the same protein but with selenomethionine rather than methionine. For the variants marked with an asterisk the enthalpy at the T_m was low for a purely two-state transition. In such cases $\Delta\Delta G$ was determined at the T_m of the mutant.

the mutant when applying the Gibbs–Helmholtz equation to compute ΔG.

3.5. Refolding kinetics

The kinetics of refolding were measured using stopped–flow mixing as described by Gassner et al. [8]. In brief, the protein was unfolded in 2.0 M urea at pH 2.0 and 20 °C and then rapidly mixed with buffer to give a final, folding solution that included 0.05 mg/ml protein, 64.4 mM $K_{3/2}H_{3/2}PO_4$, 0.1 M urea, 2.1 mM KCl, pH 6.7, also at 20 °C. Refolding was monitored by means of the CD at 223 nm as well as the total fluorescence above 338 nm as before [8]. Refolding rates were taken to be the principal decay component of a two-component exponential fit (ORIGIN, version 6) to the post burst-phase decay curves as before [8].

4. Results

4.1. Sites substituted

The sites that were substituted with methionine are shown in Fig. 1. All are within the C-terminal domain of T4 lysozyme and all are internal. Table 1 is intended to provide an up-to-date overview of all relevant methionine substitutions that have been made at buried sites within the C-terminal domain of T4 lysozyme. The table includes both single and multiple methionine mutants. In cases where the structure has been determined, the Protein Data Bank access code is included. During the detailed study of some of these mutants, we found from X-ray and sequence analysis that the spurious mutation Gly110→Arg (G110R) had been inadvertently included in some constructs (detailed in Table 1). For example, the mutant 10a-M, which has 10 introduced methionines, also includes this mutation.

4.2. Structural effects

The lysozymes with up to seven methionine substitutions could generally be crystallized readily. Variants incorporating larger numbers of methionines became successively more difficult to crystallize and, at room temperature, tended to diffract poorly. By using flash-freezing and synchrotron radiation, the resolution limit for crystals of WT* lysozyme can be increased from approximately 1.7 Å resolution to approximately 1.0 Å resolution (B.H.M.M. and B.W.M., unpublished observations). Flash-freezing and synchrotron radi-

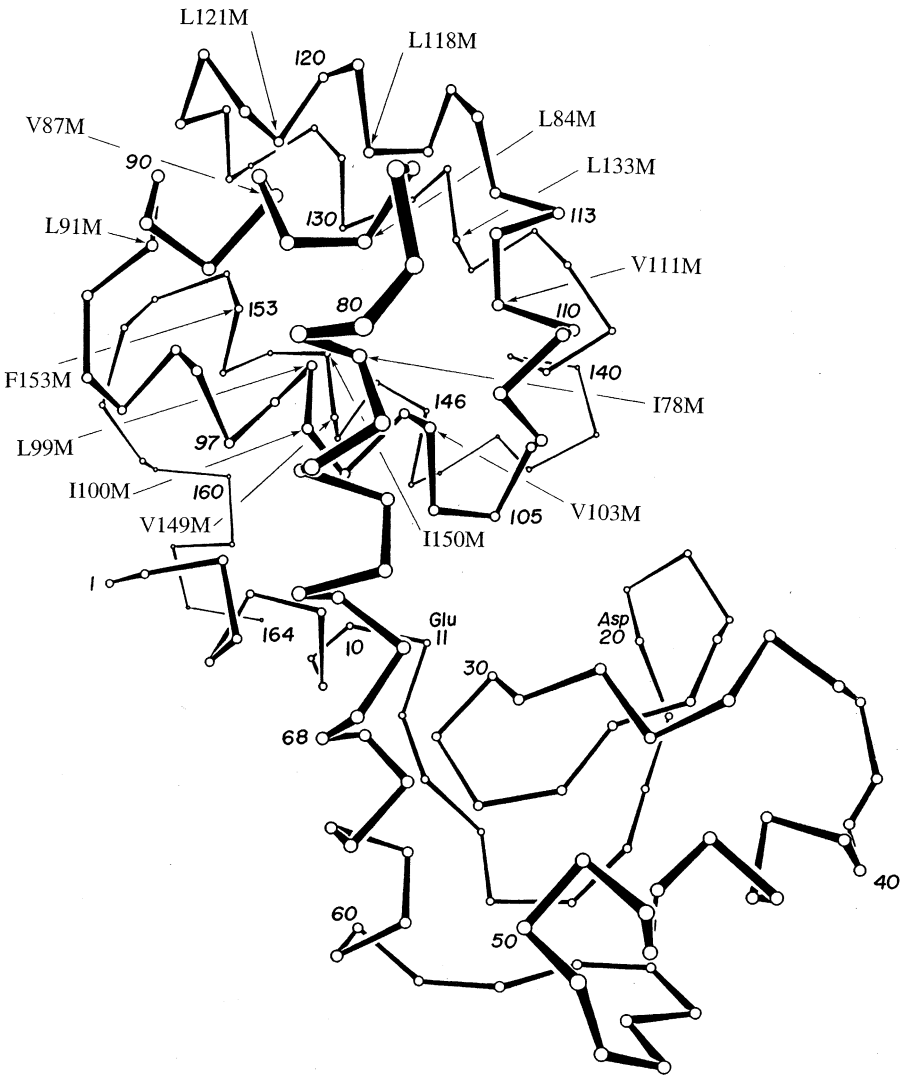

Fig. 1. Schematic drawing of the α-carbon backbone of T4 lysozyme showing the amino acids in the C-terminal domain that were substituted with methionine.

ation also improved the diffraction quality of the multiple methionine lysozymes, although not so dramatically. The 8-methionine protein, for example, increased from approximately 2.1 Å resolution to 1.7 Å resolution and the discrepancy between equivalent structure factor amplitudes decreased from 12.1% to 5.2%. Thus, flash-freezing made it feasible to collect data sets for T4 lysozyme with eight, nine and 10 substitutions. In two of these cases, selenomethionine was used instead of methionine. It has previously been shown that the structure of lysozyme substituted with seven selenomethionines is virtually identical with that substituted with seven methionines [19]. Therefore, from a structural perspective, substitution with selenomethionine or methionine is expected to make very little difference. The use of selenomethionine can, however, stabilize the protein somewhat relative to methionine [20]. Also, because selenium is more electron dense than sulfur it is

highly visible in the electron density map. If the sidechain is disordered, or has multiple conformations, this is more easily seen for selenomethionine than for methionine. Details of the X-ray data collection and refinement statistics for all the variants are given in Table 2.

Fig. 2a–d are 'shift plots' showing the changes in the protein backbone associated with the successive introduction of three, five, seven and 10 methionines or selenomethionines. The main result of introducing the first three methionines, which replace Leu84, Leu91 and Leu99, is a shift of approximately 0.5 Å near residue 84 (Fig. 2a). There is also a shift of up to 0.6 Å in the vicinity of the F helix (residues 108–114). This helix has higher crystallographic thermal factors than most of the T4 lysozyme molecule [15,21], and is often observed to shift and to become even more mobile when nearby substitutions are made in the core [22].

When the two additional mutations, Leu118→Met and Leu121→Met are included (Fig. 2b) the structural changes near residue 84 that were associated with the original triple-Met substitution are retained, and the shift in Helix F becomes slightly more pronounced. Fig. 2b shows that this helix moves in such a way that it increases its distance from residues 75–90 and 115–122 but at the same time moves closer to the rest of the protein.

Fig. 2c shows the effect of adding two more substitutions, V111M and L133M. As before, many of the structural changes seen in the 3-Met and 5-Met variants are retained. Val111 is, however, located in the F helix and its replacement with methionine causes even larger backbone changes in this vicinity. The center of mass of the F helix moves approximately 1.2 Å. In contrast, the Leu133→Met replacement causes only a small shift in the backbone at the site of the substitution, as is also the case for Leu99→Met (Fig. 2c).

The difference electron density map for the 7-Met mutant 7b-M is shown in Fig. 3a. The most obvious feature is the positive and negative density due to the shift in Helix F. Positive density features can also be seen at most of the substitution sites due to the enhanced electron density of methionine relative to leucine or valine. The backbone of the C-terminal domain of the 7b-Met mutant superimposed on that of WT* is shown in Fig. 3b. The overall close correspondence of the two structures, except for the F-helix, is very clear.

Fig. 2d is the shift plot comparing the 10-selenomethionine protein 10a-M with WT*. Relative to 7b-M (Fig. 2c) the additional substitutions are at sites 87, 100 and 103 plus the inadvertent mutation G110R. This mutant displays substantial changes in the backbone structure not only in Helix F but also in Helix G (residues 115–124). In order to characterize these structural changes in more detail, we first superimposed the C^{α} atoms of the C-terminal domain (residues 81–162) of 10a-sM on to the corresponding atoms in the low-temperature WT* structure. The root-mean-square discrepancy was 0.93 Å. If, however, the atoms including the F- and G-helices (residues 108–124) are deleted and the superposition is repeated, the root-mean-square discrepancy falls to 0.56 Å. Given this superposition (Fig. 3c) the F helix (residues 108–114) is seen to have rotated 23° and its center of mass translated approximately 1.5 Å while the G helix (residues 115–124) has translated approximately 1.2 Å and rotated 17°.

4.3. Effect on stability and folding

The thermal stabilities of the multiple methionine mutants are given in Table 3. In a number of cases the protein was also expressed incorporating selenomethionine rather than methionine. The thermal stabilities of these variants are included in the Table 3. It should be kept in mind that WT* lysozyme has five methionines, i.e. Met1, Met6, Met102, Met106 and Met120. In the selenomethionine-containing variants each of these five sites will necessarily be substituted with selenomethionine. Thus, in comparing the Met and SeMet versions of mutant 7b-M, for example, a total of 12 residues will be replaced with selenomethionine.

The trend in stability of the multiple methionine mutants is clearly downward with increasing numbers of substitutions. Equally clear is the recovery of a modicum of stability upon conversion to selenomethionine. Four multiple mutants which have the V111M mutation have higher melting temperatures (T_m) as well as concomitantly higher enthalpies of unfolding than do the other multiple

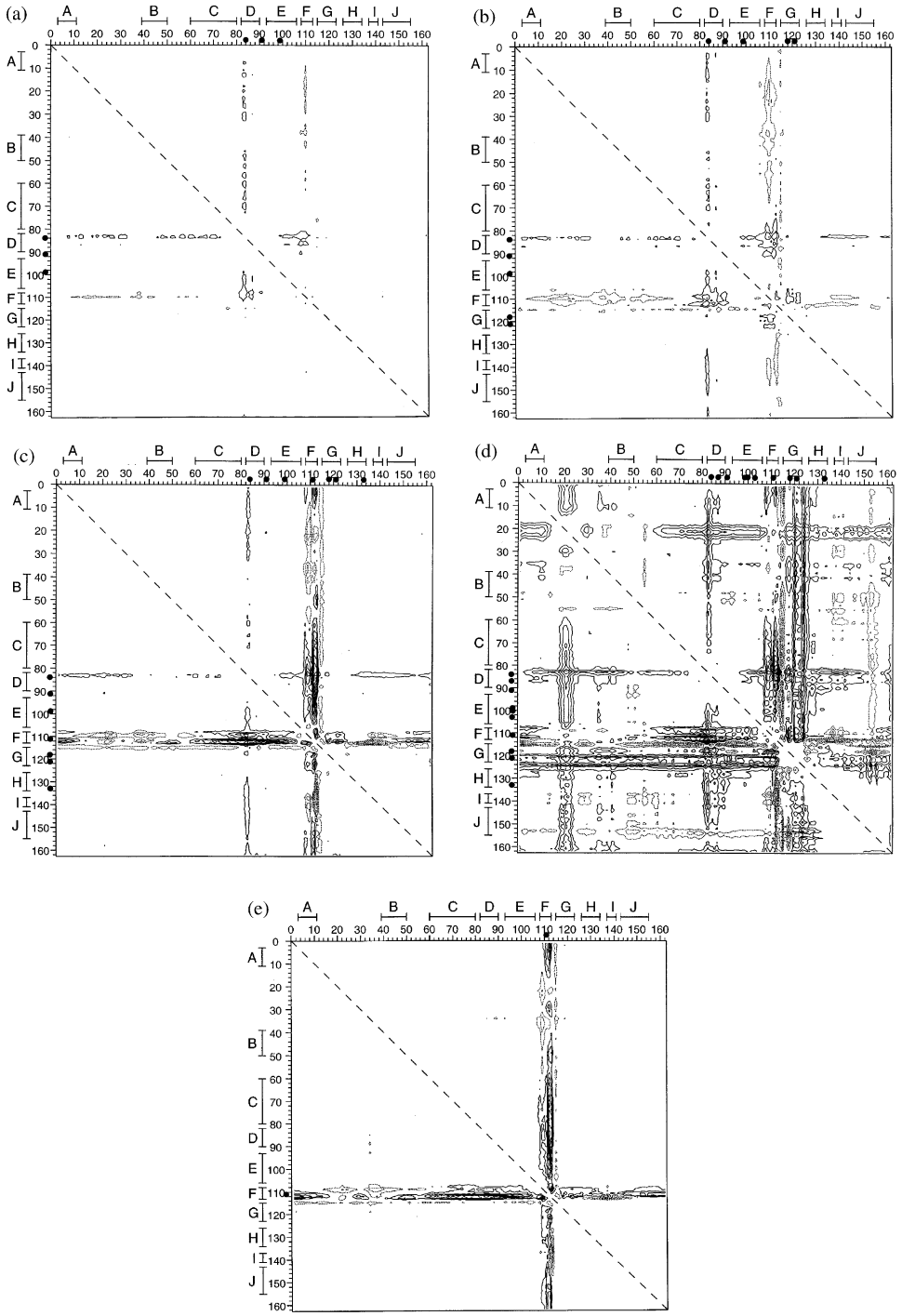

Fig. 2. 'Difference distance' plots showing the changes in distances between C^{α} atoms that occur on successive introduction of methionines. The first solid contour indicates pairs of C^{α} atoms which are 0.5 Å further apart in the mutant than in WT*. The first dashed contour indicates pairs of C^{α} atoms that move closer by 0.5 Å. Successive contours are at increments of 0.5 Å. The solid circles show the locations of the substituted methionines. In the following listing the PDB codes are given in parentheses. (a) Mutant 3-M (1KW5) vs. WT* (1L63). (b) Mutant 5-M (1D3M) vs. WT* (1L63). (c) Mutant 7b-M (1L0K) vs. WT* (1L63). (d) Frozen mutant 10-sM (1LPY) vs. frozen WT* (1LW9). (e) Mutant G110R/V111M (1CV1) vs. WT* (1L63).

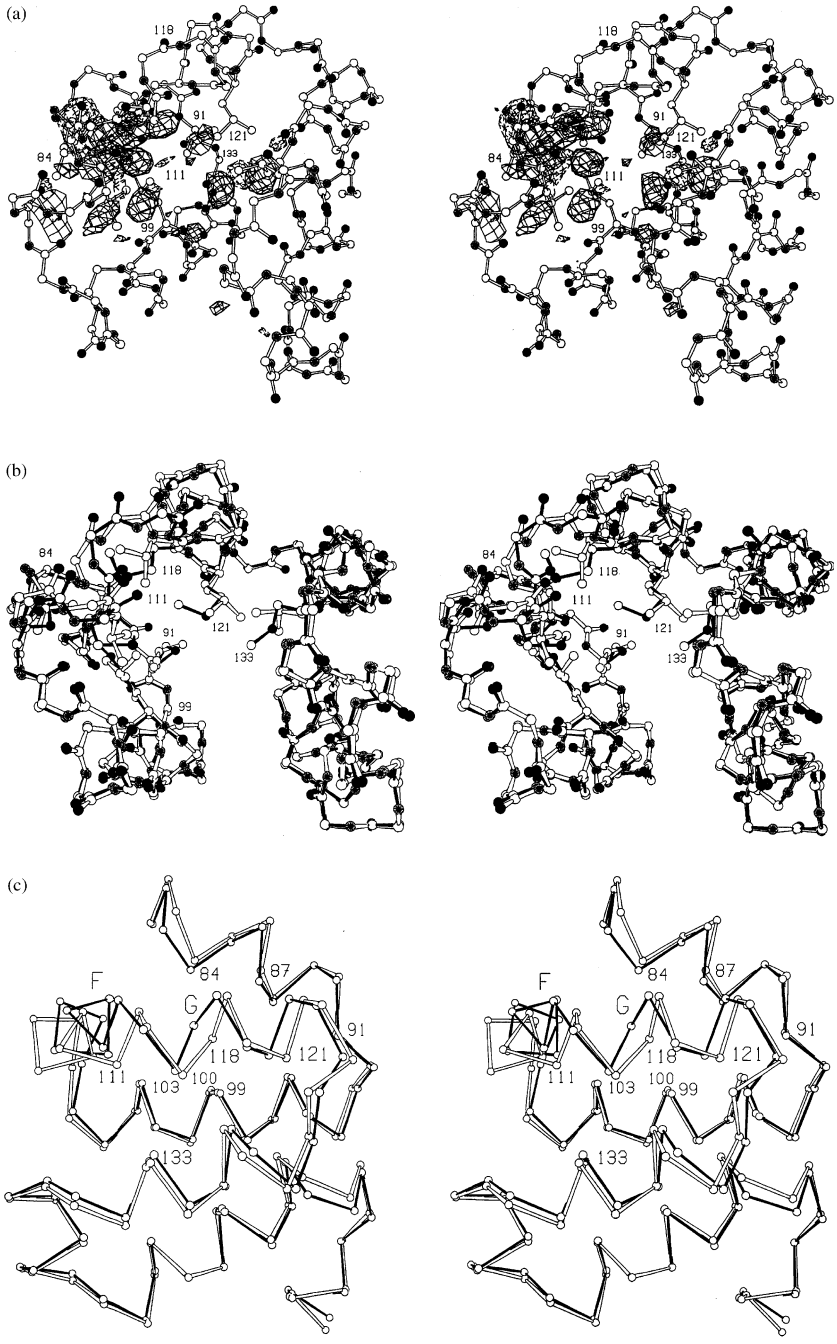

Fig. 3. (a) Difference electron density map for mutant 7b-M with methionines substituted at sites 84, 91, 99, 111, 118, 121 and 133. Coefficients are F_o(mutant)-F_o(WT*) where F_o are observed structure amplitudes and phases are from the refined structure of WT*. The map is contoured at increments of 3σ where σ is the root-mean-square density throughout the unit cell. The backbone structure of the C-terminal domain of T4 lysozyme is superimposed with nitrogen and carbon atoms drawn solid. For clarity, the sidechains are omitted except at the sites of substitution. (b) Superposition of the C-terminal backbone of mutant 7b-M (solid bonds) on that of WT* (open bonds). For clarity, all sidechains are omitted except at the sites of substitution. (c) Superposition of the backbone of the C-terminal domain (residues 81–162) of mutant 10a-M (open bonds) on WT* (solid bonds). The superposition is based on the alignment of residues 81–107 plus 125–162. The 10 residues that are substituted with selenomethionine in the mutant structure are numbered.

Table 4
Rates of refolding of methionine- and selenomethionine-substituted lysozymes

Methionine variants		Selenomethionine variants	
Protein	Rate, k (s^{-1})	Protein	Rate, k (s^{-1})
WT*	18.6	WT*-sM	17.8
L84M	8.7	L84sM	11.6
I100M	10.3	I100sM	12.4
V103M	21.1	V103sM	22.3
L133M	21.2	–	–
F153M	13.2	F153sM	19.3
G110R/V111M	13.1	–	–
2-M	11.7	–	–
3-M	5.5	3-sM	8.0
4a-M	4.3	–	–
4b-M	5.0	–	–
5-M	5.1	5-sM	8.1
6a-M	4.7	–	–
6b-M	5.6	–	–
7a-M	5.0	7a-sM	9.6
7b-M	6.6	7b-sM	12.6
8a-M	5.7	–	–
9a-M	4.4	9a-sM	7.8
I27M/L33M	17.7	–	–
I17M/I27M/L33M	25.8	–	–

The table gives the rates of refolding for single and multiple mutants of T4 lysozyme and, insofar as they have been determined, the corresponding rates for the selenomethionine-substituted protein. The rates for the additional single methionine mutants are given in Gassner et al. [8].

mutants containing identical numbers of new methionines but not V111M.

The rates of refolding of the multiple methionine-containing mutants and their seleno-methionine analogs are given in Table 4. As seen previously [8], approximately 50% of the expected CD signal developed within the deadtime of the instrument which was 11.5 ms [8]. Two-component exponential fits were done to the visible part of the decay curves to calculate the refolding rates which were taken to be the principal component and to eliminate a low-amplitude, slower component which has been ascribed to proline isomerization [8,25]. While rates determined by CD are presented here, those determined by the same method for the fluorescence signals were also computed. For the 42 mostly single-site mutants from Gassner et al. [8] plus the 24 new constructs presented here, the average ratio of the refolding rate determined by fluorescence to that determined by CD was 0.98 ($\pm$0.07), that is, essentially unity.

The refolding rates tend to decrease with stability but not in a monotonic manner at least not for the multiple methionine mutants in the C-terminal core. That is, for 3-M through 9a-M, refolding rates are uniformly decreased relative to WT* but only modest changes are apparent within the group, with clustering approximately 5 ($\pm$0.8) s^{-1}. As was the case for stabilities, the rates of refolding for the selenomethionine-substituted constructs were all greater than their methionine analogs (with the exception of WT* itself).

As was observed for single methionine substitution in the N-terminal domain [8], stabilities are decreased for multiple mutants (I27M/L33M and I17M/I27M/L33M; Table 3) although kinetics of refolding do not appear to be affected (Table 4).

5. Discussion

5.1. Tolerance to methionine substitution

One of the striking results of the present study is the tolerance of T4 lysozyme to methionine replacements within the core. The C-terminal domain of the protein can be considered to include residues 75–164 plus the N-terminal α-helix (Fig. 1). As such it contains all the naturally-occurring methionines in T4 lysozyme. Ignoring the N-terminal helix this corresponds to 89 residues of which 29 can be considered to be 'buried', i.e. the surface areas of their sidechains are at least 90% inaccessible to solvent [23]. In the 10-Met mutant 10 out of 29, or 34%, of these buried core residues are replaced with methionine. Nevertheless, the protein still folds and is active.

At the same time, the successive substitution of more methionines does result in a progressive loss of stability. On average, each additional substitution decreases stability by approximately 0.65 kcal/mol, with the individual point substitutions ranging from 0.4 kcal/mol for L99M and L133M to 2.3 kcal/mol for V87M [8]. The loss in stability can be attributed to a combination of three factors. First, a methionine sidechain has two more rotatable bonds than leucine (the most commonly

replaced residue) and is estimated to require an additional entropic free energy of approximately 0.8 kcal/mol to localize in the folded structure [6]. Second, the solvent transfer free energy of methionine is approximately 0.6 kcal/mol less than that of leucine. Finally, the substituted methionine may introduce steric clashes. In this context it is of interest to note that L84M, a highly destabilizing point mutant, distorts the local protein backbone (Fig. 2a), whereas L99M and L133M, the least destabilizing point mutants, cause very little perturbation of the backbone close to the site of substitution (Fig. 2a–c).

At the same time it should be emphasized that a large change in structure does not necessarily imply a large change in stability. In particular, as shown in Fig. 2e the mutant G110R/V111M causes a quite large (~3 Å) change in the local backbone but is only modestly destabilizing (0.6 kcal/mol).

5.2. Maintenance of core packing

Generally speaking, the core of the protein retains a well-packed, well-defined structure notwithstanding the introduction of up to 10 methionines. As can be seen in Fig. 2d there are a limited number of backbone atoms in the 10-selenomethionine structure that move up to 3.5 Å further apart than in WT*. At the same time there are also a few atoms that move approximately 1.5 Å closer to each other than in WT*. For most of the atoms, however, the shifts are much less (see also Fig. 3c). Thus, there is neither a dramatic expansion nor a collapse of the mutant structure.

As was seen previously with up to seven substitutions [6], the methionine sidechains tend to follow the paths of one or both branches of the aliphatic residues that they replace. This trend is also followed with the larger constructs. The following is a brief summary of the behavior of the 10 selenomethionines in 10a-sM. Both conformers of SeMet99 follow the CD2 branch of Leu99. The CD1 branch is not available due to the increased size of SeMet111. For SeMet91, one conformer follows each of the possible branches of Leu91. Val to SeMet substitutions are particularly disruptive. For SeMet111, despite the considerable backbone shifts, both conformers are still relatively cramped. For SeMet103, the sidechain follows the CG1 path which has additional space as a result of the substitution at Leu84. Because SeMet103 occupies volume near the CD1 branch of Leu84, SeMet84 follows the other branch (CD2). SeMet100 follows the longer of the two possible branches occupied by Ile100. The conformation of SeMet118 seems to be the least like the wildtype conformation and collapses toward the interior to fill voids left by nearby mutations at 84, 87 and 99. At position 87, the SeMet sidechain follows the CG1 path of Val87, largely because of increased volume from the Leu121 replacement. The CG2 path may be less favorable since it appears to be solvent-exposed. SeMet121 has a large backbone shift perhaps to make room for the replacements of Leu118 and Val87. SeMet133 follows the CD1 branch of Leu133.

Table 5 compares the crystallographic B-factors of the 10 substituted selenomethionines in the 10a-sM structure with the corresponding residues in the wildtype protein. As judged by the very modest

Table 5
Comparison of B-factors in wildtype lysozyme and the protein 10a-sM substituted with 10 selenomethionines

Amino acid in WT* (SeMet in mutant)	B-factor of C^{α} atom (Å^2)		Average B for sidechain atoms (Å^2)	
	WT*	10-sM	WT*	10-sM
Leu84	21	23	20	36
Val87	17	26	18	39
Leu91	13	16	16	23
Leu99	13	11	19	41
Ile100	17	9	16	18
Val103	14	11	18	21
Val111	18	22	17	25
Leu118	17	34	18	50
Leu121	15	25	17	30
Leu133	16	15	15	27
Average	16	19	17	31

The table summarizes the B-factors for each of the residues substituted in 10a-sM lysozyme. The first and third columns are taken from the refined structure of flash-frozen WT* lysozyme. The second and fourth columns are from the refined structure of the flash-frozen mutant. In the latter case the sidechains of residues 91, 99 and 111 have alternative conformations and the B-values quoted are the averages for the two conformers.

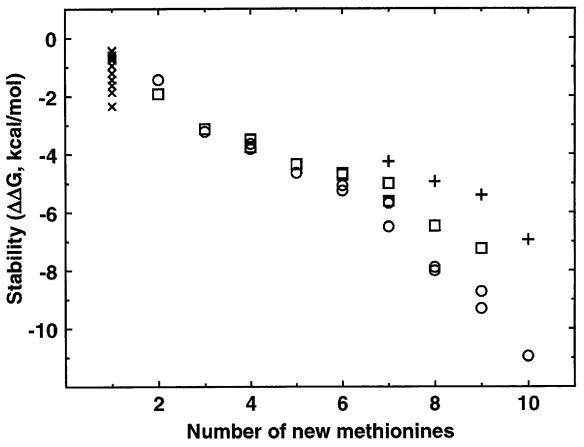

Fig. 4. Comparison of the observed stability ($\Delta\Delta G$) of each methionine multiple mutant (□) with the sum of the stabilities (○) for each of the constituent single mutants (X). Observed stabilities for the multiple mutants including V111M are indicated (+). The figure includes every multiple mutant for which the stabilities of the constituents are available (Table 3) and the data are plotted as a function of the number of introduced methionines. In some cases (e.g. 8a-M and 8b-M) the total number of methionines (eight) was achieved in two or more different ways (Table 1). This leads to two sets of $\Delta\Delta G$ values, one for mutant 8a-M and the other for 8b-M.

increases in the B-factors of the α-carbon atoms, the backbone of the mutant structure is essentially as well ordered as that in the wildtype structure. The selenomethionine sidechains in the mutant structure do show an average increase in B-factor from 17 Å^2 to 31 Å^2, relative to WT*. Also the sidechains of SeMet 91, 99 and 111 each display two alternate conformations at approximately 25:75 occupancy. Nevertheless, the interior of the mutant still retains a well-defined structure with little suggestion of molten-globule characteristics.

5.3. Additivity of changes in stability for multiple mutants

Fig. 4 compares the observed loss in stability of the multiple methionine mutants with the sum of the loss of stability of each of the constituent single mutants. Up to the point that six methionines are included these values are almost identical. This indicates that each of the methionine substitutions is acting in an independent manner. Also, as suggested in Fig. 2a,b, the structural changes accumulate in an additive fashion with little if any compensation or cancellation from one site to another. However, with seven or more methionine substitutions, the protein is somewhat more stable than predicted by summing the effects of the constituent single mutations. This effect is especially noticeable for the four variants 7b-M, 8a-M, 9a-M and 10a-M (Fig. 4). Strikingly, these are the multiple mutants that include the substitution V111M. As noted above, this replacement is located within the F helix and also causes an unusually large change in the backbone conformation (Fig. 2e). It is this same helix that undergoes the largest structural change in response to changes at other sites (Fig. 2a,b). This suggests that the effect of the V111M mutation is to move part of the F helix and to disrupt the hydrophobic contacts between this helix and the rest of the protein. When other mutants which also tend to distort the F helix are combined with V111M, the energetic cost of the distortion is shared. Thus, the combined mutants realize net gains in stability relative to their constituents. It may also be noted that the variants that include V111M (7b-M, 8a-M, 9a-M and 10a-M) all have enthalpies of unfolding that are equal to or higher than their counterparts with equal numbers of methionines. This also suggests that the presence of V111M has improved the packing in these variants.

The three residues in 10a-sM that have alternative sidechain conformations (sites 91, 99 and 111) are arranged in a line in the folded structure but the newly-introduced selenomethionine sidechains do not appear to strongly interact with each other. Depending on the actual fractions of the observed alternative conformations and whether this variable occupancy is concerted or independent, we expect a stabilizing entropic contribution to the free energy of up to approximately 1 kcal/mol relative to a situation in which such motion would not be possible. Present data are insufficient to say whether or not this effect contributes to the enhanced stability.

5.4. Kinetics of folding

Gassner et al. [8] previously determined the kinetics of folding of a series of single alanine and

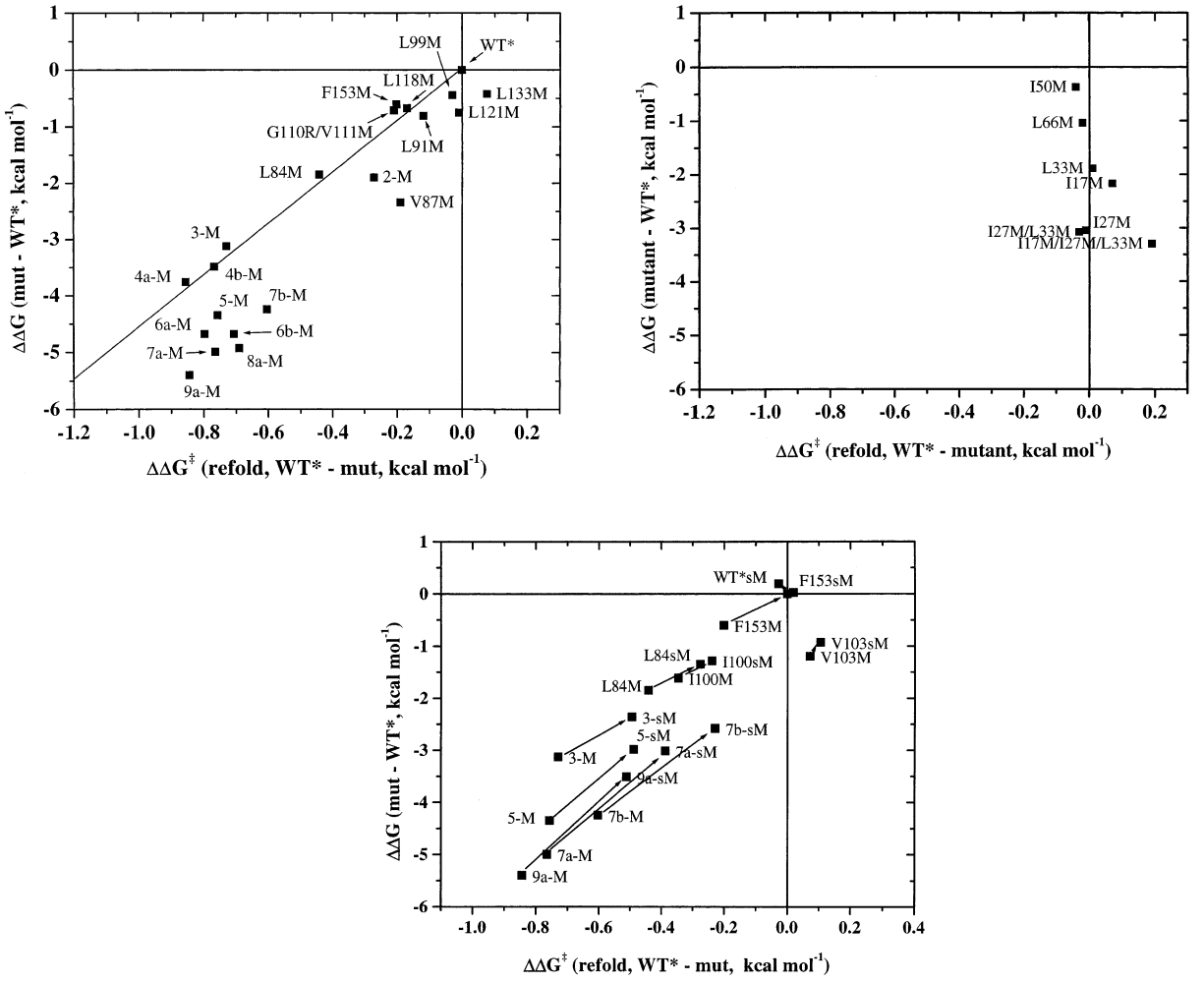

Fig. 5. Comparison of the effects of mutations on the stability of T4 lysozyme and on the kinetics of folding. $\Delta\Delta G$ is the change in the Gibbs free energy of folding of the mutant relative to WT*. $\Delta\Delta G^{\ddagger}$ is derived from the rate constants for folding (Table 4) by the equation $\Delta\Delta G^{\ddagger} = -RT\ln(k_{WT*}/k_{mut})$ where k_{mut} and k_{WT*} are the observed refolding rates of the mutant and WT*, respectively. (a) Single- and multiple-methionine mutants within the C-terminal domain of T4 lysozyme. The straight line was obtained previously [8]. (b) Single- and multiple-methionine mutants within the N-terminal domain of T4 lysozyme. (c) Comparison of methionine and selenomethionine mutants of T4 lysozyme. The arrowed lines connect the data point for the methionine-containing mutant with the data point for the corresponding selenomethionine mutant.

methionine substitutions in T4 lysozyme. It was found that many of the substitutions that destabilized the molecule and were located in the C-terminal domain also slowed down the rate of folding. In contrast, destabilizing mutations that were located in the N-terminal domain had little effect on the rate of folding. The availability of the multiple mutants described here, some of which substantially reduce stability, provides an opportunity to further examine this trend.

In Fig. 5a we compare the effect of single and multiple methionine substitutions in the C-terminal domain on the stability of T4 lysozyme and on the kinetics of folding. Fig. 5b is a similar comparison for all available methionine substitutions located within the N-terminal domain. Unfortu-

nately, the latter sample is rather limited. Nevertheless, the overall trend seems clear. Within the C-terminal domain, mutations that destabilize the protein also reduce the rate of folding. In contrast, mutations in the N-terminal domain may reduce stability but they do not slow folding. This strongly supports the idea that formation of the C-terminal domain is rate-limiting in the folding of T4 lysozyme and that the structure of this domain at the transition state is akin to that of the fully-folded protein although the hydrophobic core is not fully developed [8].

Additional support for this idea comes from a comparison of the stability and kinetic data for the methionine and selenomethionine proteins (Fig. 5c). Due to the increased hydrophobicity of selenomethionine relative to methionine, the selenomethionine-containing lysozymes are consistently more stable than their methionine counterparts [20]. At the same time, the structures of the Met and SeMet proteins are highly isomorphous [19]. Thus, the replacement of Met with SeMet makes it possible to generate pairs of proteins which have virtually identical structures but differ significantly in stability (Table 3).

As shown in Fig. 5c the replacement of Met with SeMet not only increases stability, but also increases the rate of folding. This provides further evidence that the rate-limiting step for folding of T4 lysozyme must have a structure in which the C-terminal domain is quite similar to that of the folded protein.

Within the context of this overall observation that refolding rates and stabilities are simply related, we note that some effects in Fig. 5a appear to be characteristic of the particular sites being converted to methionine. The diagonal line (with slope of $1/0.22$) is drawn from Gassner et al. [8] wherein it appeared that in the case of single mutants approximately 22% of the free energy of refolding was typically manifested during the formation of a transition state. We note that points for the multiple mutants depart from this line approximately in accord with the degree to which they contain mutations at sites 121 and 133. These two individual mutations also depart from this line. On a case-by-case basis, 2-M contains 121M and is below this line. 3-M, 4a-M and 4b-M contain neither 121M nor 133M and fall on or slightly above the line. 5-M and 6a-M contain 121M and are below it. 6b-M, 7a-M and 7b-M contain both 121M and 133M and are roughly twice as far below the line. This is also the case for 8a-M and 9a-M which both contain 121M and 133M but which also have 87M. V87M is a site which is quite atypical as a single mutation but which in the cases of these two multiples does not seem to further contribute, possibly because it is part of the same unit of secondary structure as L84M, a mutation which reduces the refolding rate to an even greater degree as a single site mutant than does V87M.

While conditions have been found for which T4 lysozyme folding appears to be two-state [24], the WT* refolding pathway in the solvent condition used here is known to have a kinetic intermediate [25,26]. The presence of such an intermediate was consistent with refolding data for T4 alanine and methionine point mutants [8,27] and has been seen for other T4 lysozyme mutants (W.A.B. and B.W.M., unpublished results). Hence, although kinetic intermediates have not yet been directly demonstrated for even the single-site T4 lysozyme methionine mutants, it is likely the rates observed here correspond to an intermediate-to-folded transition. This invites rationalization of these data in terms of stability changes of an intermediate; however, neither changes in barrier heights nor transmission efficiencies over barriers can be ruled out by our rate data.

Selenomethionine replacement of methionine increases the stability of such constructs (Table 3). It also accelerates the rate of refolding in a manner that is directly proportional to the increase in stability (see the connected arrows in Fig. 5c). In other words, it appears for each of these mutants that approximately 22% of the free energy of folding is manifest in the formation of the transition state (Fig. 5c). We believe this means that the isomorphism between methionine and selenomethionine found in the final folded proteins is also present in the transition states. Because the multiple mutants studied here are a nested set (e.g. L84M is present in all multiples in Fig. 5c), the

results may not be general. It seems, however, that selenomethionine incorporation might be a sufficiently small perturbation to allow the uncoupling of stability effects from structural effects in the study of protein folding, at least for those proteins which contain or can be made to contain a sufficient number of methionine residues.

Acknowledgments

We are most grateful to Hong Xiao, Andrew Fields and Leslie Gay for excellent technical assistance. We also thank Gerry McDermott and Keith Henderson at ALS and Aina Cohen, Paul Ellis and Mike Soltis at SSRL for technical support with the use of synchrotron radiation. This work was supported in part by NIH grant GM21967 to B.W.M.

References

[1] E.J. Cohn, J.T. Edsall, Proteins, Amino Acids and Peptides as Ions and Dipolar Ions, Reinhold Publishing Company, New York, 1943.
[2] J.T. Edsall, J. Wyman, Biophysical Chemistry, Academic Press, New York, 1958.
[3] J.U. Bowie, J.F. Reidhaar-Olson, W.A. Lim, R.T. Sauer, Deciphering the message in protein sequences: tolerance to amino acid substitutions, Science 247 (1990) 1306–1310.
[4] B.W. Matthews, Structural and genetic analysis of protein stability, Ann. Rev. Biochem. 62 (1993) 139–160.
[5] D.D. Axe, N.W. Foster, A.R. Fersht, Active barnase variants with completely random hydrophobic cores, Proc. Natl. Acad. Sci. USA 93 (1996) 5590–5594.
[6] N.C. Gassner, W.A. Baase, B.W. Matthews, A test of the 'jigsaw puzzle' model for protein folding by multiple methionine substitutions within the core of T4 lysozyme, Proc. Natl. Acad. Sci. USA 93 (1996) 12155–12158.
[7] K.W. Plaxco, D.S. Riddle, V. Grantcharova, D. Baker, Simplified proteins: minimalist solutions to the 'protein folding problem', Curr. Opin. Struct. Biol. 8 (1998) 80–85.
[8] N.C. Gassner, W.A. Baase, J.D. Lindstrom, J. Lu, F.W. Dahlquist, B.W. Matthews, Methionine and alanine substitutions show that the formation of wildtype-like structure in the carboxy-terminal domain of T4 lysozyme is the rate-limiting step in folding, Biochemistry 38 (1999) 14451–14460.
[9] M. Matsumura, B.W. Matthews, Control of enzyme activity by an engineered disulfide bond, Science 243 (1989) 792–794.
[10] A.R. Poteete, S. Dao-pin, H. Nicholson, B.W. Matthews, Second-site revertants of an inactive T4 lysozyme mutant restore activity structuring the active site cleft, Biochemistry 30 (1991) 1425–1432.
[11] N.C. Gassner, W.A. Baase, J.D. Lindstrom, B.K. Shoichet, B.W. Matthews, Isolation and characterization of multiple-methionine mutants of T4 lysozyme with simplified cores, in: D. Marshak (Ed.), Techniques in Protein Chemistry VIII, Academic Press, 1997, pp. 851–863.
[12] D.C. Muchmore, L.P. McIntosh, C.B. Russell, D.E. Anderson, F.W. Dahlquist, Expression and ^{15}N labelling of proteins for proton and nitrogen-15 NMR, Meth. Enzymol. 177 (1989) 44–73.
[13] G.D. Van Duyne, R.F. Standaert, P.A. Karplus, S.L. Schreiber, J. Clardy, Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin, J. Mol. Biol. 229 (1993) 105–124.
[14] A. Tsugita, Phage lysozyme and other lytic enzymes, in: P.D. Boyer (Ed.), Third Ed, The Enzymes, 5, 1971, pp. 343–411.
[15] A.E. Eriksson, W.A. Baase, B.W. Matthews, 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 (1993) 747–769.
[16] R. Hamlin, Multiwire area X-ray diffractometers, Meth. Enzymol. 114 (1985) 416–452.
[17] D.E. Tronrud, L.F. Ten Eyck, B.W. Matthews, An efficient general-purpose least-squares refinement program for macromolecular structures, Acta Cryst. A43 (1987) 489–503.
[18] X-J. Zhang, W.A. Baase, B.K. Shoichet, K.P. Wilson, B.W. Matthews, Enhancement of protein stability by the combination of point mutations in T4 lysozyme is additive, Prot. Engin. 8 (1995) 1017–1022.
[19] N.C. Gassner, B.W. Matthews, Use of differentially substituted selenomethionine proteins in X-ray structure determination, Acta Cryst. D55 (1999) 1967–1970.
[20] N.C. Gassner, W.A. Baase, A.C. Hausrath, B.W. Matthews, Substitution with selenomethionine can enhance the stability of methionine-rich proteins, J. Mol. Biol. 294 (1999) 17–20.
[21] L.H. Weaver, B.W. Matthews, Structure of bacteriophage T4 lysozyme refined at 1.7 Å resolution, J. Mol. Biol. 193 (1987) 189–199.
[22] J. Xu, W.A. Baase, E. Baldwin, B.W. Matthews, The response of T4 lysozyme to large-to-small substitutions within the core and its relation to the hydrophobic effect, Prot. Sci. 7 (1998) 158–177.
[23] T. Alber, S. Dao-pin, J.A. Nye, D.C. Muchmore, B.W. Matthews, Temperature-sensitive mutations of bacteriophage T4 lysozyme occur at sites with low mobility

and low solvent accessibility in the folded protein, Biochemistry 26 (1987) 3754–3758.
[24] B-L. Chen, W.A. Baase, J.A. Schellman, Low-temperature unfolding of a mutant of phage T4 lysozyme. 2. Kinetic investigations, Biochemistry 28 (1989) 691–699.
[25] J. Lu, F.W. Dahlquist, Detection and characterization of an early folding intermediate of T4 lysozyme using pulsed hydrogen exchange and two-dimensional NMR, Biochemistry 31 (1992) 4749–4756.
[26] M.J. Parker, S. Marqusee, The cooperativity of burst phase reactions explored, J. Mol. Biol. 293 (1999) 1195–1210.
[27] B-L. Chen, W.A. Baase, H. Nicholson, J.A. Schellman, Folding kinetics of T4 lysozyme and nine mutants at 12 °C, Biochemistry 31 (1992) 1464–1476.