Protein Science (1994), *3*:92–102. Cambridge University Press. Printed in the USA.

Response of dynamic structure to removal of a disulfide bond: Normal mode refinement of C77A/C95A mutant of human lysozyme

AKINORI KIDERA,[1] KOJI INAKA,[2] MASAAKI MATSUSHIMA,[1] AND NOBUHIRO GŌ[3]

[1] Protein Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565, Japan
[2] Research Laboratory of Resources Utilization, Tokyo Institute of Technology, 4259 Nagatsuda, Midori-ku, Yokohama 227, Japan
[3] Department of Chemistry, Faculty of Science, Kyoto University, Kyoto 606, Japan

Abstract

In order to investigate the response of dynamic structure to removal of a disulfide bond, the dynamic structure of human lysozyme has been compared to its C77A/C95A mutant. The dynamic structures of the wild type and mutant are determined by normal mode refinement of 1.5-Å-resolution X-ray data. The C77A/C95A mutant shows an increase in apparent fluctuations at most residues. However, most of the change originates from an increase in the external fluctuations, reflecting the effect of the mutation on the quality of crystals. The effects of disulfide bond removal on the internal fluctuations are almost exclusively limited to the mutation site at residue 77. No significant change in the correlation of the internal fluctuations is found in either the overall or local dynamics. This indicates that the disulfide bond does not have any substantial role to play in the dynamic structure. A comparison of the wild-type and mutant coordinates suggests that the disulfide bond does not prevent the 2 domains from parting from each other. Instead, the structural changes are characteristic of a cavity-creating mutation, where atoms surrounding the mutation site move cooperatively toward the space created by the smaller alanine side chain. Although this produces tighter packing, more than half of the cavity volume remains unoccupied, thus destabilizing the native state.

Keywords: cavity-creating mutation; disulfide mutant; dynamic structure; human lysozyme; normal mode refinement

Crystallographic studies on genetically engineered mutant proteins have shown how amino acid substitution affects 3-dimensional structure. In particular, numerous mutant structures of bacteriophage T4 lysozyme have been solved to high resolution by Matthews and his colleagues (Matthews, 1987).

However, X-ray data do not necessarily provide a clear-cut picture of mutation effects on the dynamic structure. The dynamic structure is usually determined in the form of the isotropic temperature factors for non-hydrogen atoms (*B*-factors), which inevitably contain contributions from the external terms, such as lattice disorder, as well as the internal atomic fluctuations (Frauenfelder et al., 1979; Finzel & Salemme, 1985; Sheriff & Hendrickson, 1987; Diamond, 1990). Usually, a mutation has an influence not only on the internal fluctuations but also on the quality of the crystal, which in turn affects the external terms. Therefore, it is difficult to decide whether changes in *B*-factors are caused by the mutation effects on the internal fluctuations, or are merely due to the difference in crystal quality (Weber et al., 1985; Soman et al., 1991).

For the purpose of solving this problem, we introduced a method of dynamic structure refinement, which we call normal mode refinement (Kidera & Gō, 1990, 1992; Kidera et al., 1992). In this method, the structure factor, F_{calc}, is given by

$$F_{\mathrm{calc}}(\mathbf{q}) = \sum_j f_j(\mathbf{q}) \exp(i\mathbf{q}\cdot\mathbf{r}_j) \times \exp\left(-\frac{1}{2}\sum_{k,l=1}^{3} q_k q_l \sum_{m,n=1}^{M+6} \phi_{jkm}\phi_{jln}\sigma_{mn}\right), \quad (1)$$

where $\mathbf{q}$ $(=\{q_k\})$ is a reciprocal lattice vector and $f_j(\mathbf{q})$ is the atomic structure factor for atom j, with $\mathbf{r}_j$ being the average coordinates. In this formula, the Debye–Waller factor describing the dynamic structure of a protein is expanded in terms of normal modes. ϕ_{jkm} is the component of the mth normal mode given theoretically by normal mode analysis, and σ_{mn} is covariance for the internal or external normal modes. The summation is over the M lowest frequency internal normal modes as well

Reprint requests to: Akinori Kidera, Protein Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565, Japan.

as the 6 external normal modes of the TLS (translation, libration, and screw) model (Schomaker & Trueblood, 1968). Here, M is determined by considering the balance between the number of experimental data and that of adjustable parameters. This formula of the structure factor is based on the theoretical model of protein dynamics, which states that conformational fluctuations occur mostly in the important conformational subspace spanned by a small number (M in Equation 1) of low frequency normal modes (Kidera & Gō, 1992).

This refinement proceeds as follows. First, for a given set of coordinates, normal mode vectors, ϕ_{jkm}, are calculated as a basis set for describing the conformational fluctuations. Then, the coefficients for this basis set, σ_{mn}, are determined in the course of the crystallographic refinement to give the best fit to the observed diffraction data. Because the normal modes reflect the detailed covalent and 3D structure of the molecule, atomic fluctuations are treated as being correlated and anisotropic, unlike in the conventional isotropic B-factor model.

In this model, the M internal normal modes and the 6 external normal modes recognize the internal and external fluctuations, respectively. Therefore, we can investigate the effects of mutations on the internal protein dynamics without any ambiguity of the external terms.

In this paper, we describe the normal mode refinement of a human lysozyme mutant, C77A/C95A, in which a disulfide bond is removed by mutations of Cys 77 and Cys 95 to alanine (Taniyama et al., 1988). In comparison to the wild type, we discuss how the removal of this disulfide bond affects the internal atomic fluctuations. It has been suggested that lysozyme has a function-related hinge-bending motion, which causes a cooperative fluctuation between the 2 structural domains (McCammon et al., 1976; Faber & Matthews, 1990; Gibrat & Gō, 1990). As shown in Figure 1 and Kinemage 1, the disulfide bond is situated at the hinge region connecting the 2 structural domains. The results of normal mode refinement should describe a possible role of the disulfide bond in the hinge-bending motion.

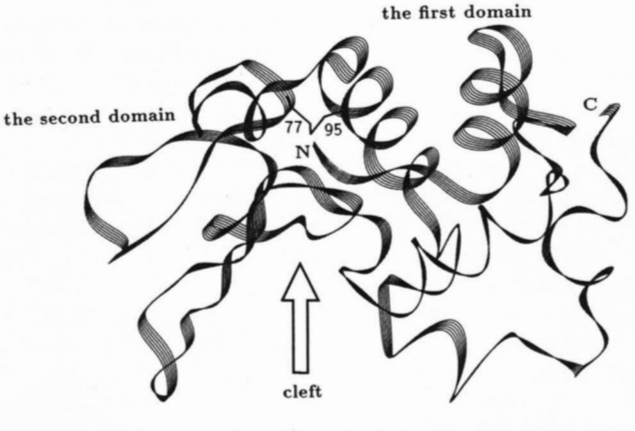

Fig. 1. Ribbon representation of the main chain trace of human lysozyme. The mutation site of the disulfide bond between Cys 77 and Cys 95 is explicitly shown. The disulfide bond connects 2 structural domains; the first domain is the right lobe in the figure containing Cys 95 (residues 1–39 and 89–129) and the second domain is the right lobe containing Cys 77 (residues 40–88). The assignment of the domains is after Brooks and Karplus (1985). A cooperative fluctuation between 2 domains causes the hinge-bending motion resulting in opening and closing the cleft. It is noted that the mutation site is at the hinge region.

Various disulfide mutants have been investigated for their effect on the thermal stability of proteins (Sauer et al., 1986; Villafranca et al., 1987; Matsumura et al., 1989; Mitchinson & Wells, 1989). Calorimetric experiments on the wild type and C77A/C95A showed that the removal of the disulfide bond reduces the thermal stability ($\Delta T_m = -14.5$ °C and $\Delta\Delta G =$ 4.6 kcal/mol at 57 °C; Kuroki et al., 1992). It is beyond the scope of the present study to discuss the stability based on the thermodynamic difference between the native and denatured states. However, detailed analyses of structural changes in the native state, not only static but also dynamic, would give an insight into the structural role of the disulfide bond.

Refinements of the wild-type and C77A/C95A structures were performed on 1.5-Å-resolution data obtained from isomorphous crystals, sufficient for the analysis of small changes in the coordinates and the dynamic structure.

Results and discussion

Results of normal mode refinements

Normal mode refinements were performed on X-ray data from the isomorphous crystals of the wild type and the C77A/C95A mutant. Data concerning the crystals used are summarized in Table 1. Superposition of the corresponding main chain atoms results in a translation by (0.01, −0.08, −0.19) Å along the (a, b, c) axes without rotation. The isomorphous condition eliminates the possible effects of lattice-induced shifts in the molecules.

Table 2 summarizes the results of normal mode refinements. As shown previously (Kidera et al., 1992), the normal mode refinement does not cause a significant shift relative to the coordinates given by the isotropic B-factor refinement. The RMS displacements of all atoms are 0.06 and 0.07 Å for the wild type and C77A/C95A, respectively. This means that the extension of the thermal factor from the isotropic to the anisotropic model has little influence on the average coordinates. The final R-factor for C77A/C95A stays about 2% higher than that of the wild type. This is because of the smaller number of reflec-

Table 1. *Data on crystals and X-ray diffraction measurements*

	Wild type	C77A/C95A
Crystals		
Space group	$P2_12_12_1$	$P2_12_12_1$
a, b, c (Å)	56.53, 60.83, 33.89	56.54, 60.67, 33.60
Crystal size (mm)	0.7 × 0.4 × 0.2	0.5 × 0.3 × 0.3
Diffraction measurements		
Number of crystals used	1	1
Incident beam	MoKα	MoKα
Resolution range (Å)	∞ to 1.46	∞ to 1.46
Number of independent reflections	16,902 (>3σ)	15,732 (>2σ)
Ratio to possible reflections	0.805	0.765
R_{merge}[a] (%)	4.19	5.18

[a] $R_{merge} = \Sigma|I - \langle I\rangle|/\Sigma\langle I\rangle$.

Table 2. *Results of the normal mode refinements*

	Wild type	C77A/C95A
Number of reflections used	15,310 (6.0–1.5 Å)	14,435 (6.0–1.5 Å)
Number of internal modes[a]		
M	100	100
m	43	43
Number of external modes[b]	6	6
Number of solvent molecules	148	111
Total number of variables for thermal factors[c]	1,320	1,246
R-factor[d] (%)	15.19	17.11
Intramolecular interaction energies (kcal/mol)[e]		
Bond (Å)	53 (0.012)	46 (0.011)
Angle (degrees)	112 (1.9)	125 (2.0)
Torsion (degrees)	157 (17)	163 (18)
Improper torsion (degrees)	25 (2.6)	25 (2.6)
Lennard-Jones	−1,114	−1,114
Electrostatic	−4,030	−3,980
Hydrogen bond	−130	−113
Intermolecular interaction energies (kcal/mol)[e]		
Lennard-Jones	−165	−147
Electrostatic	−11	43
Hydrogen bond	−14	−13
Thermal factor restraints[f]		
NM-restraint R (Å^4)	0.031	0.044
NM-restraint C	0.020	0.021

[a] The model comprises the M lowest frequency normal modes out of the 629 modes. For the m lowest frequency modes among the M modes, we consider the mode coupling.

[b] The external modes are those of the TLS model (Schomaker & Trueblood, 1968).

[c] The total number of the variables describing thermal factors is $M + m(m-1)/2 + 21 + 2$(number of solvents). The third term is for the external modes. The last term is of B-factors and occupancies for solvents.

[d] $\Sigma ||F_{obs}| - |F_{calc}|| / \Sigma |F_{obs}| \times 100$.

[e] The energy values are those of AMBER for united atoms (Weiner et al., 1984), and those in parentheses are the average deviations from the ideal values. The electrostatic energy is calculated with the dielectric constant $\epsilon = 2$.

[f] NM-restraint R and NM-restraint C are the RMS values of the normal mode restraints given by the first and the second term of Equation 29 in the paper by Kidera and Gō (1992).

tions from the C77A/C95A crystal, even after extending the data collection from $I > 3\sigma$ to $I > 2\sigma$. This difference in crystal quality affects the interpretation of dynamic structure in these proteins (see below). Statistics concerning the geometry are given in the form of interaction energies and deviations from the ideal values. Both the wild type and C77A/C95A are well within the range of known high-resolution X-ray structures in the Protein Data Bank (Bernstein et al., 1977).

Except near the mutation sites and in disordered side chains, there are few significant shifts in the average structure. RMS displacements between the 2 structures are 0.14 Å for the main chain atoms and 0.62 Å when all atoms are considered. The RMS displacement for all atoms whose B-factors are less than 20 (866 atoms out of 1,027 atoms) is only 0.24 Å.

The models of solvent structures for the wild type and C77A/C95A are in good agreement with each other, especially for those where $B < 20$ Å^2. Out of the 111 solvent sites found in C77A/C95A, 80 (which include 2 chloride ions) are within 1.0 Å of the corresponding solvent sites in the wild type. Their B-values are 21.2 ± 7.7 Å^2, whereas the other 31 solvent atoms have much higher B-factors of 31.0 ± 9.1 Å^2. Average occupancies of the 111 solvent atoms in C77A/C95A ($=0.85 \pm 0.18$) are on the same order as those of the 148 solvent atoms in the wild-type lysozyme ($=0.83 \pm 0.19$).

Mutation effects on the dynamic structure

Table 3 gives a summary of the dynamic structures of the wild type and C77A/C95A determined from normal mode refinement. The most characteristic feature found in the dynamic structure is that highly anisotropic internal fluctuations are masked by large and isotropic external terms (Kidera et al., 1992). The statistics in Table 3 for both the wild type and C77A/C95A obviously show these features: $(\langle \Delta r^2 \rangle^{1/2})_{internal} / (\langle \Delta r^2 \rangle^{1/2})_{external} = 0.81$ and 0.78 for the wild type and C77A/C95A, respectively. This indicates that a proper subtraction of the external terms from the apparent thermal factors is crucially important in the interpretation of mutation effects on the dynamic structure. The influence of disulfide bond removal on the dynamic structure appears in the differences in the values of $\langle \Delta r^2 \rangle^{1/2}$ and anisotropy listed in Table 3. While the total apparent fluctuations in C77A/C95A are larger in amplitude and anisotropy than those of the wild type $((\langle \Delta r^2 \rangle^{1/2})_{C77A/C95A} - (\langle \Delta r^2 \rangle^{1/2})_{wild} = 0.036$ Å, and $(\text{anisotropy})_{C77A/C95A} - (\text{anisotropy})_{wild} = 0.042)$, the major contribution to these differences comes from the external terms. The actual difference in the internal fluctuations is very small $((\langle \Delta r^2 \rangle^{1/2})_{C77A/C95A} - (\langle \Delta r^2 \rangle^{1/2})_{wild} = 0.009$ Å, and $(\text{anisotropy})_{C77A/C95A} - (\text{anisotropy})_{wild} = 0.007)$.

This can be seen more clearly in Figure 2, which shows residue profiles of $\langle \Delta r^2 \rangle^{1/2}$. The mutant C77A/C95A shows an increase in the apparent fluctuations at most residue sites (Fig. 2A). However, when the total fluctuations are decomposed into the external and internal fluctuations, most differences, except those near the mutation sites, can be ascribed to the external terms shown in Figure 2B. The external terms are determined mainly by the contributions that are sensitive to crystal quality, such as lattice disorder and diffuse scattering. Therefore, one can conclude that most of the differences in the apparent thermal factors originate from the difference in the external terms caused by the mutation effect on crystal quality. This effect can also be seen in the number of diffractions and the final R-factors given in Tables 1 and 2.

An amino acid substitution usually causes only a small change in the average structure localized at the mutation site (Matthews, 1987). Localization occurs also in the response of the dynamic structure to the mutation. It can be seen in Figure 2C that the effect of the mutation on the internal fluctuations is localized mostly at one of the mutation sites, the long loop region containing residue 77. The other mutation at residue 95, which is in an α-helix, does not exhibit any change in the dynamic structure. This difference between the 2 sites can be attributed to the difference in the secondary structures; one is a long loop whose flexibility is easily affected, whereas the other is a rigid α-helix that requires a large concerted fluctuation to increase flexibility.

Table 3. *Comparison of atomic fluctuations*

	$\langle \Delta r^2 \rangle^{1/2}$ (Å)[a]				Anisotropy[b]					C_{ij} (Å^2)[c]	
	Ave[d]	Dif[d]	C[d]	RMSD[d]	Ave[d]	Dif[d]	C[d]	RMSD[d]	$\langle \lvert \cos\theta \rvert \rangle$[e]	C[d]	RMSD[d]
Total[f]											
Wild type	0.695[g]				1.312						
	0.651				1.254						
	0.741				1.371						
C77A/C95A	0.731	0.036	0.916	0.073	1.354	0.042	0.757	0.150	0.860		
	0.686	0.035	0.898	0.061	1.300	0.046	0.732	0.114	0.874		
	0.776	0.035	0.913	0.084	1.409	0.038	0.743	0.180	0.845		
External[f]											
Wild type	0.533				1.263						
	0.530				1.255						
	0.537				1.272						
C77A/C95A	0.569	0.036	0.905	0.046	1.337	0.074	0.847	0.131	0.867		
	0.566	0.036	0.906	0.047	1.330	0.075	0.854	0.128	0.864		
	0.572	0.035	0.904	0.046	1.344	0.072	0.839	0.133	0.871		
Internal[f]											
Wild type	0.433				1.763						
	0.372				1.680						
	0.495				1.849						
C77A/C95A	0.442	0.009	0.911	0.079	1.770	0.007	0.554	0.387	0.820	0.772	0.023
	0.379	0.007	0.870	0.062	1.682	0.002	0.501	0.322	0.821		
	0.507	0.012	0.908	0.094	1.862	0.013	0.549	0.444	0.818		

[a] $\langle \Delta r^2 \rangle^{1/2}$ is the RMS fluctuation.

[b] Anisotropy is the measure of anisotropy of the atomic fluctuation defined by $[2\lambda_1/(\lambda_2 + \lambda_3)]^{1/2}$, where λ_1, λ_2, and λ_3 are the variances along the first, second, and third principal axes of the thermal ellipsoid, respectively.

[c] C_{ij} is the covariance $\langle \Delta r_i \cdot \Delta r_j \rangle$ for a pair of Cα atoms.

[d] Comparison between the wild type and C77A/C95A is done in terms of Ave, Dif, C, and RMSD, which are the average, the difference between 2 average values, the correlation coefficient, and the RMS difference, respectively, calculated over all the 1,027 corresponding non-hydrogen atoms for $\langle \Delta r^2 \rangle^{1/2}$ and anisotropy, and all the 8,385 Cα pairs for C_{ij}.

[e] $\langle \lvert \cos\theta \rvert \rangle$ is the average value of absolute cosine of the angle between 2 principal axes.

[f] Total means fluctuation involving both external and internal fluctuations. External is fluctuation that is determined by the external terms. Internal is internal fluctuation after external terms are removed from Total.

[g] There are 3 entries for each term: top for all atoms, middle for main chain atoms, and bottom for side chain atoms.

Figure 3 shows the ORTEP drawings (Johnson, 1976) of thermal ellipsoids for Cα atoms that express the internal fluctuations. Comparing the sizes and directions of thermal ellipsoids of the wild type and C77A/C95A, we notice that the atomic fluctuations with large amplitude and large anisotropy are only affected between residues 60 and 80. In particular, we observe the conservation of fluctuations representing the hinge-bending motion, an opening and closing fluctuation between the 2 domains. Thermal ellipsoids around the cleft region do not show any significant change in amplitude and direction. These similarities in the dynamic structure imply that the disulfide bond connecting the 2 domains at the hinge region does not play an important role in the hinge-bending motion. This finding may have some relevance to the experimental result showing that C77A/C95A has the same lytic activity as the wild type (Taniyama et al., 1988).

More detailed pictures of the dynamic structures at the mutation sites are given in Figure 4. They are the ORTEP drawings of the internal fluctuations of the main chain atoms of residues 62–65, 70–80, and 90–98. In C77A/C95A, residues 70–80 show amplified fluctuations with anisotropy almost identical to those in the wild type; the average value of absolute cosine between the principal axes of the corresponding thermal ellipsoids ($\langle \lvert \cos\theta \rvert \rangle$) is 0.95 for residues 71–77. For C77A/C95A, the atoms of residues 62–65 appear to move in concert with residues 70–80. On the other hand, the α-helix that includes the other mutation site (residue 95) has essentially the same dynamic structure, with small and rather isotropic fluctuations.

Another piece of valuable information from normal mode refinement is correlation in the interatomic fluctuations. In Figure 5, the covariance, C_{ij} ($=\langle \Delta \mathbf{r}_i \cdot \Delta \mathbf{r}_j \rangle$), for a pair of Cα atoms is plotted, where $\Delta \mathbf{r}_i$ represents the internal fluctuation vector of atom i around the average position. This shows how the fluctuations of 2 residues are correlated. Positive and negative correlation is for a pair of residues moving simultaneously in the same and opposite directions, respectively. The pattern of correlation in the residue fluctuations of C77A/C95A is essentially identical to that of the wild type. The covariance map shows that positive covariances for atoms covalently connected are at the diagonal region of the map, and negative covariances suggesting the hinge-bending motion are between pairs of residues, 40–80 and 100–130, located in different lobes. This map also suggests that the disulfide bond between residues 77 and 95 does not play a substantial role in the hinge-bending motion. The

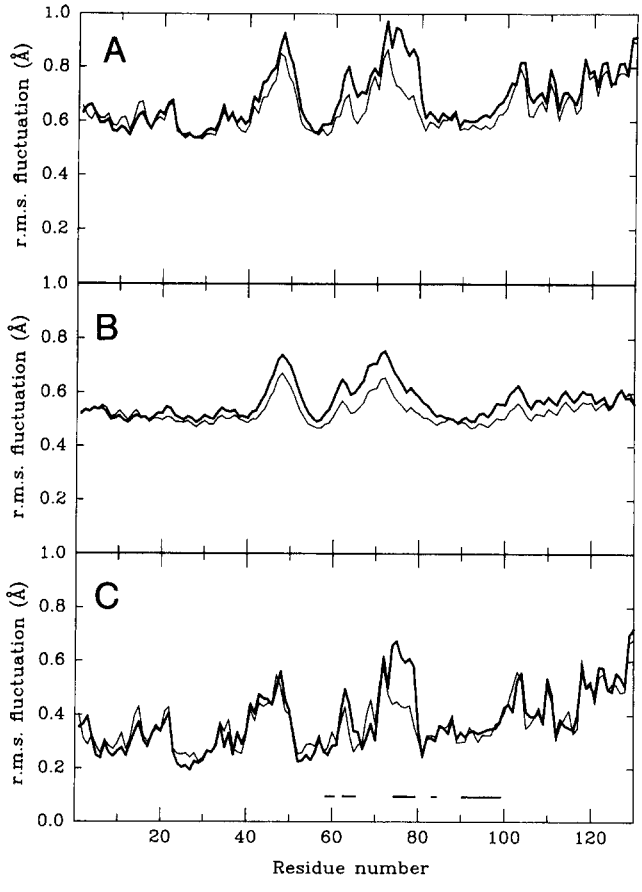

Fig. 2. Comparisons of RMS fluctuations $\langle \Delta r^2 \rangle^{1/2}$ of main chain atoms between the wild-type human lysozyme (thin curves) and the C77A/C95A mutant (thick curves). The values of $\langle \Delta r^2 \rangle^{1/2}$ are determined by normal mode refinement and averaged within each residue. **A:** Total apparent fluctuations. **B:** External terms. **C:** Internal fluctuations. These values are related by $\langle \Delta r^2 \rangle_{\text{total}} = \langle \Delta r^2 \rangle_{\text{external}} + \langle \Delta r^2 \rangle_{\text{internal}}$. The horizontal lines in C indicate the residues near the mutation site; a residue is considered close to the mutation site when any atom in the residue is within 8 Å from the midpoint between Sγ77 and Sγ95.

difference in the map can be found mainly around residues 70–80. The increase in the fluctuations at residues 70–80 of C77A/C95A is reflected in the increase of covariances involving these residues.

Role of the disulfide bond in the structure of human lysozyme

Substitution of Cys 77 and Cys 95 by 2 alanines is expected to affect protein conformation and dynamics through 2 possible mechanisms: cleavage of a disulfide bond connecting 2 domains, and cavity formation due to replacement of 2 bulky sulfur atoms by 2 small hydrogen atoms. The response of the protein structure to accommodate these changes is discussed below.

Role in the dynamic structure

How does the cleavage of the disulfide bond between Cys 77 and Cys 95 contribute to increasing fluctuations of the loop region at residues 70–80? A possible mechanism is the following. In the wild type, the disulfide bridge restricts the fluctuations around the loop region by connecting the loop to a less flexible α-helix. In C77A/C95A, the cleavage of the disulfide crosslink frees the loop from the α-helix, so increasing its flexibility. To examine this possibility, we evaluated the correlation coefficient, CC, of the fluctuations between Cα77 and Cα95,

$$CC = \frac{\langle \Delta \mathbf{r}_{C\alpha 77} \cdot \Delta \mathbf{r}_{C\alpha 95} \rangle}{(\langle \Delta r^2_{C\alpha 77} \rangle \langle \Delta r^2_{C\alpha 95} \rangle)^{1/2}}, \quad (2)$$

before and after cleavage of the disulfide bond. Normal mode refinement gives the CC values 0.35 and 0.26 for the wild type and C77A/C95A, respectively.

The values of CC for Cα atom pairs that are sequentially more than 10 residues apart but located within 6 Å are found to be 0.25 ± 0.18, whereas the CC values for Cα atoms located over 10 residues apart and over 6 Å apart are −0.05 ± 0.14. These values mean that a pair of spatially neighboring Cα atoms have a positive correlation coefficient for their fluctuations. Therefore, the CC value of 0.35 in the wild type would appear to be merely due to the spatial proximity of these 2 atoms (the distance between Cα77 and Cα95 is 5.34 and 5.37 Å in the wild type and C77A/C95A, respectively). The CC values for 2 Cα atoms separated by 1, 2, 3, and 4 residues sequentially are 0.73 ± 0.12, 0.53 ± 0.16, 0.42 ± 0.21, and 0.32 ± 0.22, respectively. The correlation CC = 0.35 is as weak as the correlation between the Cα atoms 4 residues apart or 12 atoms apart. This weak correlation was obtained also in the molecular dynamics simulation of wild-type human lysozyme in water (M. Saito, pers. comm.), where the correlation coefficient for the motion between Cα77 and Cα95 is 0.31 for the OPLS parameter (Jorgensen & Tirado-Rives, 1988).

This small value for CC in the wild type indicates that the loop containing residue 77 is not affected by the restriction of the disulfide bond and fluctuates almost independently of the α-helix even in the wild type. Therefore, the mechanism proposed above is not correct in the present case, and the cleavage of the disulfide bond is not the main cause of the increase in fluctuations of the loop region.

Instead of removing a disulfide bond, proteins with artificial crosslinks have been used to investigate the structural role of crosslinks in proteins. These include chemically crosslinked ribonuclease A (Weber et al., 1985), disulfide mutants of dihydrofolate reductase (Villafranca et al., 1987), and phage T4 lysozyme (Pjura et al., 1990). Comparisons of thermal factors before and after crosslinking showed that the engineered crosslinks actually increase *B*-factors of the crosslinked residues in these proteins. These observations agree with our conclusion that crosslinking does not impose a restriction on the fluctuations of the crosslinked residues.

In total, there are 4 disulfide bonds in lysozyme (Kinemage 1). The other 3 disulfide bonds give similar but slightly larger values for the correlation coefficients: CC = 0.43, 0.38, and 0.41 for Cα6–Cα128, Cα30–Cα116, and Cα65–Cα81, respectively. These values remain unaffected by the mutation of C77A/C95A. The CC value for Cα77–Cα95 (0.35 in the wild type) is the smallest of the four because this connects 2 structural domains, each of which fluctuates independently. The other 3 disulfide bonds are within the domains.

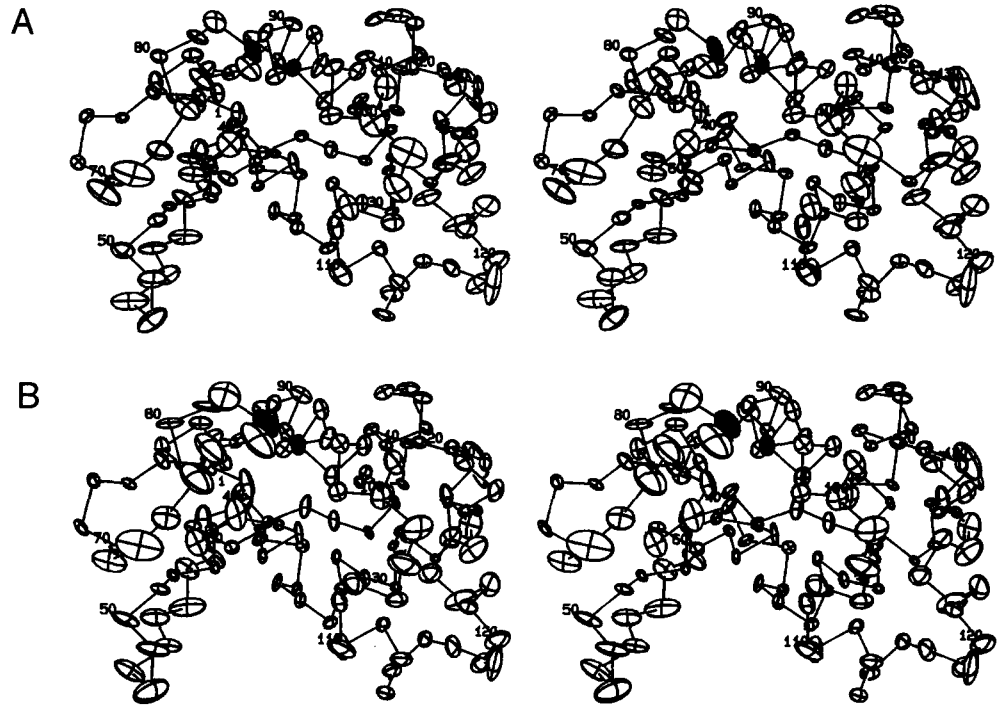

Fig. 3. Stereo ORTEP drawings of thermal ellipsoids (at a 3σ level) for the internal fluctuations of Cα. **A:** Wild type. **B:** C77A/C95A. Cα77 and Cα95 are indicated by the shaded atoms.

Role in the static average structure

The implication of replacing Cys 77 and Cys 95 with 2 alanines can be explained by the coordinate shifts (the change in the average static structure) caused by the mutation. Figure 6 and Kinemage 1 show the coordinate shift of the main chain atoms after the superposition of the 2 structures. The changes in the coordinates are localized around the mutation sites. Unlike the change in the dynamic structure, the coordinates of the α-helix, including Cys 95, are also influenced by the mutation. The main chain trace in Figure 6B shows the direction and magnitude of the changes. The chain appears to shift cooperatively in order to fill a void caused by the replacement of 2 bulky sulfur atoms by 2 small hydrogen atoms.

More quantitative analyses are given in Table 4, where C77A/C95A is compared with "Model," which is a model mutant having the same coordinates as the wild type but with the same amino acid sequence as C77A/C95A. To avoid an ambiguity in the comparison of superposed structures and to see if atoms really move toward the mutation site, we calculated the distance between the 2 domains, d_{domain}, by the average distance defined as:

$$d_{\text{domain}} = \frac{1}{N} \sum_{i,j}{}' d_{ij}, \tag{3}$$

where the summation is for N distances d_{ij}, and atoms i and j are both within spheres given in Table 4 and in the first and the second domain, respectively (see Fig. 1). The domain distance has decreased by about 0.2 Å after the mutation. Because the difference of the average distances within each domain is 1 order smaller, −0.03 Å, this shift appears to result from the 2 domains approaching each other cooperatively without changing their individual structures appreciably. This can also be observed by looking at the radius of gyration, $\langle S^2 \rangle^{1/2}$, of the atoms surrounding the mutation site:

$$\langle S^2 \rangle^{1/2} = \left[N^{-2} \sum_{i<j} d_{ij}^2 \right]^{1/2}. \tag{4}$$

The value of $\langle S^2 \rangle^{1/2}$ for C77A/C95A decreases by 0.06–0.09 Å. A similar decrease was observed as the response to the removal of a disulfide bond in a mutant of bovine pancreatic trypsin inhibitor in which Cys 30 and Cys 51 were replaced by Ala (Eigenbrot et al., 1990). This contraction produces tighter atom packing in C77A/C95A, which can be seen in the van der Waals interaction energy, E_{vdW}. We find that E_{vdW}(C77A/C95A) < E_{vdW}(model). It should be remarked that E_{vdW}(wild type) takes an even lower value than E_{vdW}(C77A/C95A) (−155.4 kcal/mol for a 10-Å sphere). However, it is not meaningful to compare the energy values of the wild type and its mutant unless the energy value of a reference state (e.g., the denatured state) could be evaluated correctly.

The atom packing should be reflected in the cavity volume. The cavity inside a protein is defined here by grid cubes (a 0.2-Å grid) that are located outside the van der Waals spheres of constituent atoms but are not accessible to solvent. As seen in Table 4, the removal of the 2 sulfur atoms creates a hole of 20 Å^3 inside lysozyme (the difference between the wild type and Model). Comparison between Model and C77A/C95A clearly shows that the atoms of C77A/C95A move toward the mutation site to fill the vacated hole and produce a tighter atom packing. The decrease in the cavity exceeds 20 Å^3 (−8 Å^3 difference

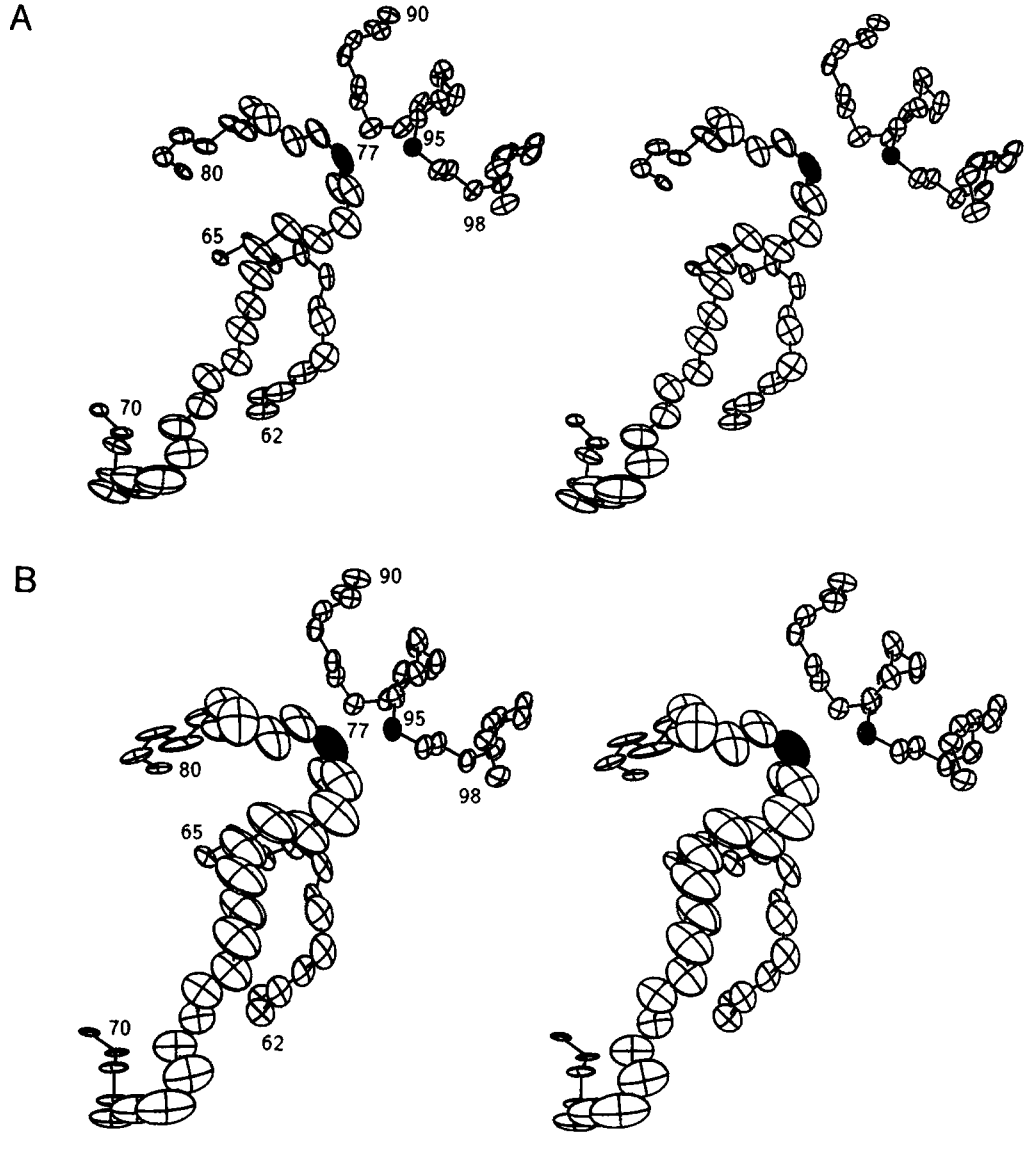

Fig. 4. Stereo ORTEP drawings of thermal ellipsoids (at a 1.8σ level) for the internal fluctuations of N, Cα, and C for residues 62–65, 70–80, and 90–98. **A:** Wild type. **B:** C77A/C95A. Cα77 and Cα95 are indicated by the shaded atoms.

from the wild type for the 9- and 10-Å spheres). However, when we look at the mutation site carefully, we notice that more than half of this volume still remains at the mutation site (13 Å^3). The tighter packing of the surrounding atoms cannot fill the cavity created by the mutation.

These coordinate shifts caused by the Cys to Ala mutations coincide with those of the "cavity-creating mutation" (Eriksson et al., 1992). Mutations that involve the replacement of bulky side chains with smaller side chains are characterized by: (1) the surrounding atoms moving toward the vacated space, (2) a cavity always remaining, and (3) the native state being destabilized. This coincidence with our observations, as well as the destabilization of the native state, leads us to conclude that the removal of the disulfide bond should be regarded as a cavity-creating mutation. The tighter atom packing in C77A/C95A may not be sufficient to compensate for the energy loss caused by the cavity remaining at the mutation site. Furthermore, both in the static average structure and in the dynamic structure, we find no evidence that the disulfide bond plays a role in preventing the 2 domains from parting from each other.

Materials and methods

Crystallization and X-ray diffraction data collection

The mutant C77A/C95A was prepared by the yeast expression system described by Taniyama et al. (1988). Wild-type human lysozyme was purchased from the Green Cross Corporation (Osaka, Japan). Crystals were grown in a solution containing 20 mg/mL protein, 2.5 M NaCl as a precipitant, and 30 mM sodium phosphate buffer, pH 6.0, at 13.0 °C. The data for crystals used in the experiments are summarized in Table 1. The crystals are in the orthorhombic space group $P2_12_12_1$ and are isomorphous with each other. X-ray diffraction intensity data

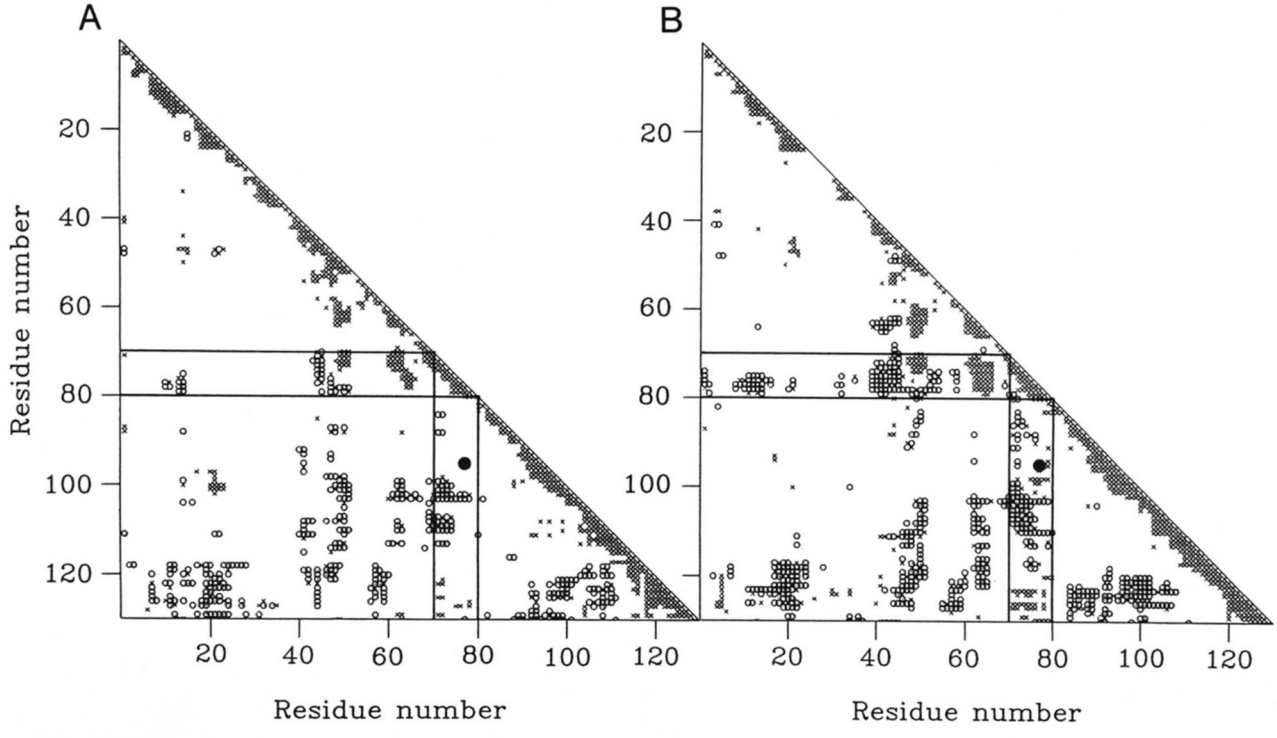

Fig. 5. Covariances in the atomic fluctuations $\langle \Delta \mathbf{r}_i \cdot \Delta \mathbf{r}_j \rangle$ between a pair of Cα atoms obtained by normal mode refinement. ×, Positive covariance > 0.04 Å^2; ○, negative covariance < −0.04 Å^2. The mutation sites, Cα77 and Cα95, are indicated by a filled circle. **A:** Wild type. **B:** C77A/C95A.

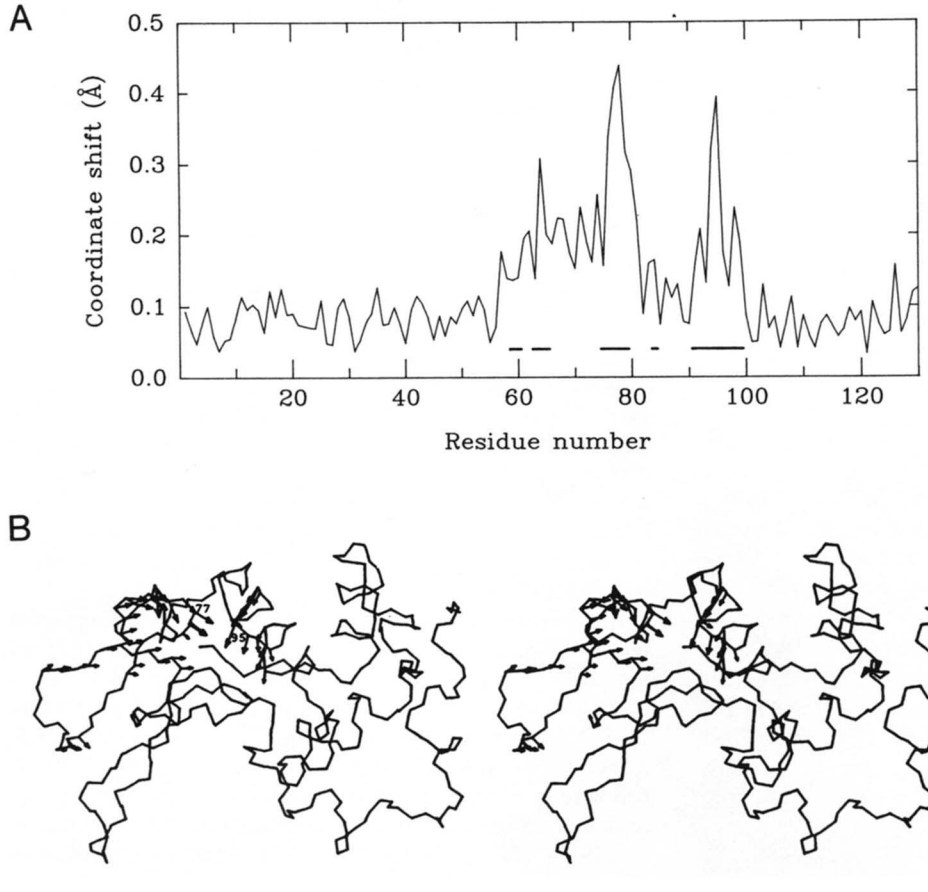

Fig. 6. Coordinate shift between the wild type and C77A/C95A. The main chain atoms of the residues 1–60 and 100–130 are superposed. The atoms of the residues 61–99 having large shifts are excluded from the superposition. **A:** RMS displacements are averaged within each residue. The thick horizontal lines indicate the residues near the mutation site (same as those in Fig. 2C). **B:** Vector representation of the atomic displacement of C77A/C95A from the wild type. All displacements of N, Cα, and C larger than 0.2 Å are given by arrows. Arrows are drawn 5 times larger than the actual displacements.

Table 4. *Comparison of the coordinates around the mutation sites*

Sphere[a] (Å)	Number of atoms[b] (united atom)	Domain distance[c] (Å)			Radius of gyration[f] (Å)			van der Waals energy[g] (kcal/mol)		
		Model[d]	C77A/C95A	Dif[e]	Model[d]	C77A/C95A	Dif[e]	Model[d]	C77A/C95A	Dif[e]
5	27	6.77	6.54	−0.23	4.19	4.10	−0.09	−5.5	−6.0	−0.5
6	54	7.92	7.70	−0.22	4.96	4.87	−0.09	−19.2	−20.4	−1.2
7	93	9.17	8.99	−0.18	5.67	5.60	−0.07	−46.0	−47.3	−1.3
8	131	10.04	9.85	−0.19	6.26	6.20	−0.06	−71.7	−74.1	−2.4
9	184	10.99	10.80	−0.19	6.95	6.87	−0.08	−110.6	−114.3	−3.7
10	241	11.89	11.73	−0.16	7.60	7.53	−0.07	−147.7	−151.8	−4.1

Sphere[a] (Å)	Number of atoms[h] (all atom)	Cavity volume[i] (Å^3)			
		Wild type	Model[d]	C77A/C95A	Dif[j]
5	59	190	210	203	−7 (13)
6	100	350	370	358	−12 (8)
7	157	514	534	518	−16 (4)
8	224	746	766	747	−19 (1)
9	298	1,035	1,055	1,027	−28 (−8)
10	380	1,356	1,376	1,348	−28 (−8)

[a] Comparison is done for the atoms within a sphere of a specified radius whose center is at the midpoint between Sγ77 and Sγ95 of the wild type.
[b] Number of atoms is the number of corresponding atoms (including polar hydrogen atoms) within a specified sphere. The coordinates of polar hydrogen atoms were those determined in the refinement.
[c] Domain distance d_{domain} is defined by Equation 3.
[d] Model has the same coordinates as the wild type, but Sγ77 and Sγ95 are each replaced by a hydrogen atom.
[e] Dif is the difference C77A/C95A − Model.
[f] The radius of gyration $\langle S^2 \rangle^{1/2}$ for atoms located within a specified sphere is defined by Equation 4.
[g] van der Waals energy is calculated by counting all interactions within a specified sphere with the AMBER parameter for polar hydrogen atoms (Weiner et al., 1984).
[h] Number of atoms is the number of corresponding atoms (including all hydrogen atoms) within a specified sphere. The coordinates of nonpolar hydrogen atoms were generated by the molecular simulation program PRESTO (Morikami et al., 1992) with the "all atom" AMBER parameter set (Weiner et al., 1986) and followed by an energy minimization with the coordinates of all the other atoms fixed.
[i] Cavity volume is the volume occupied by grid cubes that are recognized as cavity cubes inside the protein. A cavity cube inside a protein is defined by a grid cube (a 0.2-Å grid size is used) that is located entirely outside the van der Waals radii of any atom, including hydrogen atoms, and has no solvent accessibility (i.e., no room for a water molecule). Solvent accessibility is evaluated by the method of Shrake and Rupley (1973).
[j] The quantity in parentheses is the difference C77A/C95A − Wild type.

were collected using only 1 crystal for each molecule up to 1.46 Å resolution by an automated oscillation camera system equipped with an imaging plate (DIP100, MAC Science, Japan; Miyahara et al., 1986; Amemiya et al., 1988) on a rotating anode generator with MoKα radiation (λ = 0.7106 Å).

Refinement

The process of normal mode refinement consists of 3 steps: (1) the conventional *B*-factor refinement, (2) normal mode analysis, and (3) normal mode refinement. Refinements by PROLSQ (Hendrickson, 1985) were first performed against the diffraction data of a 6.0–1.5-Å resolution range. The initial models are taken from the X-ray structures of wild type and C77A/C95A, respectively, refined against 1.8-Å-resolution data (Inaka et al., 1991; the mutant structure is 1LHM in the Protein Data Bank). Solvent molecules were identified from the peaks of electron density satisfying the criterion

$$O_i \sum_k^5 a_{ki}\left(\frac{4\pi}{b_{ki}+B_i}\right)^{2/3} > \sum_k^5 a_{ki}\left(\frac{4\pi}{b_{ki}+50}\right)^{2/3}, \qquad (5)$$

where O_i is the occupancy of the atom i that is also refined during the refinement, a_{ki} and b_{ki} are the parameters of the 4 Gaussian-type atomic scattering factors for atom i ($b_{5i} = 0$), and B_i is the isotropic *B*-factor. This criterion means that the peak electron density of a solvent atom should be larger than that of $O_i = 1$ and $B_i = 50$ Å^2. The results of the refinement were then transferred to the normal mode refinement program, NM-REF (Kidera & Gō, 1992), to perform a further *B*-factor refinement. By this additional refinement step, the coordinates of the least-squares refinement from PROLSQ were adjusted to those of the energy refinement of NM-REF, which uses the AMBER parameters for united atoms (Weiner et al., 1984).

In the next step, 2 sets of normal modes for the wild type and C77A/C95A were calculated in the dihedral angle space. This step includes the regularization of bond lengths and bond angles to the ECEPP/2 standards (Némethy et al., 1983) by DADAS (Braun & Gō, 1985), energy minimization by FEDER (Wako & Gō, 1987), and normal mode analysis (Gō et al., 1983), followed by a transformation of the normal mode eigenvectors of dihedral angles to those of Cartesian coordinates (Noguti & Gō, 1983). The RMS displacements of all non-hydrogen atoms

after the energy minimization were 1.44 and 1.31 Å for the wild type and C77A/C95A, respectively.

The results of the preceding steps, coordinates, *B*-factors and occupancies for solvent atoms, and normal modes were used in normal mode refinement by NM-REF. The model of the dynamic structure in the normal mode refinement is the same as in the previous study (Kidera et al., 1992). This is described by the normal modes representing the internal fluctuations that comprise the 100 lowest frequency normal modes with mode coupling among the 43 lowest frequency modes as well as the 6 modes for the external fluctuations (the TLS model; Schomaker & Trueblood, 1968). Solvent atoms are described by the isotropic *B*-factors and the external normal modes with variable occupancies. The details of the method of normal mode refinement are described in Kidera and Gō (1992). The results of the refinement are summarized in Table 2.

Significance test of the model of dynamic structure

To test the significance of the dynamic structure determined by normal mode refinement, we calculated the free *R*-factor, R_{free}, (Brünger, 1992, 1993). Because R_{free} is not subjected to overfitting the model, it should be a good measure for the significance test. As the normal mode refinement does not result in a significant shift in coordinates from the *B*-factor model, we calculated values of R_{free} only for the dynamic structure.

First, we define a test set by randomly selecting 10% of the observed reflections. Refinement of the remaining 90% was made using either the *B*-factor model or the normal mode model. The former adjusts *B*-factors of all non-hydrogen protein and solvent atoms as well as occupancies of solvent atoms. The thermal factor restraints used were 1.5 and 2.0 $Å^2$ for bonded and angle-related backbone atoms, and 2.0 and 2.5 $Å^2$ for bonded and angle-related side chain atoms. The normal mode model refines the variance-covariance of protein normal modes as well as *B*-factors and occupancies of solvent atoms. The initial values are set to be 15 $Å^2$ for *B*-factors, 1.0 for occupancies, and variance-covariance given by the theoretical normal modes (all covariances are set to be 0). The coordinates are fixed as those refined against all the diffraction data. After 50 cycles of quasi-Newton minimization, we calculated both R_{free} and the conventional *R*-factor for the test set. These results are summarized in Table 5.

Table 5. *Free R-factor for the dynamic structure*

	R_{free}[a]	R[b]	$R_{free} - R$
Wild type			
Isotropic *B*-factor model	17.40	16.00	1.40
Normal mode model	16.02	15.33	0.69
C77A/C95A			
Isotropic *B*-factor model	19.67	17.99	1.68
Normal mode model	17.84	17.13	0.71

[a] Free *R*-factor is calculated for the 10% diffraction data that were not used in the refinement.

[b] *R*-factor is calculated against the remaining 90% diffraction data to which the model was refined.

Although these values are derived from only 1 test set, it is clearly seen, both in the wild type and in the mutant, that the difference between the values of R_{free} and the conventional *R*-factor ($R_{free} - R$) is much smaller in the normal mode model than in the isotropic *B*-factor model. This means that the normal mode model is much more resistant to overfitting than the conventional isotropic *B*-factor model. In this sense, the normal mode refinement is a reliable method. One of the reasons for this reliability is that we ensured that the number of adjustable parameters in the model do not exceed those in the isotropic *B*-factor model. More essentially, it should be due to the basic assumption of the normal mode model that only takes into account the fluctuations occurring in the important conformational subspace defined by the low frequency normal modes (Kidera & Gō, 1992). The remaining fluctuations are regarded by the model either as external motions or as noise that raises the *R*-factor.

Acknowledgments

We thank Dr. M. Saito for providing us with the result of a simulation and Drs. S. Hayward and M.B. Swindells for reading the manuscript carefully. This work was supported in Kyoto University by grants from MESC and HFSP to N.G.

References

Amemiya Y, Matsushita T, Nakagawa A, Satow Y, Miyahara J, Chikawa J. 1988. Design and performance of an imaging plate system for X-ray diffraction study. *Nuclear Instrum Methods A 266*:645–653.

Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 112*:535–542.

Braun W, Gō N. 1985. Calculation of protein conformations by proton–proton distance constraints. A new efficient algorithm. *J Mol Biol 186*:611–626.

Brooks B, Karplus M. 1985. Normal modes for specific motions of macromolecules: Application to the hinge-bending mode of lysozyme. *Proc Natl Acad Sci USA 82*:4995–4999.

Brünger AT. 1992. Free *R* value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature 355*:472–475.

Brünger AT. 1993. Assessment of phase accuracy by cross validation: The free *R* value. Method and application. *Acta Crystallogr D 49*:24–36.

Diamond R. 1990. On the use of normal modes in thermal parameter refinement: Theory and application to the bovine pancreatic trypsin inhibitor. *Acta Crystallogr A 46*:425–435.

Eigenbrot C, Randal M, Kossiakoff AA. 1990. Structural effects induced by removal of a disulfide-bridge: The X-ray structure of the C30A/C51A mutant of basic pancreatic trypsin inhibitor at 1.6 Å. *Protein Eng 3*: 591–598.

Eriksson AE, Baase WA, Zhang XJ, Heinz DW, Blaber M, Baldwin EP, Matthews BW. 1992. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science 255*:178–183.

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

Finzel BC, Salemme FR. 1985. Lattice mobility and anomalous temperature factor behaviour in cytochrome *c'*. *Nature 315*:686–688.

Frauenfelder H, Petsko GA, Tsernoglou D. 1979. Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. *Nature 280*:558–563.

Gibrat JF, Gō N. 1990. Normal mode analysis of human lysozyme: Study of the relative motion of the two domains and characterization of the harmonic motion. *Proteins Struct Funct Genet 8*:258–279.

Gō N, Noguti T, Nishikawa T. 1983. Dynamics of a small globular protein in terms of low-frequency vibrational modes. *Proc Natl Acad Sci USA 80*:3696–3700.

Hendrickson WA. 1985. Stereochemical restrained refinement of macromolecular structures. *Methods Enzymol 115*:252–270.

Inaka K, Taniyama Y, Kikuchi M, Morikawa K, Matsushima M. 1991. The crystal structure of a mutant human lysozyme C77/95A with increased secretion efficiency in yeast. *J Biol Chem 266*:12599–12603.

Johnson CK. 1976. Ortep II, a FORTRAN thermal ellipsoid plot program for crystal structure illustrations. *Oak Ridge National Laboratory Report ORNL-5138.*

Jorgensen WL, Tirado-Rives J. 1988. The OPLS potential function for proteins. Energy minimization for crystals of cyclic peptides and crambin. *J Am Chem Soc 110*:1657–1666.

Kidera A, Gō N. 1990. Refinement of protein dynamic structure: Normal mode refinement. *Proc Natl Acad Sci USA 87*:3718–3722.

Kidera A, Gō N. 1992. Normal mode refinement: Crystallographic refinement of protein dynamic structure I. Theory and test by simulated diffraction data. *J Mol Biol 225*:457–475.

Kidera A, Inaka K, Matsushima M, Gō N. 1992. Normal mode refinement: Crystallographic refinement of protein dynamic structure II. Application to human lysozyme. *J Mol Biol 225*:477–486.

Kuroki R, Inaka K, Taniyama Y, Kidokoro S, Matsushima M, Kikuchi M, Yutani K. 1992. Enthalpic destabilization of a mutant human lysozyme lacking a disulfide bridge between Cys-77 and Cys-95. *Biochemistry 31*:8323–8328.

Matthews BW. 1987. Genetic and structural analysis of the protein stability problem. *Biochemistry 26*:6885–6888.

Matsumura M, Becktel WJ, Levitt M, Matthews BW. 1989. Stabilization of phage T4 lysozyme by engineered disulfide bonds. *Proc Natl Acad Sci USA 86*:6562–6566.

McCammon JA, Gelin B, Karplus M, Wolynes PG. 1976. The hinge-bending mode in lysozyme. *Nature 262*:325–326.

Mitchinson C, Wells JA. 1989. Protein engineering of disulfide bonds in subtilisin BPN'. *Biochemistry 28*:4807–4815.

Miyahara J, Takahashi K, Amemiya Y, Kamiya N, Satow Y. 1986. A new type of X-ray area detector utilizing laser stimulated luminescence. *Nuclear Instrum Methods A 246*:572–578.

Morikami K, Nakai T, Kidera A, Saito M, Nakamura H. 1992. PRESTO (protein engineering simulator): A vectorized molecular mechanics program for biopolymers. *Comput Chem 16*:243–248.

Némethy G, Pottle MS, Scheraga HA. 1983. Energy parameters in polypeptides. 9. Updating of geometrical parameters, nonbonded interactions, and hydrogen bond interactions for the naturally occurring amino acids. *J Phys Chem 87*:1883–1887.

Noguti T, Gō N. 1983. Dynamics of native globular proteins in terms of dihedral angles. *J Phys Soc (Jpn) 52*:3283–3288.

Pjura PE, Matsumura M, Wozniak JA, Matthews BW. 1990. Structure of a thermostable disulfide-bridge mutant of phage T4 lysozyme shows that an engineered cross-link in a flexible region does not increase the rigidity of the folded protein. *Biochemistry 29*:2592–2598.

Sauer RT, Hehir K, Stearman RS, Weiss MA, Jeitler-Nilsson A, Suchanek EG, Pabo CO. 1986. An engineered intersubunit disulfide enhances the stability and DNA binding of the N-terminal domain of λ repressor. *Biochemistry 25*:5992–5998.

Schomaker V, Trueblood KN. 1968. On the rigid-body motion of molecules in crystals. *Acta Crystallogr B 24*:63–76.

Sheriff S, Hendrickson WA. 1987. Description of overall anisotropy in diffraction from macromolecular crystals. *Acta Crystallogr A 43*:118–121.

Shrake A, Rupley JA. 1973. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *J Mol Biol 79*:351–371.

Soman J, Iismaa S, Stout CD. 1991. Crystallographic analysis of two site-directed mutants of *Azotobacter vinelandii* ferredoxin. *J Biol Chem 266*: 21558–21562.

Taniyama Y, Yamamoto Y, Nakao M, Kikuchi M, Ikehara M. 1988. Role of disulfide bonds in folding and secretion of human lysozyme in *Saccharomyces cerevisiae*. *Biochem Biophys Res Commun 152*:962–967.

Villafranca JE, Howell EE, Oatley SJ, Xuong NH, Kraut J. 1987. An engineered disulfide bond in dihydrofolate reductase. *Biochemistry 26*: 2182–2189.

Wako H, Gō N. 1987. Algorithm for rapid calculation of Hessian of conformational energy function of proteins by supercomputer. *J Comput Chem 8*:625–635.

Weber PC, Sheriff S, Ohlendorf DH, Finzel BC, Salemme FR. 1985. The 2-Å resolution structure of a thermostable ribonuclease A chemically cross-linked between lysine residue 7 and 41. *Proc Natl Acad Sci USA 82*:8473–8477.

Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, Profeta S Jr, Weiner P. 1984. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J Am Chem Soc 106*:765–784.

Weiner SJ, Kollman PA, Nguyen DT, Case DA. 1986. An all atom force field for simulations of proteins and nucleic acids. *J Comput Chem 7*:230–252.