ELSEVIER

Biophysical Chemistry 50 (1994) 25–31

Biophysical Chemistry

Dynamic structure of human lysozyme derived from X-ray crystallography: normal mode refinement

Akinori Kidera *[a], Masaaki Matsushima [a], Nobuhiro Gō [b]

[a] *Protein Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565, Japan*
[b] *Department of Chemistry, Faculty of Science, Kyoto University, Kyoto 606, Japan*

(Received 3 January 1994)

Abstract

X-ray crystallography provides a wealth of information about the dynamic as well as static protein structure. A new method of dynamic structure refinement of protein X-ray crystallography, normal mode refinement, is proposed. In this method, the Debye–Waller factor is expanded in terms of the low-frequency internal normal modes and external normal modes, whose amplitudes and couplings are optimized in the process of crystallographic refinement. The internal and external contributions to the atomic fluctuations can be separated. Also, anisotropic atomic fluctuations and their inter-atomic correlations can be determined experimentally even with a relatively small number of adjustable parameters. The method is applied to the analyses of experimental data of human lysozyme and its mutant, C77A/C95A, to reveal its dynamic structure.

Key words: Normal modes; Crystallographic refinement; Protein dynamics; Temperature factor; Human lysozyme

1. Introduction

X-ray crystallography provides information about the dynamic structure in the form of crystallographic temperature factors [1]. However, there remain two problems in the determination of the protein dynamics from temperature factors.

(1) The *isotropic B*-factors, conventionally used in protein crystallography, give only the magnitude of fluctuations but lose information of its direction and correlation.

(2) Various factors, such as static disorder [2] and rigid body external motion of a protein [3], which we call external terms, mix up with the real internal dynamics in temperature factors.

For the purpose of solving these problems, we introduced a method of dynamic structure refinement, normal mode refinement [4–6]. First, we will explain the principle of normal mode refinement focusing on how the two problems are solved. Then, two examples of normal mode refinement are presented; human lysozyme and its mutant (C77A/C95A). We will see the real inter-

* Corresponding author.

SSDI 0301-4622(94)00006-6

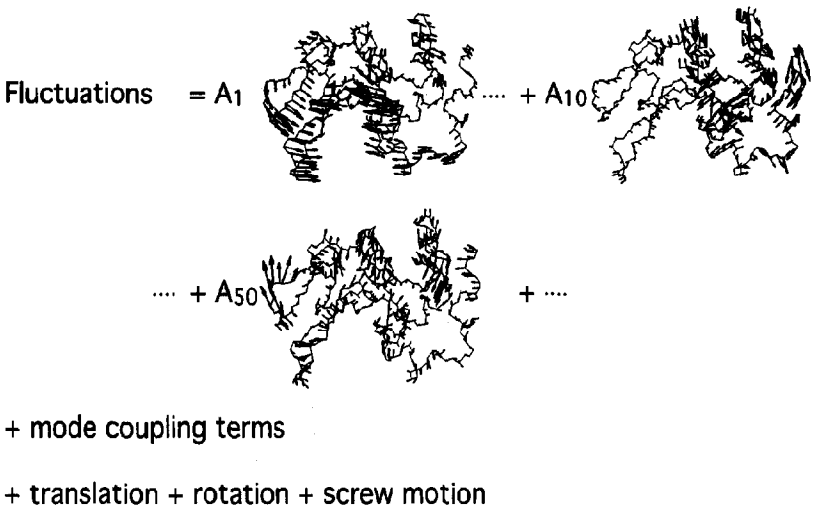

Fig. 1. A schematic explanation of normal mode refinement of the protein dynamic structure. Fluctuations are described by a sum of internal modes, mode coupling terms and external modes. Arrow representations of eigenvectors, ϕ_{jkm}, are given for the lowest, 10th lowest and 50th lowest internal modes. During the course of the refinement, amplitudes for these terms are optimized so as to give the best fit to the experimental data.

nal dynamics of the protein derived from X-ray diffraction data. The comparison with the mutant lysozyme will show the response of dynamic structure to the mutation.

2. Theory

In normal mode refinement [4,5], the structure factor, F_{calc}, is given by

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

where $\boldsymbol{q}$ $(=\{q_k\})$ is a reciprocal lattice vector, and $f_j(\boldsymbol{q})$ is the atomic structure factor for atom j, with $\boldsymbol{r}_j$ 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 [7] and σ_{mn} is covariance for

Table 1
Results of the normal mode refinements

	Wild-type	C77A/C95A
number of reflections used [a]	15310	14435
resolution range (Å)	6.0–1.5	6.0–1.5
number of internal modes [b]		
M	100	100
m	43	43
number of external modes [c]	6	6
number of solvent molecules	148	111
total number of variables for thermal factors [d]	1320	1246
R-factor [e] (%)	15.19	17.11
deviation from the ideal values [f]		
bond (Å)	0.012	0.011
angle (deg)	1.9	2.0
torsion (deg)	17	18
improper torsion (deg)	2.6	2.6

[a] The details of the diffraction experiments are described in ref. [10].
[b] 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.
[c] The external modes are those of the TLS model [8].
[d] The total number of the variables describing thermal factors is $M+\frac{1}{2}m(m-1)+21+2\times$(number of solvents). The third term is for the external modes. The last term is of B-factors and occupancies for solvents.
[e] $\Sigma||F_{obs}|-|F_{calc}||/\Sigma|F_{obs}\times 100$.
[f] The ideal values are those of AMBER for united atoms [11].

the internal or external normal modes. The summation is taken over the M lowest frequency internal normal modes as well as the 6 external normal modes of the TLS model [8]. 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 Eq. (1)) of low frequency normal modes.

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

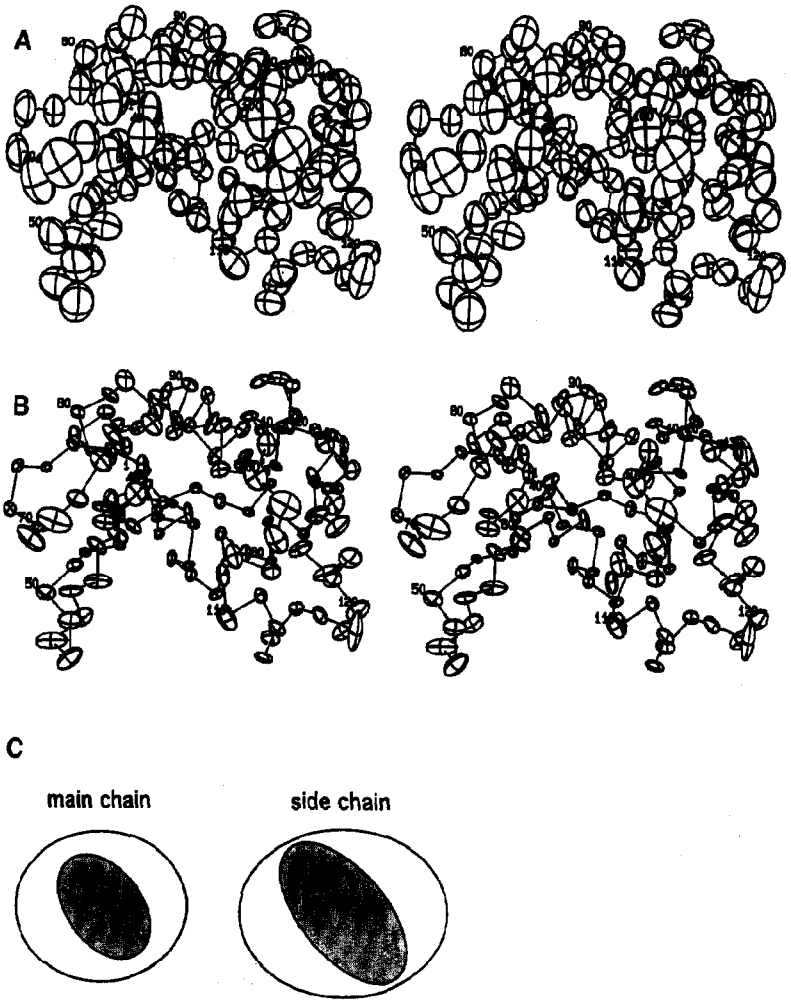

Fig. 2. Stereo ORTEP [11] drawings of thermal ellipsoids of C_α atoms (at a 3σ level) for the total apparent, A, and the internal, B, fluctuations. The average shapes of thermal ellipsoids for main chain atoms and side-chain atoms, C, are given in two dimensional representation. The values of $\langle \Delta r \rangle^{1/2}$ are 0.651 (0.741) and 0.372 (0.495) for the total and internal fluctuations, respectively, for main chain atoms (side chain atoms). Anisotropy of the atomic distribution is defined by the axial ratio $[2\lambda_1/(\lambda_2 + \lambda_3)]^{1/2}$, where λ_1, λ_2 and λ_3 are the variances along the first, second and third principal axes of a thermal ellipsoid. These values are 1.254 (1.371) and 1.680 (1.849) for the total and internal fluctuations, respectively. The average cosine of the angle between the directions of the total and internal fluctuations, $\langle \cos \theta \rangle$, is found to be 0.45 for both main chain atoms and side chain atoms. This value is almost the random level of 0.5.

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 (Fig. 1). Because the normal modes reflect the detailed covalent and three-dimensional structure of the molecule, atomic fluctuations are treated as being correlated and anisotropic unlike in the conventional isotropic B-factor model. It is noticed that intensities at Bragg angles do not contain information concerning inter-atomic correlation. Nevertheless, once σ_{mn} has been determined in the refinement, the inter-atomic correlation (between atoms i and j) can be described by

$$\langle \Delta r_{ik} \Delta r_{jl} \rangle = \sum_{m,n=1}^{M} \phi_{ikm} \phi_{jln} \sigma_{mn}. \qquad (2)$$

This means that not only the anisotropic intra-atomic distribution, but also inter-atomic correlation is inferred from the diffraction intensities.

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.

3. Computations

We applied Eq. (1) as an expression for the Debye–Waller factor in the later stage of crystallographic refinement. In the initial stage, the conventional isotropic B-factor refinement was done by PROLSQ [9]. Then, by using the structure thus determined, the normal mode analysis [7] was carried out to calculate normal mode

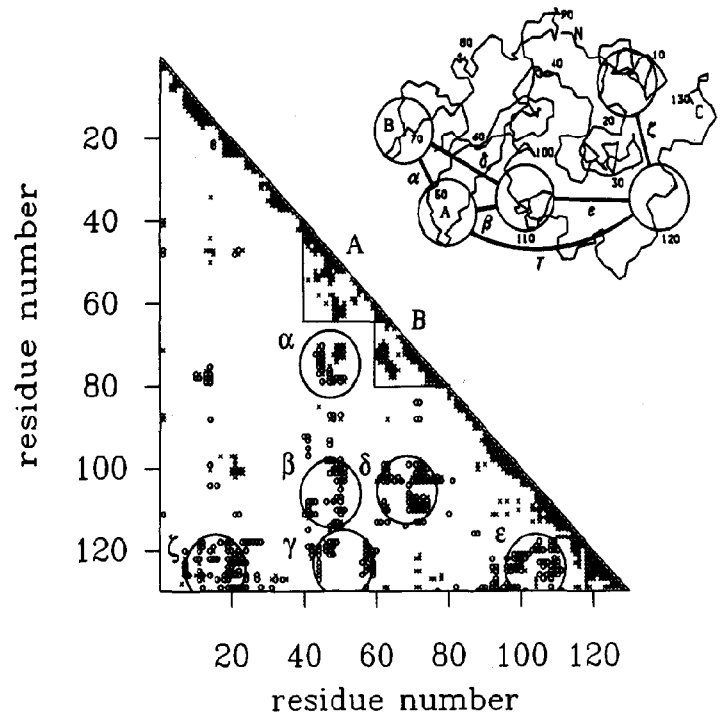

Fig. 3. Covariances in the atomic fluctuations $C_{ij}(=\langle \Delta r_i \cdot \Delta r_j \rangle)$ between a pair of C_α atoms i and j. (×), positive covariance > 0.04 Å^2; (○), negative covariance < -0.04 Å^2. The loci of highly negative covariances α–ζ correspond to the correlation between two segments illustrated in the figure of the main chain trace. Those of β, γ and δ suggest the existence of the hinge-bending motions. Triangular areas of highly positive covariances correspond to two loops of A and B.

eigenvectors ϕ_{jkm}, which was further followed by the normal mode refinement based on Eq. (1). In this stage, the 100 ($=M$) lowest frequency normal modes with their 903 covariances corresponding to the 43 lowest frequency modes for σ_{mn} and 21 parameters for external modes are used as adjustable parameters. And the sum of conformational energy and the crystallographic residual was minimized. The power of this new method of refinement has been tested theoretically by using simulated diffraction data [4,5] and experimentally by free-R factor [10]. The statistics of the refinement are summarized in Table 1.

4. Dynamic structure of lysozyme

In this section, the normal mode refinement of wild type human lysozyme is summarized. Fig. 2 shows the ORTEP [12] drawing of the total apparent and the internal fluctuations of C_α atoms together with the two-dimensional representation of the average shape of thermal ellipsoids. The internal fluctuations are much smaller than the apparent ones. This means that the apparent fluctuations, in which the internal fluctuations are heavily masked by the external ones, does not necessarily reflect the real protein dynamics. The internal fluctuations show very large anisotropy, while the apparent anisotropy, being masked by the external terms, is much closer to 1.0, that of the isotropic distributions. Directions of internal fluctuations are found to have no correlation with those of the external terms, i.e. the average cosine of the angle between the two directions, $\langle \cos\theta \rangle$, is found to be 0.45, very close to the random level of 0.5. The directions of the thermal ellipsoids of the internal fluctuations indicate the existence of the hinge-bending motion. The principal axes of the residues 40–50 and 70–80 in one lobe and those of the residues 105–115 in the other lobe appear to be aiming at the cleft region.

In addition, the correlation of inter-atomic fluctuations in protein can also be deduced from the normal mode refinement. Fig. 3 shows the covariance, $C_{ij}(=\langle \Delta \boldsymbol{r}_i \cdot \Delta \boldsymbol{r}_j \rangle)$, for a pair of C_α atoms. This figure shows how fluctuations of two residues are correlated. Positive and negative correlations are for a pair of residues moving simultaneously in the same and in the opposite directions, respectively. The hinge-bending motion in lysozyme should appear as a negative correlation between pairs of residues located in the different lobes forming the cleft. The loci giving negative correlations at β, δ and γ in Fig. 3 suggest the existence of the hinge-bending motion. However, the fluctuation does not seem to be a simple rigid body motion between the two lobes. In the first lobe (the right side of the main chain trace), negatively correlated fluctuations are found in ϵ and ζ. In the second lobe (the left side), there is negative correlation between two loops (α), each of which exhibits a rigid body motion appearing as positive correlations within each loop (A and B).

By the normal mode refinement, the anisotropic and correlated picture of the internal dy-

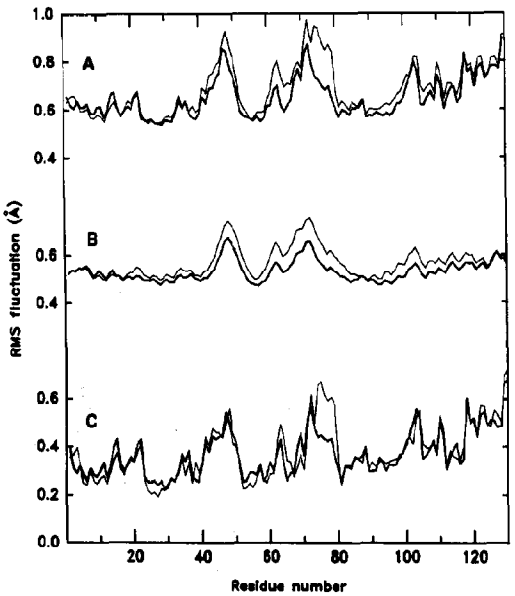

Fig. 4. Comparisons of root mean square fluctuations $\langle \Delta r^2 \rangle^{1/2}$ of main chain atoms between the wild-type human lysozyme (thin curves) and the C77A/C95A mutant (thick curves). The values of $\langle \Delta r^2 \rangle^{1/2}$ are determined by normal mode refinement and averaged within each residue. (A) The total apparent fluctuations. (B) The external fluctuations. (C) The internal fluctuations. These values are related by $\langle \Delta r^2 \rangle_{\text{total}} = \langle \Delta r^2 \rangle_{\text{external}} + \langle \Delta r^2 \rangle_{\text{internal}}$.

namics is determined from diffraction data after removing the influences from the external terms.

5. Mutation effects on the dynamic structure

In this section, we describe the normal mode refinement of a human lysozyme mutant, C77A/C95A, in which a disulfide bond is removed by mutations of Cys77 and Cys95 to alanine [10]. In comparison to the wild-type, we discuss how the removal of this disulfide bond affects the internal atomic fluctuations.

Fig. 4 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. 4A). However, when the total apparent 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 Fig. 4B. 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 Table 1.

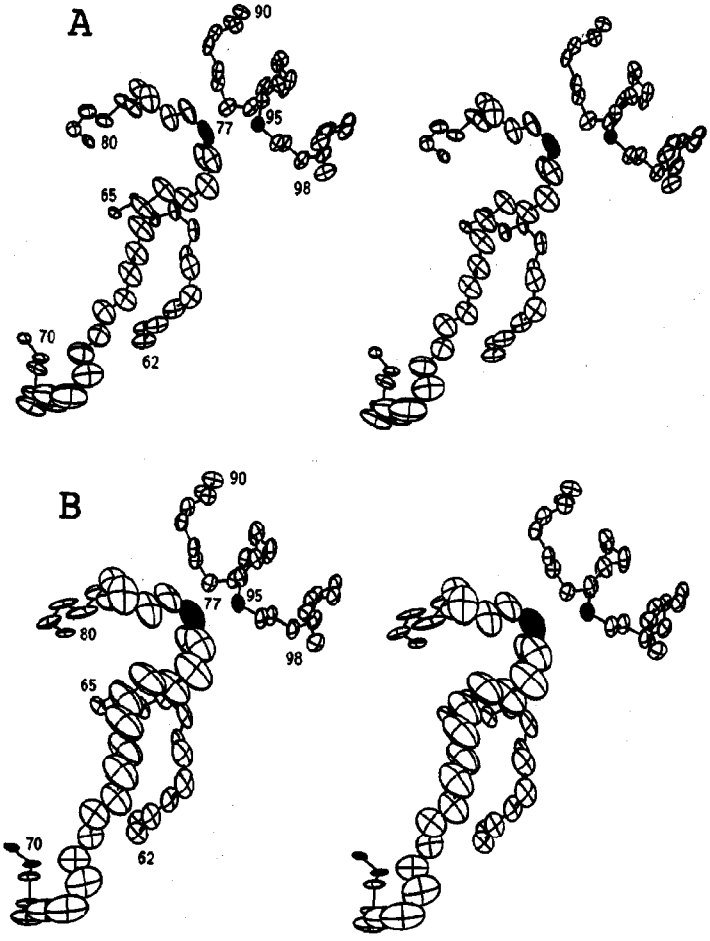

Fig. 5. Stereo ORTEP drawing of thermal ellipsoids (at a 1.8σ level) for the internal fluctuations of N, Ca, and C for the residues 62–65, 70–80, and 90–98. (A) Wild-type. (B) C77A/C95A. $C_\alpha 77$ and $C_\alpha 95$ are indicated by the shaded atoms.

An amino acid substitution usually causes only a small change in the average structure localized at the mutation site [13]. Localization occurs also in the response of the dynamic structure to the mutation. It can be seen in Fig. 4C 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 two 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.

Detailed pictures of the dynamic structures at the mutation sites are given in Fig. 5. They are the ORTEP drawings of the internal fluctuations of the main chain atoms of the 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 ($\langle|\cos\theta|\rangle$) is 0.95 for the residues 71–77. For C77A/C95A, the atoms of the residues 62–65 appear to move in concert with the 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.

The normal mode refinement presented here is expected to enhance the applicability of the protein X-ray crystallography to detect functionally important collective fluctuations and detailed response of dynamic structure to mutations.

Acknowledgement

This work was supported in Kyoto University by grants from MESC and HFSP to NG.

References

[1] G.A. Petsko and D. Ringe, Ann. Rev. Biophys. Bioeng. 13 (1984) 331.
[2] H. Frauenfelder, G.A. Petsko and D. Tsernoglou, Nature 280 (1979) 558.
[3] C. Edwards, S.B. Palmer, P. Emsley, J.R. Helliwell, I.D. Glover, G.W. Harris and D.S. Moss, Acta Cryst. A 46 (1990) 315.
[4] A. Kidera and N. Gō, Proc. Nat. Acad. Sci. USA 87 (1990) 3718.
[5] A. Kidera and N. Gō, J. Mol. Biol. 225 (1992) 457.
[6] A. Kidera, K. Inaka, M. Matsushima and N. Gō, J. Mol. Biol. 225 (1992) 477.
[7] N. Gō, T. Noguti and T. Nishikawa, Proc. Nat. Acad. Sci. USA 80 (1983) 3696.
[8] V. Schomaker and K.N. Trueblood, Acta Cryst. B 24 (1968) 63.
[9] W.A. Hendrickson, Meth. Enzymol. 115 (1985) 252.
[10] A. Kidera, K. Inaka, M. Matsushima and N. Gō, Protein Sci. 3 (1994) 92.
[11] S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona, S. Profeta Jr. and P. Weiner, J. Am. Chem. Soc. 106 (1984) 765.
[12] C.K. Johnson, Report ORNL-5138, Oak Ridge National Laboratory, Tennessee, USA (1976).
[13] B.W. Matthews, Biochemistry 26 (1987) 6885.