Analysis of unfolded state of protein by the small angle X-ray scattering and molecular modeling

Yasutaka SEKI$^{1,3}$, Kaoru ICHIMURA$^2$, Kunitugu SODA$^3$, Yuzuru HIRAGI$^{*4}$

$^1$Nagoya Univ., Nagoya 464-8602, Japan, $^2$Dokkyo Univ. Sch. Med., Mibu, Tochigi 321-0293, Japan
$^3$Nagaoka Univ. Tech., Nagaoka, Niigata 940-2188, Japan, $^4$Kyoto Univ., Uji, Kyoto 611-0011, Japan

Introduction

The small angle X-ray scattering (SAXS) method has unique advantage over that of spectroscopic and NMR. It offers information on the global structure of protein in solution. Unfolded state of protein is very important for the study of protein stability and folding mechanism. By SAXS method, unfolded state structure is clearly distinguished from the native one with a mean square radius $R_g$ from Guinier analysis and molecular shape (globule or chain like etc) from Kratky plot$^1$. Present study aims to extract more detailed structural information on unfolded state structure by the combination of SAXS and molecular modelling method.

Experiment of SAXS measurement

SAXS measurements of two unfolded states of horse apomyoglobin, acid unfolded (AU) state in 20mM HCl solution and urea unfolded (UU) state in 5M urea, were carried out at BL-10C for 30 minutes using a flow cell (Hiragi, unpublished). Aggregation owing to the radiation damage by X-rays was not observed. Intermolecular spatial correlation$^{2,3}$ by Coulomb repulsion had been observed in AU state. This was corrected by the procedure of Zimm with the scattering curves of six different protein concentrations,

$$\frac{B \cdot c}{I_{exp}(K, c)} = \frac{1}{M \cdot I_n(K)} + 2A_cQ_{Exp}(K) \cdot c$$

where $c$ is the protein concentration, $B$ the instrumental factor, $I_{exp}(K,c)$ the scattered intensities at $c$, $M$ the molecular mass of a protein and $I_n(K)$ the scattering by a single particle, respectively. Second term in the right equation indicates that the inverse SAXS intensity per unit protein is a linear function of $c$ and $2A_cQ_{Exp}(K)$ which depends on the scattering vector $K$. Least squares fitting was made for the equation at each $K$ and $I_n(K)$ was obtained from the slope. On the other hand, the UU state solutions in different protein concentrations (20, 40 mg/ml) showed no inter particular interactions. $R_{sq}$ values of AU state corrected for the interparticular interaction and UU state measured at 40mg/ml were obtained from the Guinier plot.

Modelling of Random Chain Conformations

Random chain generation algorism (RCGA) was developed for the structural analysis of unfolded state of proteins$^2$. Detail of the method is as follows: 1) all the bond lengths and angles were fixed to the most suitable value; 2) dihedral angles of main chain group ($\phi, \psi$) were randomly chosen from the ensemble; 3) atomic collision is avoided. Conformations by RCGA are random chain with amino acid residue as one unit but have various degree and frequency in bending of peptide chain.

Three hundred and seventy nine native conformations not having a sequence homology and structural similarity were chosen from the native protein structures registered in PDB. The ($\phi, \psi$) ensemble is obtained from these conformations. They are classified in the following groups: all, all; $\alpha$, $\alpha$-helix region; $t$, hydrogen bonded turn region; $\beta$, $\beta$-strand region; $c$, coil region (having no secondary structure). We get eight ensembles of the ($\phi, \psi$) using these groups: all, $\alpha$, $\alpha$+c, $\alpha$+$\beta$+c, $t$+$\beta$+c, $\alpha$+t+c, $t$+c, $\beta$+c. Thousand conformations for each of above eight ensembles were generated and $R_{sq}$ was averaged over all conformations. Standard error of averaged $R_{sq}$ is 0.02 to 0.04 nm and well exact to compare with experimental $R_{sq}$ values.

Result and Discussion

Experimental $R_{sq}$ value is 4.54 ($\pm$0.11) nm for AU state and 5.12 ($\pm$0.30) nm for UU state. $R_{sq}$ values calculated from molecular modelling, in nm, are 3.36 for all, 3.23 for $\alpha$, 3.46 for $\alpha$+$\beta$+c, 3.98 for $t$+$\beta$+c, 3.20 for $\alpha$+c, 3.63 for $t$+c, 5.18 for $\beta$+c, 4.58 for $c$, respectively. Calculated $R_{sq}$ values vary from 3.20nm to 5.18nm. Hence, molecular modelling method samples a large conformational space, $R_{sq}$ of AU state agree with that of c model within the experimental error. However, SAXS profiles of AU state and c model differ each other on Kratky plot (data not shown). Therefore, AU state of apomyoglobin is not random chain but perturbed chain having some long-range interaction among residues. This result would give further information on unfolded structure.

Experimental $R_{sq}$ value of UU state agrees well with that of $\beta$+c model; both of SAXS profiles are also similar (data not shown). This indicates that main chain group is exposed to solvent owing to the contact of urea molecules with amide and carbonyl groups of the main chain.

references


*hiragi@mbi1.kuicr.kyoto-u.ac.jp

$^{*}$