Global Structure Analysis of Acid-Unfolded Myoglobin with Consideration to Effects of Intermolecular Coulomb Repulsion on Solution X-ray Scattering†

Yasutaka Seki,‡ Tadashi Tomizawa,§ Yuzuru Hiragi,|| and Kunitsugu Soda*,‡,§

Department of Bioengineering, Nagaoka University of Technology, Nagaoka, Niigata 940-2188, Japan, Protein Research Group, RIKEN Genomic Sciences Center, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan, and Department of Physics, Kansai Medical University, Moriguchi, Osaka 570-8506, Japan

Received August 4, 2006; Revised Manuscript Received November 5, 2006

ABSTRACT: To obtain information on the global structure of protein in the acid-unfolded (AU) state, the structure of apomyoglobin (apoMb) was analyzed by using the solution X-ray scattering (SXS) method. SXS profiles were obtained over a wide range of protein concentrations, 1−18 mg mL^{-1}, under strongly acidic conditions. From analysis of the SXS profile extrapolated to a zero protein concentration, the mean square radius, R_{sq}, of AU-apoMb at 20 mM HCl was estimated to be 4.81 ± 0.31 nm. This estimate is more than 1.3 nm larger than those of 3.0−3.5 nm reported thus far. The difference originates from the fact that effects of Coulomb repulsive forces acting between AU-apoMb molecules have not been correctly taken into account in the conventional analysis. In fact, even at a low protein concentration of 1 mg mL^{-1} close to the limit of measurement in the present SXS method, the solution condition applicable to estimating accurately structural parameters of AU-apoMb is very limited. At HCl concentrations lower than 10 mM, the scattering intensity at a small scattering vector decreases remarkably through the effect of intermolecular repulsive forces and the forward scattering intensity is significantly lower than the estimate from the partial specific volume of protein. On the other hand, at HCl concentrations higher than 50 mM, some compact molten-globule-like structures emerge. As a result, the intermediate concentration of 20 mM HCl is the best choice of the solution condition for determining R_{sq} of AU-apoMb. The effect of intermolecular Coulomb repulsion on the SXS profile of AU-apoMb is at its maximum for forward scattering and decreases monotonously with an increase in the scattering angle to be virtually negligible at $K \approx 0.63$ nm^{-1}. Whereas urea-denatured apoMb shows a SXS profile typical of Gaussian chains, the intrinsic SXS profile of AU-apoMb differs significantly from those of Gaussian chains.

Both the kinetic initial state of protein folding and the equilibrium-denatured state as the reference state to the native one are non-native states. These non-native states are known to have different structural characteristics depending upon how they are generated. Each of them consists of an ensemble of protein molecules having diverse conformations. Among a number of proteins, many researchers have chosen apomyoglobin (apoMb)[1] as a target molecule for analyzing physical mechanisms of the structural stabilization and folding of protein (*1*−*4*). Many studies have been carried out on the acid-unfolded (AU) state of apoMb (*5*−*7*). The AU state of a protein is one of the denatured states where its three-dimensional structure is broken by intramolecular Coulomb repulsive forces between positively charged basic groups under strongly acidic conditions. The global structure of AU-apoMb has been considered well-represented by an unfolded "random coil" chain (*6*, *7*). From analysis by circular dichroism (CD) and nuclear magnetic resonance (NMR) spectra, however, it has been indicated that an AU-apoMb molecule has local structures containing α-helical fragments (*5*). Therefore, it is very important to study details of both global and local structures of AU-apoMb for elucidating structural characteristics of the AU state as one of the non-native states.

Solution X-ray scattering (SXS) (*8*, *9*) is a very useful method for analyzing the globular structure of proteins in solution. From the measurement of a SXS profile, i.e., dependence of the scattered X-ray intensity on the scattering angle, various information is obtained on the structure of the solute protein, such as the mean square radius, R_{sq}, the molecular shape as globular or chain-like, and the distance distribution function. These types of information on the global structure of proteins are complementary to those on their local structure from spectroscopic measurements such as NMR, CD, and fluorescence spectra. They give us frequently an important clue for elucidating the non-native structure of proteins. When the SXS method is applied to an ensemble of solute molecules having different conformations, it gives their structural information averaged over the

† This work was supported by a Grant-in-Aid for Scientific Research on Priority Areas of "Genome Information Science" (number 12208006) and "Water and Biomolecules" (number 16041214) of the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

* To whom correspondence should be addressed. Telephone and Fax: +81-258-47-9424. E-mail: soda@vos.nagaokaut.ac.jp.

‡ Nagaoka University of Technology.

§ RIKEN Genomic Sciences Center.

|| Kansai Medical University.

[1] Abbreviations: AU-apoMb, acid-unfolded apomyoglobin; N-holoMb, native holomyoglobin; UU, urea unfolded; MG, molten globule; SXS, solution X-ray scattering; CD, circular dichroism; NMR, nuclear magnetic resonance.

10.1021/bi061578+ CCC: $37.00
Published on Web 01/02/2007

ensemble. With this advantage, various non-native states of proteins such as the denaturant-, acid-, heat-denatured states, and the denaturation intermediate state have been analyzed by SXS (*10*). In recent years, high-brightness X-ray sources of synchrotron orbit radiation (SOR) and 2D X-ray detectors have become available, which has markedly improved the accuracy of measurement and enabled us to analyze precisely SXS profiles in the intermediate to large K regions (*11*, *12*). As a result, it has been becoming possible to obtain not only information on the global structure but also detailed information on the local structure of proteins, the latter of which has thus far been difficult to obtain (*13*, *14*).

For a dilute solution where the solute concentration is so low that the influence of intermolecular attractive and/or repulsive interactions can be neglected, we can safely assume that solute molecules distribute in solution mutually independently. Under such a condition, the SXS profile obtained from the experiment reflects the structure of a single solute molecule accurately. On the other hand, when the influence cannot be neglected, some effect of the spatial correlation between solute molecules appears on the measured SXS profile. This effect appears strongly on the shape of a SXS profile in the small K range. The SXS intensity decreases with increasing repulsive forces and increases with increasing attractive forces. The long-range interaction that acts between protein molecules and affects most strongly their spatial distribution is the Coulomb interaction between charged groups. The magnitude of the effect varies depending upon the concentration, size, and number of charged groups of the protein molecule and the pH and salt concentration of the solution. In the SXS measurement, to ensure the scattering intensity needed for precise profile analysis, the protein concentration, c, must usually be higher than about 1 mg mL^{-1}, which is more than 1 order of magnitude higher than that in spectroscopic measurements such as CD and fluorescence spectra. On the other hand, to obtain a SXS profile reflecting only the structural characteristics of a single solute molecule, a measurement condition must be chosen so that the effect of spatial correlation between solute molecules can be neglected. The non-native structure of a protein usually changes greatly with solution conditions such as the pH and salt concentration. The AU state taken in this study is one of the denatured states of proteins and comes out under the condition of a low pH and low salt concentration (*15*). Effects of the electrostatic repulsion between protein molecules can be very large under these solution conditions. In fact, it is observed in SXS measurements of AU-apoMb that the forward scattering intensity decreases greatly with an increase in the protein concentration (*6*). In such a case, to extract the structural information of a single protein molecule from analysis of experimental SXS profiles, it is necessary to measure SXS profiles for different protein concentrations and obtain a SXS profile extrapolated to $c = 0$.

The purpose of this study is to obtain a SXS profile of AU-apoMb extrapolated to $c = 0$ and to obtain detailed information on its average global structure. To achieve this aim, we first performed SXS measurements varying both HCl and NaCl concentrations in solution to determine the solvent condition that enables us to obtain information on the global structure of AU-apoMb with high accuracy. Then, carrying out SXS measurements of solutions with varying protein concentrations, we obtained a SXS profile extrapolated to $c = 0$. Independently of the above, we measured the partial specific volumes of AU-apoMb and native holomyoglobin (N-holoMb) to test consistency between these data and the estimates from SXS forward scattering intensities for the two samples. Finally, on the basis of the accurate SXS profile thus obtained at the limit of $c = 0$, we discuss details of the global structure of the AU-apoMb molecule.

MATERIALS AND METHODS

Preparation of Sample Solutions. Horse heart N-holoMb was purchased from the Nacalai Tesque Company in Japan. Apomyoglobin samples were prepared from N-holoMb through the following procedure: Heme was extracted and removed from N-holoMb by mixing an equal volume of 2-butanone and an aqueous holoMb solution with a concentration of 5% (w/v) at pH 2.21. This procedure was repeated 4 times (*6*, *7*). The resultant solution of apoMb was dialyzed against pure water and 10 mM ammonium carbonate buffer at pH 7.0. After purification by gel filtration with Sephadex G-50, the sample solution was desalted by dialysis against water and finally lyophilized.

Solvents for SXS measurement are as follows: The solvent for N-holoMb is 10 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES) buffer at pH 6.0. An aqueous solution with 20 mM HCl was chosen for the solvent for most of the AU-apoMb samples. Those with 5, 10, and 50 mM HCl were used for studying the effect of the counterion concentration on SXS profiles. The solvent for urea-unfolded (UU) apoMb is 10 mM HEPES buffer at pH 6.0 containing 5 M urea.

Protein solutions for the SXS measurement were prepared by the following procedure: N-holoMb samples were purified by gel filtration with Sephadex G-50. Powdery apoMb samples were dissolved in the solvent for measurement and equilibrated with an excess amount of the same solvent to set the pH of solutions equal to that of the solvent. Prior to the SXS experiment, each of the N-holoMb and AU- and UU-apoMb solutions was filtered through a membrane filter with a pore size of 0.45 μm to remove aggregates and dirt. The concentrations of N-holoMb and AU-apoMb molecules in solution were determined from optical absorption measurements using the molar extinction coefficient at $\lambda = 408$ nm, $\epsilon = 1.60 \times 10^5$ (*16*), and that at $\lambda = 280$ nm, $\epsilon = 1.43 \times 10^5$ (*6*), respectively. The weight density of each solution was determined with the density meter, Anton Paar model DMA5000.

SXS Measurement. SXS experiments have been carried out using the beam lines at two SOR facilities, i.e., BL10C of Photon Factory at KEK and BL40B2 of Spring-8 at JASRI in Japan. To minimize the effect of damage from X-ray radiation to protein samples, SXS profiles were measured with a flow cell through which all of the sample solutions were allowed to flow with a rate of 3 μL s^{-1}. At BL10C, a 1D-detector position sensitive proportional counter (PSPC) was used for measuring the scattering intensity. The measurement time was chosen to be 30−90 min depending upon the solvent and protein concentration. The X-ray wavelength is 0.1488 nm, and the sample−detector distance is about 0.8 m. The scattering angle was calibrated with the position of layers from collagen fiber. At BL40B2, a sensitive 2D

detector of the imaging plate was used and the measurement time was chosen to be 10 min. Use of the 2D detector gave us a more sensitive measurement in a shorter time than at BL10C. The X-ray wavelength is 0.1500 nm, and the sample−detector distance is 1.00 m. The minimum value of the measurable K range (K_{min}) is 0.31 nm^{-1} at both beam lines, and the data in the range of $0.31 < K < 2.50$ nm^{-1} were taken for analysis. In this K range, both beam lines yielded us SXS profiles with sufficient accuracy and reproducibility needed for the later precise analysis. Both profiles for the solution and solvent were measured, and the difference between them divided by the protein concentration, c, was assumed to be the raw SXS profile of a single solute molecule.

The intensity of scattered X-ray in the SXS measurement depends upon the measuring conditions such as the intensity of incident X-ray, the sample−detector distance, the thickness of the sample cell, as well as the characteristic parameters of the sample solution. To enable a quantitative comparison of data from the two measuring systems of BL40B2 and BL10C, the intensities of X-ray scattered from the same solution of N-holoMb with $c = 5$ mg mL^{-1} were measured at both systems. From the X-ray transmittance measurement at BL10C, it was confirmed that the difference in X-ray absorption could safely be neglected between the 10 mM HEPES buffer for holoMb and the aqueous solution of 20 mM HCl for apoMb.

Analysis of SXS Profiles. We will consider a sample solution of a species of protein molecule dissolved in an aqueous solvent. When the protein concentration, c, is low enough for the effect of the spatial correlation between protein molecules to be negligible, the SXS profile for the protein, which will be denoted by $I(K, c)$, is given by the following equation (*8*):

$$I(K, c) = \tilde{K}M(\Delta z)^2 I_n(K) \tag{1}$$

It is well-known that, when the effect of intermolecular interactions is significant, $I(K, c)$ is represented by the formula of a power series of c developed by Zimm (*17*) and Flory and Bueche (*18*)

$$\frac{\tilde{K}(\Delta z)^2}{I(K, c)} = \frac{1}{MI_n(K)} + 2A_2 Q_{FB}(K)c + ... \tag{2}$$

Equation 2 is reduced to eq 1 when c approaches 0. The factor $\tilde{K}$ is a constant defined by (*8*)

$$\tilde{K} = \frac{I_e N_A d I_i}{a^2} \tag{3}$$

where I_e is the Thomson factor, N_A the Avogadro number, a is the sample−detector distance, d is the thickness of the sample, and I_i is the intensity of incident X-ray. In eqs 1 and 2, M is the molecular mass of the protein (in g mol^{-1}) and c is the weight concentration of the protein (in g mL^{-1}). The quantity Δz is the excess number of electrons per unit weight of the protein defined by

$$\Delta z \equiv z - \rho v_p \tag{4}$$

where z (in mol g^{-1}) is the number of electrons per unit weight of the protein, ρ (in mol mL^{-1}) is the mean number density of electrons in the solvent, and v_p (in mL g^{-1}) is the partial specific volume of the protein. The function $I_n(K)$ is the normalized scattering factor of the protein, where $I_n(0) = 1$. In eq 2, the parameter A_2 is the second virial coefficient of osmotic pressure and the function $Q_{FB}(K)$ is a function that represents the K dependence of the first-order term and approaches 1.0 at the limit of $K = 0$ (*18*). The expression for $I(K, c)$ is obtained from eq 2 as

$$I(K, c) = \frac{\tilde{K}M(\Delta z)^2 I_n(K)}{1 + 2MA_2 Q_{FB}(K) I_n(K) c} \tag{5}$$

and the forward scattering intensity, $I(0, c)$, i.e., the scattered X-ray intensity at $K = 0$, is given by

$$I(0, c) = \lim_{K \to 0} I(K, c) = \frac{\tilde{K}M(\Delta z)^2}{1 + 2MA_2 c} \underset{c \to 0}{\to} \tilde{K}M(\Delta z)^2 \tag{6}$$

We can see from eqs 5 and 6 that the scattering profile $I(K, c)$ is approximated at sufficiently low c as

$$I(K, c) \approx I(0, c) I_n(K) \tag{7}$$

When the intrinsic forward scattering intensity of the protein, $I_0(0)$, at the limit of $c \to 0$ is defined as

$$I_0(0) = \lim_{c \to 0} I(0, c) = \tilde{K}M(\Delta z)^2 \tag{8}$$

and the relative forward scattering intensity r_f is defined by the ratio of $I(0, c)$ for a sample molecule S and $I_0(0)$ for a reference molecule R as

$$r_f \equiv \frac{I(0, c; S)}{I_0(0; R)} \tag{9}$$

we obtain the following expression for r_f at the limit of $c \to 0$ as

$$r_f = \frac{M_S(\Delta z_S)^2}{M_R(\Delta z_R)^2} \quad \text{with } c \to 0 \tag{10}$$

Substituting the known values of M_R, M_S, Δz_R, and the measured value of r_f at $c \to 0$ into eq 10, we can estimate $\Delta z_S = z_S - \rho v_{p,S}$ and then the partial specific volume of the sample molecule, $v_{p,S}$. If we apply a value of r_f measured for a finite value of c to eq 10, we can estimate an apparent value of $v_{p,S}$ for the protein concentration, c. Adopting N-holoMb as the reference molecule, we applied this method to the sample molecule of AU-apoMb to test the validity of the assumptions in the analysis of SXS profiles. The value of z is 0.5345 for holoMb and 0.5536 for apoMb. The value of ρ was estimated to be 0.5537 for 10 mM HEPES buffer and 0.5535 for aqueous 20 mM HCl solution from solvent composition and weight density. The reported value of 0.745 ± 0.003 mL g^{-1} (*19*) was taken for the value of v_p of N-holoMb. With these preparations, we first determined the intrinsic forward scattering intensity $I_0(0;$ N-holoMb) for N-holoMb by extrapolating measured SXS profiles to $c = 0$. The forward scattering intensity of each sample was represented by a relative value to $I_0(0;$ N-holoMb), chosen

as the reference intensity. These means enable us to compare quantitatively scattering-intensity data between different samples.

Guinier and Debye Analyses. Two approximate formulas described below were applied to the analysis of measured SXS profiles for deriving the values of $I(0, c)$ and R_{sq}

$$I(K, c) = I(0, c)\exp\left(-\frac{x}{3}\right) \quad (11)$$

$$I(K, c) = I(0, c)\frac{2}{x^2}\{x - 1 + \exp(-x)\} \quad (12)$$

where $x = (KR_{sq})^2$. Equation 11 or the Guinier formula (*20*) holds for any molecule in dilute solution at sufficiently small K where the relation $x \ll 1$ holds. This formula has been widely used for determining the values of $I(0, c)$ and R_{sq} from measured SXS profiles. Equation 12, which is called the Debye formula (*21*), describes the SXS profile of a Gaussian coil that is the model for a random-coil molecule. It gives a better approximation for a coil chain than the Guinier formula and reproduces its profile in a wider range of K. In this study, the Guinier, the Debye, and both formulas were applied to N-holoMb, UU-apoMb, and AU-apoMb, respectively, for determining the values of their $I(0, c)$ and R_{sq}. The range of K for the least-squares fit in the analysis of each profile is described in the Results.

RESULTS

Dependence of the SXS Profiles of AU-apoMb on the HCl Concentration. The acid denaturation of a protein is caused by an increase in its intramolecular Coulomb repulsion owing to the protonation of acidic residues. To examine dependences on the HCl concentration of $I(0, c)$ and R_{sq} for the AU-apoMb molecule, the SXS measurement was carried out on its aqueous solutions with different HCl concentrations of 5, 10, 20, and 50 mM at a protein concentration of 1 mg mL^{-1}. These experiments were performed at BL40B2 with a measuring time of 10 min. Guinier plots for the obtained SXS profiles are shown in Figure 1. The mean square radius, R_{sq}, for AU-apoMb and the value of r_f (eq 9) with R = N-holoMb and S = AU-apoMb, i.e.,

$$r_f = \frac{I(0, c; \text{AU-apoMb})}{I_0(0; \text{N-holoMb})} \quad (13)$$

are listed in Table 1. The value of r_f characterizes the apparent X-ray scattering intensity of a single AU-apoMb molecule in protein solution with a concentration of c relative to the corresponding real value for a single N-holoMb molecule.

The value of $I(0, c;$ N-holoMb) for N-holoMb could be determined accurately with Guinier analysis. For AU-apoMb, both Guinier and Debye analyses yielded values of $I(0, c)$ that agree with each other within experimental error at all of the four HCl concentrations. On the other hand, Debye analysis gave a larger estimate than Guinier analysis for R_{sq} at a HCl concentration of 20 mM. This will result from the situation that, although Guinier analysis is only applicable to the Guinier region or a narrower range of K than Debye analysis, both analyses were applied to the same range of K and the Guinier analysis underestimated the value of R_{sq}.

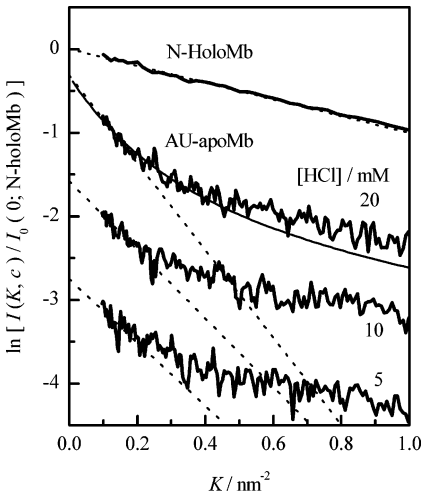

FIGURE 1: Guinier plots of the SXS intensity of the protein divided by the intrinsic forward scattering intensity of N-holoMb $I(K, c)/I_0(0;$ N-holoMb). The functions $I(K, c)$ are the SXS profile for N-holoMb extrapolated to infinite dilution, and those for AU-apoMb with $c = 1$ mg mL^{-1} and different HCl concentrations. Four thick solid lines are the experimental profiles for N-holoMb in 10 mM HEPES buffer at pH 6.0, and AU-apoMb at 20, 10, and 5 mM HCl from top to bottom, respectively. For clarity, the profiles for AU-apoMb at 10 and 5 mM HCl are shifted 0.5 and 1.0 unit downward, respectively. Broken and thin solid lines are the theoretical profiles obtained from least-squares fits of the Guinier (eq 11) and Debye (eq 12) formulas to the experimental ones, respectively. The K ranges of the least-squares fits are 0.31−0.42 nm^{-1} for Guinier analysis and 0.31−0.61 nm^{-1} for Debye analysis.

Therefore, we will hereafter discuss only the result of the analysis from the Debye formula.

The values of R_{sq} for AU-apoMb in 5 and 10 mM HCl solutions were found nearly equal to the reported value of 3.0−3.5 nm (*6*, *7*). In contrast, the value of R_{sq} at 20 mM HCl is estimated to be 4.32 ± 0.15 nm, which is significantly larger than the above. The value of $I(0, c)$ also varies considerably with the HCl concentration. The forward intensity ratio, r_f, has a very low value of 0.44 at the lowest HCl concentration of 5 mM. It increases with an increase in the HCl concentration to 0.72 at 20 mM HCl. The partial specific volume, υ_p, for AU-apoMb could be estimated from the experimental value of r_f obtained by applying Debye analysis to its SXS profile. Here, we will denote the change in υ_p as a result of the N−AU structural transition by $\Delta\upsilon_p$, that is,

$$\Delta\upsilon_p \equiv \upsilon_p(\text{AU-apoMb}) - \upsilon_p(\text{N-holoMb}) \quad (14)$$

The estimated values of both υ_p and $\Delta\upsilon_p$ are shown in Table 2.

From the measurement of the density of protein solutions, the volume change, $\Delta\upsilon_p$, of sperm whale myoglobin accompanying the N−AU transition was reported to be +0.008 mL g^{-1} (*22*). Markedly different from it, all of the estimates from our SXS measurement, $\Delta\upsilon_p = +0.030$ to +0.072 mL g^{-1}, are more than about 4 times larger than the above value. This will result from the condition that, at all HCl concentrations, the scattering intensity in the small K range decreases owing to electrostatic repulsion between protein molecules, and as a result, apparently small values of r_f are estimated. Even with a relatively higher concentration of Cl$^-$ ion at 20 mM HCl, where a decrease in Coulomb repulsion as a result

Table 1: Relative Forward Scattering Intensity, r_f, and Mean Square Radius, R_{sq}, of AU-apoMb

[HCl] (mM)	protein concentration, c (mg mL^{-1})	Guinier analysis		Debye analysis	
		r_f[a]	R_{sq} (nm)	r_f[a]	R_{sq} (nm)
5	1	0.47 (±0.05)	3.37 (±0.32)	0.44 (±0.02)	3.28 (±0.15)
10	1	0.56 (±0.05)	3.51 (±0.28)	0.52 (±0.02)	3.53 (±0.15)
20	1	0.73 (±0.05)	3.96 (±0.18)	0.72 (±0.03)	4.32 (±0.15)
50	1	0.71 (±0.05)	3.39 (±0.21)	0.76 (±0.03)	3.95 (±0.13)
20	0[b]	0.79 (±0.11)	3.92 (±0.36)	0.86 (±0.08)	4.70 (±0.31)
20	5	0.54 (±0.00)	3.28 (±0.25)	0.43 (±0.02)	3.29 (±0.15)

[a] $r_f \equiv I(0, c; \text{AU-apoMb})/I_0(0; \text{N-holoMb})$, relative forward scattering intensity (eq 9). The value of $I_0(0; \text{N-holoMb})$ was determined by Guinier analysis of the SXS profile for N-holoMb extrapolated to $c = 0$. The K ranges of least-squares fits are 0.31−0.42 nm^{-1} for Guinier analysis and 0.31−0.61 nm^{-1} for Debye analysis. [b] Derived from the SXS profile for AU-apoMb extrapolated to $c = 0$.

Table 2: Partial Specific Volumes, υ_p, of N-holoMb and AU-apoMb Estimated from the Relative Forward Scattering Intensity

source	[HCl] (mM)	protein concentration, c (mg mL^{-1})	partial specific volume, υ_p (mL g^{-1})		difference, $\Delta\upsilon_p$ (mL g^{-1})
			N-holoMb	AU-apoMb	
sperm whale			0.734[a]	0.742[a]	+0.008[a]
horse heart	5	1	0.745 (±0.003)[b]	0.817 (±0.005)[c]	+0.072[c]
horse heart	10	1		0.804 (±0.005)[c]	+0.059[c]
horse heart	20	1		0.775 (±0.007)[c]	+0.030[c]
horse heart	20	0		0.758 (±0.013)[d]	+0.013[d]

[a] From ref *22*. [b] From ref *19*. [c] From this work. The parameters taken for the estimation of υ_p are described in the Materials and Methods. [d] These values were derived in this work from analysis of the SXS profile extrapolated to $c = 0$.

of electrostatic shielding effects is expected, an increase in $\Delta\upsilon_p$ as much as +0.030 mL g^{-1} was estimated. With still lower counter-ion concentrations at 5 and 10 mM HCl, the effect of intermolecular repulsion will be very strong even at a low protein concentration of $c = 1$ mg mL^{-1}, which is near the lowest limit of c in our SXS measurement. Thus, it is expected to be difficult to obtain the mean SXS profile of a single protein molecule from extrapolation to $c = 0$ under the two solvent conditions.

Contrastingly to the above, at 50 mM HCl, a slightly larger value of $I(0, c)$ and a smaller value of R_{sq} than the respective values at 20 mM HCl were observed. A small but distinct peak could also be seen in the Kratky profile or the $K^2I(K)$ versus K plot. This result shows that the solution contains a small amount of solute molecules having globular shape. Besides, when 500 mM NaCl is added to the solvent with 20 mM HCl, the value of $I(0, c)$ is nearly unchanged from that for the initial protein solution not containing NaCl; however, the value of R_{sq} decreases considerably, and a peak even larger than the above was observed in its Kratky profile. This will probably show that an acidic molten-globule (MG) structure, which is frequently observed under the solution condition of a low pH and high salt concentration, is formed in solution (*2*). These observations suggest strongly that a significant amount of MG-like structures of apoMb are formed under the solution condition of 50 mM HCl (*15*). Using spectroscopic methods of CD and light absorption spectra, Goto et al. studied the acid-induced U_A to A state folding transition of proteins caused by the increase in the HCl concentration in solution (*2*). Their U_A and A states correspond to our AU and MG states, respectively. We can estimate from their data the fractions of apoMb molecules in the MG state at 20 and 50 mM HCl to be 6 and 24%, respectively. From these results, we can see that (a) a small but finite amount of species with MG conformation are contained in solution even at 20 mM HCl and (b) as much as 4 times the above MG components are formed at 50 mM HCl. We could also estimate from their data that the HCl concentration necessary for the MG content to be less than 1% is as low as 5.9 mM, which is hardly applicable for the reason described before.

From those stated above, we can conclude that, among the four solvent conditions studied, the "20 mM HCl" is the only one that enables us to obtain a SXS profile reflecting the structure of a single AU-apoMb molecule at the lowest protein concentration of 1 mg mL^{-1}. At low HCl concentrations of 5 and 10 mM, strong intermolecular repulsive forces prevent us from obtaining the SXS profile intrinsic to a single AU-apoMb molecule. On the other hand, at 50 mM HCl, the intramolecular repulsion as well as the intermolecular one is weakened by the shielding effect of counter ions. As a result, a significant fraction of apoMb molecules have an MG-like globular structure distinct from that of the AU state.

Dependence of the SXS Profile of AU-apoMb on the Protein Concentration. To examine the dependence of the SXS profiles of N-holoMb and AU-apoMb on the protein concentration, c, we carried out SXS experiments on N-holoMb for five values of $c = 0.8$, 1.8, 2.7, 5.3, and 10.0 mg mL^{-1} and on AU-apoMb for six values of $c = 1.2$, 3.3, 4.9, 7.5, 10.6, and 17.6 mg mL^{-1}. The solvent for AU-apoMb is an aqueous solution of 20 mM HCl. The measurements were performed with BL10C, and the measuring times are 90 min for a N-holoMb sample with $c = 0.8$ mg mL^{-1} and an AU-apoMb sample with $c = 1.2$ mg mL^{-1} and 30 min for the other samples. SXS profiles were analyzed by using eq 2 (*17*, *18*). Effects of the terms higher than the first order in c were neglected. It is because the value of AIC (*23*), which evaluates validity of the least-squares fit, was found not to depend upon whether the second-order term was incorporated or not, and the respective values of $I_n(K)$ and $2A_2Q_{FB}(K)$ obtained from the two analyses agreed with each other. For a similar reason as above, the first-order term was also neglected at $K > 0.63$ nm^{-1}.

Guinier plots of the SXS profiles for AU-apoMb at three protein concentrations of $c = 4.9$, 10.6, and 17.6 mg mL^{-1}

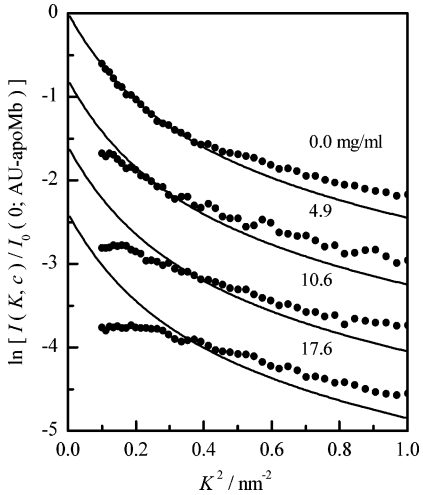

FIGURE 2: Guinier plots of $I(K, c)/I_0(0;$ N-holoMb) for AU-apoMb in solutions with different protein concentrations at a HCl concentration of 20 mM. The number in the figure indicates the protein concentration, c, in units of mg mL^{-1}. Each plot for the solution with nonzero c is shifted 0.4 units downward. The solid line for each plot is the same theoretical profile obtained from a least-squares fit of the Debye formula to the SXS profile for AU-apoMb extrapolated to infinite dilution. The range of K for the least-squares fit is 0.31−0.61 nm^{-1}.

and extrapolated to $c = 0$ in aqueous solution of 20 mM HCl are shown in Figure 2.

For visual convenience, each SXS profile other than that at the limit of $c = 0$ is drawn with its vertical scale being shifted 0.8 unit downward. We can see from Figure 2 that the scattering intensity at $K^2 < 0.4$ nm^{-2} decreases with an increase in the protein concentration. Especially, as the profiles at $c \geq 10$ mg mL^{-1} have upward convex forms, we cannot estimate the value of R_{sq} from Guinier or Debye analysis at this range of c. In contrast, the SXS profiles nearly coincide with each other in the range of $K^2 > 0.4$ nm^{-2}, where the internal structure of the protein is mainly reflected in the profile. These results show that, whereas electrostatic repulsive forces significantly affect the long-range structure of protein solutions, they hardly affect the structure of the protein interior at this range of the protein concentration.

The values of R_{sq} and r_f obtained from Guinier and Debye analyses of the SXS profile extrapolated to $c = 0$ mg mL^{-1} are also listed in Table 1. The Debye analysis gave the values of 4.70 nm and 0.86 for R_{sq} and r_f, respectively, which are significantly larger than the corresponding values of 4.32 nm and 0.72 at $c = 1$ mg mL^{-1}. The value of Δv_p was estimated to be +0.013 mL g^{-1} from the value of r_f. It is consistent, within experimental error, with data of the change in the partial specific volume, Δv_p, accompanying the structural transition from N-holoMb to AU-apoMb for sperm whale myoglobin. This result means that the SXS profile of AU-apoMb extrapolated to its infinite dilution includes practically no contribution from intermolecular repulsion and reflects almost correctly the average structure of a single AU-apoMb molecule. Consequently, the value of 4.70 ± 0.31 nm thus obtained can be regarded as a close approximation to R_{sq} for a single AU-apoMb molecule under the condition of 20 mM HCl. More accurately, we know that (a), according to the data by Goto et al. (*2*), an apoMb solution contains 6% MG component at 20 mM HCl and (b) the value of R_{sq} for the MG state of apoMb is estimated to be 2.31 nm by

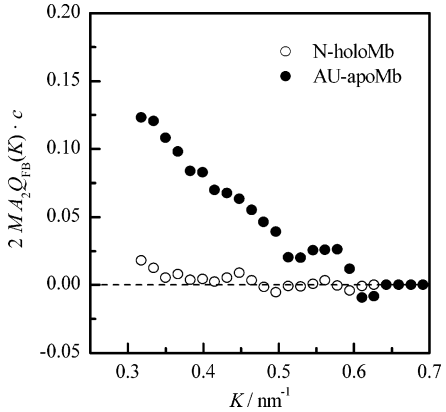

FIGURE 3: First-order term $2MA_2Q_{FB}(K)c$ of the expansion of the apparent inverse scattering factor $I_0(0)/I(K, c)$ (eqs 2 and 8) for N-holoMb and AU-apoMb into a geometric series of protein concentration, c. (○) N-holoMb in 10 mM HEPES buffer at pH 6.0. (●) AU-apoMb in aqueous 20 mM HCl solution.

Kataoka et al. (*6*). When these data are combined with the approximation that the values of r_f for AU-apoMb and MG-apoMb are equal to each other, we finally have an estimate of R_{sq} for AU-apoMb as 4.81 nm. This will be, at present, the most probable estimate of R_{sq} for a single molecule of AU-apoMb at the solution condition of 20 mM HCl.

The effect of intermolecular interactions between protein molecules on their SXS profile is reflected almost in the first-order term, $2A_2Q_{FB}(K)c$, of eq 2. Combining eq 2 with eq 8, we obtain the following expression:

$$2MA_2Q_{FB}(K)c \approx \frac{I_0(0)}{I(K, c)} - \frac{1}{I_n(K)} \qquad (15)$$

The values of $I_0(0)$ and $I_n(K)$ can be derived from the SXS profile obtained by extrapolating measured profiles to $c = 0$. Substituting these data into eq 15, we can estimate the magnitude of the first-order term, $2MA_2Q_{FB}(K)c$, from the difference between the apparent value $I_0(0)/I(K, c)$ and the real one $1/I_n(K)$ for the inverse scattering factor of a single protein molecule. Since it holds that $Q_{FB}(0) = 1$, the magnitude of this effect is equal to $2MA_2c$ at $K = 0$. The dependence on K of the effect is shown in Figure 3 for both solutions of AU-apoMb and N-holoMb with $c = 1$ mg mL^{-1}.

At a finite protein concentration, the scattering intensity, $I(K, c)$, decreases from that at infinite dilution owing to the effect of intermolecular repulsive interactions. We can see from Figure 3 that, at $c = 1$ mg mL^{-1}, (a) the effect is largest at the smallest K and decreases monotonously with increasing K and (b) the effect for AU-apoMb is much larger than that for N-holoMb in the range of K where the former is significantly nonzero. However, it vanishes virtually at $K \approx 0.63$ nm^{-1}, which indicates that intermolecular repulsive forces are not involved in determining the SXS profile at $K > 0.63$ nm^{-1}. This is consistent with the fact shown in Figure 2 that the SXS profile at $K^2 > 0.4$ nm^{-2} does not depend upon the protein concentration. In contrast to the above, at $K < 0.63$ nm^{-1}, the value of Q_{FB} varies notably with K, which means that the apparent values of $I(0, c)$ and R_{sq} estimated from a SXS profile at a finite c higher than 1 mg mL^{-1} deviate significantly from their respective real values.

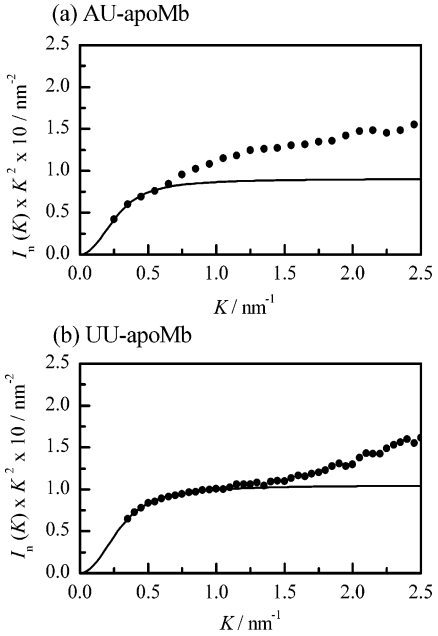

FIGURE 4: Kratky profiles for unfolded apomyoglobin in the two denatured states. (a) AU-apoMb, where the SXS profile is a combined profile of the one extrapolated to infinite dilution at $K \leq 1.5$ nm^{-1} and the other with $c = 5$ mg mL^{-1} at $K > 1.5$ nm^{-1}, with 20 mM HCl. (b) UU-apoMb, where $c = 5$ mg mL^{-1}, with 10 mM HEPES buffer containing 5 M urea at pH 6.0. The solid line in each figure is the theoretical profile of the Debye formula for a Gaussian chain with the same value of R_{sq} as the corresponding unfolded chain.

The ratio of the forward scattering intensity with $c = 1$ mg mL^{-1} to that with $c = 0$ can be estimated from data at 20 mM HCl in Table 2 to be $0.86/0.72 \approx 1.19$. Putting $K = 0$ in eq 15 and using the relations of $I_n(0) = 1$ and $Q_{FB}(0) = 1$, we can estimate the value of $2MA_2c$ with $c = 1$ mg mL^{-1} to be

$$2MA_2c = \frac{I_0(0)}{I(0, c)} - 1 = \frac{0.86}{0.72} - 1 \approx 0.19 \qquad (16)$$

In other words, this result shows that the forward scattering intensity, $I(0, c)$, for AU-apoMb decreases to $(1/(1 + 0.19)) \times 100 \approx 84\%$ of that at infinite dilution. It predicts that, with a higher c of 5 mg mL^{-1}, the $I(0, c)$ decreases more greatly to $(1/(1 + 0.95)) \times 100 \approx 51\%$ of $I_0(0)$. Extrapolating linearly the derived data on $2MA_2Q_{FB}(K)c$ at $K \geq 0.31$ nm^{-1} (Figure 3) to $K = 0$, we have an estimate of $2MA_2c \approx 0.27$, which is significantly larger than the above estimate of 0.19. This erroneous estimate indicates that, approaching $K = 0$, the function $Q_{FB}(K)$ actually tends to be saturated.

Kratky Profile of AU-apoMb. The Kratky profiles for AU-apoMb and UU-apoMb at the limit of $c = 0$ are shown in parts (a) and (b) of Figure 4, respectively, where the profile at $c = 5$ mg mL^{-1} is substituted for the zero c profile for AU-apoMb at $K > 1.5$ nm^{-1}.

For comparison, a theoretical Kratky profile (*21*) is also shown in each figure for the Gaussian chain having the same value of R_{sq} as that of the unfolded myoglobin. The Kratky profile for UU-apoMb coincides with the theoretical profile for the corresponding Gaussian chain in the range of $K < 1.1$ nm^{-1}. In contrast, the range of K where the theoretical profile reproduces well the experimental one for AU-apoMb is estimated to be $K < 0.61$ nm^{-1}, which is much narrower than the above. Showing no plateau region characteristic of the Kratky profile of Gaussian chains, the profile for AU-apoMb rises monotonously with K. This suggests that there is some significant difference between the average conformations of AU-apoMb and UU-apoMb. This difference will be discussed in later sections.

DISCUSSION

Applicable K Range of Guinier and Debye Analyses. In the usual analysis of the SXS profile of a protein solution, Guinier analysis with eq 11 is applied to the experimental profile at a small K region for deriving the two important parameters of $I(0, c)$ and R_{sq}. Guinier analysis has been usefully applied to native proteins. It is because the Guinier region, i.e., the range of K where Guinier analysis is applicable, is relatively wide for the particles with globular structure. Specifically, the range of K required for deriving the value of R_{sq} within an error of 2.5% is given by $KR_{sq} < 1.57$ for a spherical particle having uniform electron density. On the other hand, the Guinier region for a coil molecule is much narrower than that for a globular one. The corresponding K range for a Gaussian chain is given by $KR_{sq} < 0.83$, which is nearly half the range for a spherical particle. For example, for a chain molecule with $R_{sq} = 4.81$ nm, which is equal to that of AU-apoMb, the above condition is represented as $K_{max} < 0.17$ nm^{-1}. This value of K_{max} is smaller than $K_{min} = 0.31$ nm^{-1}, the minimum of the range of K where a SXS profile can be measured in this study. Accordingly, with this measuring range of scattering angles, it is practically impossible to obtain an accurate estimate of R_{sq} from Guinier analysis for a chain molecule having R_{sq} larger than $0.83/0.32 = 2.6$ nm. It has been reported, however, that Debye analysis (*21*) with eq 12 is applicable in the wider K range of $KR_{sq} < 3.0$ for a random-coil molecule such as a protein in the unfolded state (*24*, *25*). As shown in Table 1, the value of R_{sq} for AU-apoMb at 20 mM HCl was estimated from Guinier analysis to be significantly smaller than that from Debye analysis. We can conclude from the above that the discrepancy results from the fact that Guinier analysis was applied in the range of K exceeding the Guinier region.

The estimate of R_{sq} from Guinier analysis for AU-apoMb at 20 mM HCl and the limit of $c = 0$ was found to be 83% of that from Debye analysis. This is nearly consistent with the theoretical prediction that the application of the Guinier analysis (eq 11) to the theoretical SXS profile of a Gaussian chain with the same fit range as the above, i.e., $1.5 < KR_{sq} < 2.0$ nm^{-1}, yields an estimate equal to 78% of the real value of R_{sq}. This result confirms that our Debye analysis is performed on the basis of rational grounds. On the other hand, at a low HCl concentration of 5 mM, the value of R_{sq} (3.37 nm) obtained from Guinier analysis was found to be nearly equal to that from Debye analysis (3.28 nm). In this case, the fit ranges of K were taken to be $1.0 < KR_{sq} < 1.4$ nm^{-1} for Guinier analysis and $1.0 < KR_{sq} < 2.0$ for Debye analysis. The result seems to be paradoxical because, although the range of K for Guinier analysis definitely exceeds the Guinier region for the chain molecule, Guinier analysis yields nearly the same estimate as Debye analysis. At this low HCl concentration, the electrostatic shielding by chloride ions is weak and the spatial correlation between protein molecules owing to intermolecular repulsive forces

strongly affects the SXS profile. Even at a low protein concentration of 1 mg mL^{-1}, the scattering intensity in the small *K* region decreases remarkably and the SXS profile is markedly transformed from that of a single chain-like protein molecule. The above apparent paradox will be interpreted as a result that, even with the Debye analysis where an appropriate range of *K* was chosen, a value of R_{sq} much smaller than the real value of a single isolated molecule was estimated. It is well-known that Debye analysis is more useful than Guinier analysis for chain molecules because the applicable range of *K* of the former is wider than the latter. However, even when a Guinier plot shows an apparently linear dependence on *K* in the small *K* region, the correct value of R_{sq} cannot be estimated from the initial slope of the observed SXS profile if it includes some systematic deviation.

Forward Scattering Intensity and Partial Specific Volume. The value of $I(0, c)$ obtained from the SXS measurement is an important indicator for judging whether the experimental SXS profile correctly reflects the structure of a single solute molecule or not. The apparent molecular weight of a sample molecule can be estimated from the ratio between the value of $I(0, c)$ for the sample molecule and that for a reference molecule with known molecular weight. Agreement between the estimated and real molecular weights would be definite evidence that effects of the intermolecular attraction and/or repulsion are negligibly small and the SXS profile reflects the average structure of a single solute molecule (*6*, *26*). This verification is based on the assumption that Δz in eq 4 has a constant value not dependent upon protein species, which is empirically known to hold approximately. However, since Δz depends upon the mean electron density of solvent, ρ, and the partial specific volume, υ_p, as seen from eq 4, strictly speaking, it varies with protein species and their structure. For example, in the urea denaturation of protein, the value of ρ increases greatly with the addition of a large amount of urea into solution, which leads to a change in the so-called contrast effect or a decrease in scattering intensity. The difference in this effect can be neglected between N-holoMb and AU-apoMb in our study, because the values of ρ for the respective solvents, 10 mM HEPES buffer and aqueous 10 mM HCl solution, are nearly equal to each other (see the Materials and Methods). On the other hand, the contribution to $I(0, c)$ of the change in υ_p accompanying the denaturation of a protein needs to be taken into account. In the studies reported thus far, different views have been presented on the volume change accompanying protein denaturation. Whereas a result of analysis was presented that the volume change is small and less than 0.5% (*27*), experimental data showing a significant volume change was also given (*28*). Therefore, it must be taken into account that the assumption that both native and denatured states have the same value of υ_p may lead to a large error in the estimate of $I(0, c)$. To avoid this possibility, we verified the reliability of the estimated value of $I(0, c)$ for AU-apoMb by (a) evaluating the value of υ_p from r_f obtained from Debye analysis of the SXS profile, where N-holoMb is used as the reference molecule, and (b) comparing the value with the reported one.

According to Makhatadze et al. (*22*), the change in partial specific volume, $\Delta\upsilon_p$, accompanying the N−AU transition of apoMb is +0.008 mL g^{-1} or $\Delta\upsilon_p/\upsilon_p \approx +1.1\%$. This must yield a large decrease in r_f of 7.3% as described in the following: The number densities of *z* and $\rho\upsilon_p$ (eq 4) for N-holoMb are estimated to be 0.535 and 0.413 mol g^{-1}, respectively, at 25 °C, where the value of υ_p is taken from the reported experimental value of 0.745 mL g^{-1} (*19*). From these data, the difference Δz is found to be 29.5% of $\rho\upsilon_p$ for N-holoMb molecules in the buffer solution. Because the number *z* is virtually unchanged with a structural transition even if the value of υ_p is changed, contribution from the contrast effect to the relative forward scattering intensity, r_f, at the limit of $c \rightarrow 0$ (eq 10) is estimated to be

$$\frac{(\Delta z_{apo})^2}{(\Delta z_{holo})^2} = \frac{(z - \rho\upsilon_{p,apo})^2}{(z - \rho\upsilon_{p,holo})^2} = \frac{(z - \rho 1.011\upsilon_{p,holo})^2}{(z - \rho\upsilon_{p,holo})^2} \approx \frac{(1.295 - 1.011)^2}{(1.295 - 1.000)^2} \approx 0.927 \quad (17)$$

The molecular weight of apoMb is 3.6% lower than that of holoMb and the ratio M_S/M_R is 0.964. Combination of this factor and eq 17 yields from eq 10 an estimate of $0.964 \times 0.927 = 0.894$ for the value of r_f. We can see from this estimation that the partial specific volume, υ_p, is a very important factor in determining the forward scattering intensity, $I(0, c)$. Especially, to derive detailed structural information as to the denatured state of the protein from an experimental SXS profile, it is essential to examine the consistency between the values of $I(0, c)$ obtained from analysis of the SXS profile and υ_p measured directly under the same solvent condition as that in the SXS measurement.

The values of υ_p estimated from those of r_f at different HCl concentrations and a protein concentration of 1 mg mL^{-1} were found to be significantly different from the reported experimental values. As shown in Table 2, for example, a value of 0.804 mL g^{-1} was estimated for υ_p from the value of r_f at 10 mM HCl. However, it deviates considerably from the value reported for sperm whale AU-apoMb as well as the range of values of υ_p, 0.700−0.750 mL g^{-1} (*29*, *30*), for most water-soluble proteins. Although there is still room for discussion on the change in protein volume accompanying denaturation as described before, it seems hardly possible for a denatured protein molecule to have a large value of υ_p such as the above. The large estimate of υ_p as 0.804 mL g^{-1} is considered to be an apparent result of a decrease in the scattering intensity in the small *K* range caused by the effect of intermolecular electrostatic repulsive forces. It is observed from the experiment that the value of r_f increases or the apparent value of υ_p decreases with an increasing HCl concentration, which supports the above inference. This will be due to the fact that, with an increase in the Cl^- concentration, the electrostatic shielding effect increases and the intermolecular repulsion decreases. As a result, among the solvents where apoMb is in the unfolded state, the solvent with a HCl concentration of 20 mM minimizes the effect of repulsive forces between myoglobin molecules. We derived the value of r_f from the SXS profile obtained by extrapolating those for AU-apoMb in this solvent to $c = 0$. The value of $\Delta\upsilon_p$ estimated from this r_f agreed with the reported experimental value within error. This result confirms that our SXS profile at infinite dilution represents fairly well the structure of a single unfolded apoMb molecule. Since $\Delta\upsilon_p$ includes a rather large error of ±0.013 mL g^{-1}, it will be difficult to

discuss in more detail, for example, the effect of differences in the primary sequence of myoglobin between sperm whale and horse.

Effect of Intermolecular Electrostatic Repulsion. The number of ionic charges on apoMb is estimated to be 0 at neutral pH if we assume that all of the histidine residues are deprotonated. On the other hand, in the AU state at pH 1.72 with 20 mM HCl, the number is as large as 33. As a result, the effect of intermolecular repulsion on the SXS profile becomes significant and makes it difficult to extract the structural information of a single AU-apoMb molecule. The magnitude of the effect will vary with the ratio ρ_d defined by

$$\rho_d \equiv r_D/d_n \tag{18}$$

where r_D and d_n are the Debye shielding radius of counter ions and the mean distance between neighboring protein molecules in solution, respectively. The larger ρ_d will lead to the larger effect of intermolecular repulsion. The value of r_D is inversely proportional to the square root of the HCl concentration or $r_D \propto [HCl]^{-1/2}$. Since the mean volume per protein molecule is inversely proportional to the protein concentration, c, the value of d_n is inversely proportional to the cube root of c or $d_n \propto c^{-1/3}$. Consequently, the dependence of ρ_d on the concentrations of protein and HCl will be given as follows:

$$\rho_d \propto c^{1/3}/[HCl]^{1/2} \tag{19}$$

Our experiment shows that the effect of intermolecular repulsion is fairly small under the solution condition of $c = 1$ mg mL^{-1} and 20 mM HCl. We can see from eq 19 that, to attain the same value of ρ_d as that in this solution, SXS profiles needs to be measured with very low protein concentrations of $c = (10/20)^{3/2} = 0.35$ mg mL^{-1} at 10 mM HCl and $c = (5/20)^{3/2} = 0.13$ mg mL^{-1} at 5 mM HCl. However, with these protein concentrations, scattered X-ray intensity is so weak that it is very difficult to obtain SXS profiles with high accuracy. In other words, a relatively high HCl concentration of $20 \times 5^{2/3} = 58$ mM is required at $c = 5$ mg mL^{-1} where sufficiently accurate SXS profiles can be obtained in the wide range of $K < 2.5$ nm^{-1}. The solvent condition, however, cannot be taken for the measurement of the AU-state because the fraction of MG structures increases at this HCl concentration. We conclude from those above that 20 mM HCl is a pin point of the solvent condition that enables us to obtain molecular information of AU-apoMb from the analysis of experimental SXS profiles obtained with the present measuring system.

A small but finite difference was observed between the estimates for the value of υ_p of apoMb from direct measurement and the SXS forward scattering intensity. It means that a weak but finite effect of intermolecular repulsion is also included in the SXS profile for the solution with $c = 1$ mg mL^{-1} and 20 mM HCl. Moreover, it was impossible to obtain a SXS profile with high accuracy at $c = 1$ mg mL^{-1} in the medium to large K range of $K > 1$ nm^{-1}, because the scattering intensity was considerably lower compared with that in the small K range. Extrapolation to $c = 0$ concentration at 20 mM HCl is required to obtain the SXS profile of a single AU-apoMb molecule. It is important to determine in what range of K there appears a significant effect of intermolecular repulsion when the protein concentration is increased. In the range of c lower than 18 mg mL^{-1}, which is the maximum concentration in the extrapolation to $c = 0$ at 20 mM HCl, the effect of intermolecular repulsion on the SXS profile was virtually negligible at $K > 0.63$ nm^{-1}. As shown in Figure 4, when the SXS profile extrapolated to $c = 0$ at $K < 1.5$ nm^{-1} is combined with the profile with $c = 5$ mg mL^{-1} at $1.5 < K < 2.5$ nm^{-1}, we could obtain the accurate SXS profile of a single AU-apoMb that covers the value of K in the small to medium range.

Using a method similar to ours, Kataoka et al. (*6*) obtained an "intrinsic" profile extrapolating their experimentally obtained SXS profiles to $c = 0$. Their solution condition of 10 mM HCl at pH 2.0 is different from our condition of 20 mM at pH 1.72. Their result of analysis agrees qualitatively with ours in that the value of $I(0, c)$ for AU-apoMb shows a strong dependence on the protein concentration. However, the values of r_f and $\Delta\upsilon_p$ obtained from their SXS measurement are 1.05 ± 0.05 and -0.009 mL g^{-1}, respectively, which differ significantly from our results. In addition, both the numerical value of R_{sq} and the Kratky profile extrapolated to $c = 0$ differ remarkably. Our estimate of R_{sq} for AU-apoMb is 4.81 nm at 20 mM HCl, but their estimate is much smaller than our estimate of 3.0 nm despite the severer solution condition of 10 mM HCl. Whereas the height of our Kratky profile at $K = 2.5$ nm^{-1} is about 35% of the peak height of our Kratky profile for N-holoMb at $K = 1.0$ nm^{-1} (data not shown), the corresponding height reaches 70% in the profile of Kataoka et al. One of the most probable causes of the large difference between the SXS profiles will be some difference in the AU-apoMb structure owing to the difference in the solution condition. The pH values of both solutions are sufficiently low, and all of the carboxyl groups on the side chains of acidic residues and the C terminus can be safely assumed to be fully protonated. In the sample solution of Kataoka et al., the intramolecular electrostatic repulsion is stronger and the unfolded polypeptide chains should be expanded more than those in our solution, because the Cl^- concentration is 10 mM or one-half that of our solution. Therefore, the result that their estimate of R_{sq} is smaller than ours cannot be explained by a structural difference of AU-apoMb owing to the difference in the solution condition. According to our measurement, in the solution of 10 mM HCl, the forward scattering intensity, $I(0, c)$, decreases to 51% of that for N-holoMb even at the protein concentration of 1 mg mL^{-1}, lowest among our sample solutions. Kataoka et al. estimated the rate of decrease in $I(0, c)$ to be 70−80%. Practically, at such a low Cl^- concentration as 10 mM HCl, it is very difficult to obtain the correct SXS profile extrapolated precisely to $c = 0$ from the profiles of protein solutions with c higher than 1 mg mL^{-1}. Conversely, it is highly probable that such an extrapolation leads us to an erroneous prediction of the SXS profile for a single solute protein. The following is concluded from the above discussion: The considerably lower estimate of R_{sq} for AU-apoMb by Kataoka et al. is due to the situation that the SXS profile is not extrapolated to $c = 0$ correctly. As a result, the K dependence of the extrapolated profile is not accurate and, especially, the SXS intensity at a large K region is estimated to be much larger than it really is. In other words, their functional form of $Q_{FB}(K)$ should differ

greatly from ours, although they did not give it explicitly. As shown in Figure 3, we determined experimentally the function $Q_{FB}(K)$ that decreases monotonously with an increase in K and becomes virtually 0 near $K = 0.63$ nm^{-1}. If the function $Q_{FB}(K)$ decreases with an increase in K more gently than that we determined, it would predict a smaller value of R_{sq} and a higher level of the SXS profile at a large K region.

Tcherkasskaya et al. (*7*) obtained SXS profiles with very low protein concentrations of 5−20 μM or 0.085−0.34 mg mL^{-1} for AU-apoMb at 10 mM HCl and pH 2. They derived a value of $R_{sq} = 3.5$ nm from the analysis of the profiles. It is unknown, however, whether or not the effect of intermolecular repulsion is included in the value of R_{sq}, because they did not refer to the consistency of the value of $I(0, c)$ estimated from the measured SXS profile with the value of υ_p from its direct measurement. Besides, the K range of $0.98 < KR_{sq} < 1.58$ that they took for the least-squares fit in Guinier analysis is completely away from the Guinier region of $KR_{sq} < 0.83$ for a chain molecule. Accordingly, it will be impossible to obtain the true value of R_{sq} from this analysis, even if no effect of intermolecular repulsion is included in the measured SXS profile.

Cytochrome *c* is known to form an ensemble of AU structures at a low pH and low salt concentration in a similar way to myoglobin. Both cytochrome *c* and myoglobin have the α-type native structure and no disulfide bond. Damaschun et al. (*31*) studied the AU state of apocytochrome *c* in 20 mM sodium phosphate hydrochloride solution with pH 2.3. They obtained an estimate of $R_{sq} = 4.6$ nm for AU-apocytochrome *c*. We can see that, whereas our estimate of $R_{sq} = 4.81$ nm for AU-apoMb is significantly larger than those reported thus far as described before, it is near the estimate for AU-apocytochrome *c* by Damaschun et al. Although the number of amino acid residues of cytochrome *c*, 104, is about $^2/_3$ of that of myoglobin, 153, these values of R_{sq} are near each other. It means that the effect of intramolecular repulsion is stronger in apocytochrome *c* than in apomyoglobin. This difference will be partly caused by the difference in the fraction of basic residues between the two molecules. At strongly acidic pH like this, cytochrome *c* and myoglobin have 25 (24.0%) and 33 (21.6%) positive ionic charges, respectively, and therefore, the positive charge density is 11% higher in cytochrome *c* than in myoglobin.

Damaschun et al. (*31*) determined the value of R_{sq}, extrapolating it to the zero concentration of apocytochrome *c*. They combined data from SXS with those from dynamic light scattering (DLS) at a low protein concentration, where the SXS measurement cannot be effectively applied. In protein solutions like these, the effect of intermolecular repulsion on their SXS profiles is so strong that it may be difficult to obtain the SXS profile and R_{sq} for a single molecule. In such cases, it is essential to examine the validity of the SXS experiment by measuring υ_p of the target protein by densitometry and to supplement SXS data with some other experiment such as DLS. In our study, the intrinsic SXS profile of AU-apoMb was determined in the following way: First, we selected a solution condition of 20 mM HCl, which retains the molecular-structure characteristic of the AU state of the protein and minimizes the effect of intermolecular interactions. Second, we extrapolated the measured profiles to the profile at the limit of $c = 0$. The function $Q_{FB}(K)$ derived from the above profiles shows a physically reasonable dependence on K, and the value of $I(0, c)$ is consistent with the value of υ_p determined by densitometry. Moreover, our resultant SXS profile was confirmed to agree quantitatively with the profile predicted from a molecular-modeling analysis (to be published). Consequently, this SXS profile can be surely assumed to be the one that reflects the real structure of AU-apoMb.

In their work on the denatured states of yeast phosphoglycerate kinase, Damaschun et al. (*32*) pointed out that (a) the concentration dependence of the SXS profiles for unfolded proteins is much stronger than that for folded globular proteins and (b) the methods of extrapolation to $c = 0$, which are successfully applied to the SXS profiles of globular proteins, fail in the case of unfolded proteins and yield erroneous results. Gast et al. (*33*) also showed that the concentration dependence of the hydrodynamic radius of AU-apoMb is 20 times stronger than that of N-apoMb. Applying linear extrapolation, they obtained an estimate of $R_{sq} = 4.67$ nm for AU-apoMb at pH 2.0 in 10 mM HCl at 20 °C. It is a harder solvent condition than the optimum one. Recognizing that the extrapolation is imperfect, they noted very carefully that the obtained value of $R_{sq} = 4.67$ nm is possibly somewhat smaller than the real value for AU-apoMb.

Global Structure of AU-apoMb. The Kratky profile of AU-apoMb (Figure 4a) is obviously different in shape from that of Gaussian chains. The Gaussian chain is a model for ideal random-coil chains whose distance distribution between the neighboring chain units obeys a Gaussian distribution. In a real protein molecule, however, the distance distribution between two residues near each other on the primary sequence deviates from a Gaussian distribution because of the potential of internal rotations intrinsic to the polypeptide chain (*34*). We observe that the Kratky profile for UU-apoMb begins to deviate from the theoretical profile for a Gaussian chain at $K \approx 1.1$ nm^{-1}. It means that the deviation from a Gaussian distance distribution is not negligible in the range of the distance, $r \approx K^{-1} < 0.9$ nm. In contrast, the corresponding deviation for AU-apoMb is significant at the distance of $r \approx K^{-1} < 1/0.61 \approx 1.6$ nm. It begins to appear from a nearly twice longer distance in AU-apoMb than in UU-apoMb. In other words, the AU-apoMb chain has a longer persistence length than the UU-apoMb chain. A UU-apoMb molecule is unfolded by local short-range interactions with urea molecules. On the other hand, an AU-apoMb molecule is unfolded by long-range Coulomb repulsive forces acting between positively charged basic side chains, on account of which the deviation will begin to appear from a longer distance in AU-apoMb. The Kratky plot for AU-apoMb shows a dependence on K increasing monotonously with an increase in K in the range of $K > 0.6$ nm^{-1}. The Kratky profile of a rod-like particle is given by a linear expression of K. These results being combined, it is inferred that the structure of AU-apoMb can be modeled by a globally random-coil chain with a locally extended conformation. This is consistent with the molecular mechanism described above that AU conformations are generated by internal Coulomb repulsion. It is difficult, however, to elucidate the molecular detail of how intramolecular Coulomb interactions influence

both the global and local structures of an unfolded protein molecule only from its measured SXS profile. Analysis combining accurate experimental SXS profiles and a molecular-modeling method can be applied effectively for such purposes (*25*, *35*−*39*), which will be reported elsewhere.

ACKNOWLEDGMENT

Y.S. and K.S. express sincere thanks to Drs. Nobuhiro Go and Mitiko Go for their continuous support for our research work. Y.S. is grateful to JSPS of Japan for financial support. K.S. is very grateful to Drs. Akiyoshi Wada and Shigeyuki Yokoyama for their support. This study was carried out by using SOR facilities of the Photon Factory at KEK and the Spring 8 at JASRI in Japan.

REFERENCES

1. Barrick, D., and Baldwin, R. L. (1993) Three-state analysis of sperm whale apomyoglobin folding, *Biochemistry 32*, 3790−3796.
2. Goto, Y., Takahashi, N., and Fink, A. L. (1990) Mechanism of acid-induced folding of proteins, *Biochemistry 29*, 3480−3488.
3. Goto, Y., Calciano, L. J., and Fink, A. L. (1990) Acid-induced folding of proteins, *Proc. Natl. Acad. Sci. U.S.A. 87*, 573−577.
4. Eliezer, D., Yao, J., Dyson, H. J., and Wright, P. E. (1998) Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding, *Nat. Struct. Biol. 5*, 148−155.
5. Yao, J., Chung, J., Eliezer, D., Wright, P. E., and Dyson, H. J. (2001) NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding, *Biochemistry 40*, 3561−3571.
6. Kataoka, M., Nishii, I., Fujisawa, T., Ueki, T., Tokunaga, F., and Goto, Y. (1995) Structural characterization of the molten globule and native states of apomyoglobin by solution X-ray scattering, *J. Mol. Biol. 249*, 215−228.
7. Tcherkasskaya, O., and Uversky, V. N. (2001) Denatured collapsed states in protein folding: Example of apomyoglobin, *Proteins 44*, 244−254.
8. Glatter, O., and Kratky, O. (1982) *Small Angle X-ray Scattering*, Academic Press, London, U.K.
9. Feigin, L. A., and Svergun, D. I. (1987) *Structure Analysis by Small-Angle X-ray and Neutron Scattering*, Plenum Press, New York.
10. Kataoka, M., and Goto, Y. (1996) X-ray solution scattering studies of protein folding, *Fold. Des. 1*, R107−R114.
11. Hirai, M., Iwase, H., Hayakawa, T., Miura, K., and Inoue, K. (2002) Structural hierarchy of several proteins observed by wide-angle solution scattering, *J. Synchrotron Radiat. 9*, 202−205.
12. Hirai, M., Koizumi, M., Hayakawa, T., Takahashi, H., Abe, S., Hirai, H., Miura, K., and Inoue, K. (2004) Hierarchical map of protein unfolding and refolding at thermal equilibrium revealed by wide-angle X-ray scattering, *Biochemistry 43*, 9036−9049.
13. Seki, Y., Tomizawa, T., Khechinashvili, N. N., and Soda, K. (2002) Contribution of solvent water to the solution X-ray scattering profile of proteins, *Biophys. Chem. 95*, 235−252.
14. Seki, Y., Tomizawa, T., Hiragi, Y., Inoue, K., and Soda, K. (2001) Solution structure and hydration states of proteins by precise analysis of X-ray scattering profiles in the wide Q-range, SPring-8 User Experiment Report Number 8, 175−175.
15. Goto, Y., and Fink, A. L. (1990) Phase diagram for acidic conformational states of apomyoglobin, *J. Mol. Biol. 214*, 803−805.
16. Fasman, G. D. (1989) *Practical Handbook of Biochemistry and Molecular Biology*, CRC Press, Boca Raton, FL.
17. Zimm, B. H. (1948) The scattering of light and the radial distribution function of high polymer solutions, *J. Chem. Phys. 16*, 1093−1099.
18. Flory, P. J., and Bueche, A. M. (1958) Theory of light scattering by polymer solutions, *J. Polym. Sci. 27*, 219−229.
19. Chalikian, T. V., Totrov, M., Abagyan, R., and Breslauer, K. J. (1996) The hydration of globular proteins as derived from volume and compressibility measurements: Cross correlating thermodynamic and structural data, *J. Mol. Biol. 260*, 588−603.
20. Guinier, A., and Fournet, G. (1955) *Small-Angle Scattering of X-Rays*, Wiley, New York.
21. Debye, P. (1947) Molecular-weight determination by light scattering, *J. Phys. Colloid Chem. 51*, 18−32.
22. Makhatadze, G. I., Medvedkin, V. N., and Privalov, P. L. (1990) Partial molar volumes of polypeptides and their constituent groups in aqueous solution over a broad temperature range, *Biopolymers 30*, 1001−1010.
23. Akaike, H. (1974) New look at the statistical model identification, *IEEE Trans. Automat. Control 19*, 716−723.
24. Calmettes, P., Durand, D., Desmadril, M., Minard, P., Receveur, V., and Smith, J. C. (1994) How random is a highly denatured protein? *Biophys. Chem. 53*, 105−113.
25. Petrescu, A. J., Receveur, V., Calmettes, P., Durand, D., and Smith, J. C. (1998) Excluded volume in the configurational distribution of a strongly-denatured protein, *Protein Sci. 7*, 1396−1403.
26. Semisotnov, G. V., Kihara, H., Kotova, N. V., Kimura, K., Amemiya, Y., Wakabayashi, K., Serdyuk, I. N., Timchenko, A. A., Chiba, K., Nikaido, K., Ikura, T., and Kuwajima, K. (1996) Protein globularization during folding. A study by synchrotron small-angle X-ray scattering, *J. Mol. Biol. 262*, 559−574.
27. Harpaz, Y., Gerstein, M., and Chothia, C. (1994) Volume changes on protein folding, *Structure 2*, 641−649.
28. Tamura, Y., and Gekko, K. (1995) Compactness of thermally and chemically denatured ribonuclease A as revealed by volume and compressibility, *Biochemistry 34*, 1878−1884.
29. Gekko, K., and Noguchi, H. (1979) Compressibility of globular proteins in water at 25 °C, *J. Phys. Chem. 83*, 2706−2714.
30. Gekko, K., and Hasegawa, Y. (1986) Compressibility−structure relationship of globular proteins, *Biochemistry 25*, 6563−6571.
31. Damaschun, G., Damaschun, H., Gast, K., Gernat, C., and Zirwer, D. (1991) Acid denatured apo-cytochrome *c* is a random coil: Evidence from small-angle X-ray scattering and dynamic light scattering, *Biochim. Biophys. Acta 1078*, 289−295.
32. Damaschun, G., Damaschun, H., Gast, K., and Zirwer, D. (1998) Denatured states of yeast phosphoglycerate kinase, *Biochemistry (Moscow) 63*, 259−275.
33. Gast, K., Damaschun, H., Misselwitz, R., Müller-Fohne, M., Zirwer, D., and Damaschun, G. (1993) Compactness of protein molten globules: Temperature-induced structural changes of the apomyoglobin folding intermediate, *Eur. Biophys. J. 23*, 297−305.
34. Flory, P. J. (1969) *Statistical Mechanics of Chain Moleclues*, John Wiley and Sons, New York.
35. Calmettes, P., Roux, B., Durand, D., Desmadril, M., and Smith, J. C. (1993) Configurational distribution of denatured phosphoglycerate kinase, *J. Mol. Biol. 231*, 840−848.
36. Petrescu, A. J., Calmettes, P., Durand, D., Receveur, V., and Smith, J. C. (2000) Change in backbone torsion angle distribution on protein folding, *Protein Sci. 9*, 1129−1136.
37. Kamatari, Y. O., Ohji, S., Konno, T., Seki, Y., Soda, K., Kataoka, M., and Akasaka, K. (1999) The compact and expanded denatured conformations of apomyoglobin in the methanol−water solvent, *Protein Sci. 8*, 873−882.
38. Hiragi, Y., Seki, Y., Ichimura, K., and Soda, K. (2002) Direct detection of the protein quaternary structure and denatured entity by small-angle scattering: Guanidine hydrochloride denaturation of chaperonin protein GroEL, *J. Appl. Crystallogr. 35*, 1−7.
39. Higurashi, T., Hiragi, Y., Ichimura, K., Seki, Y., Soda, K., Mizobata, T., and Kawata, Y. (2003) Structural stability and solution structure of chaperonin GroES heptamer studied by synchrotron small-angle X-ray scattering, *J. Mol. Biol. 333*, 605−620.

BI061578+