# Hierarchical Map of Protein Unfolding and Refolding at Thermal Equilibrium Revealed by Wide-Angle X-ray Scattering

Mitsuhiro Hirai,*,‡ Masaharu Koizumi,‡ Tomohiro Hayakawa,‡ Hiroshi Takahashi,‡ Satoru Abe,‡ Harutaka Hirai,‡ Keiko Miura,§ and Katsuaki Inoue§

*Department of Physics, Gunma University, Maebashi 371-8510, Japan, and Japan Synchrotron Radiation Research Institute, Sayo, Hyogo 679-5198, Japan*

*Received January 5, 2004; Revised Manuscript Received April 12, 2004*

ABSTRACT: Hierarchical features of the thermal unfolding−refolding structural transition of hen egg white lysozyme (HEWL) have been studied in the temperature range from 13 to 84 °C by using high-resolution wide-angle X-ray scattering (WAXS) measurements at a third-generation synchrotron source. We have gathered high-statistic WAXS data of the reversible unfolding−refolding process of HEWL in the $q$ range from ∼0.05 to ∼3 Å$^{-1}$ [$q = (4\pi/\lambda)\sin(\theta/2)$, where $\theta$ is the scattering angle and $\lambda$ the wavelength]. This measured $q$ range corresponds to the spatial distance from ∼2 to ∼125 Å, which covers all hierarchical structures of a small globular protein such as HEWL, namely, tertiary, domain, and secondary structures. Because of this, we have found that the pH dependence of the thermal structural transition of HEWL is well characterized by the various hierarchical levels and the transition concurrence among them. In this report, we present a new hierarchical map depiction of unfolding−refolding transitions. Using scattering with various ranges of $q$ values, we determine the molar ratio of native-like protein structure defined by the data in each range, thus producing a map of the amount of native-like structure as a function of the hierarchical level or resolution. This map can visualize a detailed feature of the unfolding−refolding transition of a protein depending on various structural hierarchical levels; however, the exact meaning of the map will await sharpening by additional works.

The mechanism of protein folding is one of the major problems in structural biology. Numerous studies on protein folding have been executed through advances both in experimental methods and in theoretical approaches (*1*). Especially the "folding funnel" or "energy landscape" model presented by theoretical studies affords us important physical insight into protein folding kinetics and stability (*1*−*3*). For clarifying how proteins actually fold into those native structures, it is important to observe the detailed nature of the folding−unfolding processes. The experimental resolutions in time and space have also been improving continuously. In one such experimental improvement, high-brilliant X-ray radiation from a third-generation synchrotron radiation (SR)[1] source has opened a new horizon not only in protein crystallography but also in protein structure analyses in solutions. A recent time-resolved small-angle X-ray scattering (SAXS) study using a third-generation SR source has successfully demonstrated that a collapse process of a folding protein at unequilibrium can now be observed at time and spatial resolutions of ∼1 ms and ∼30 Å, respectively, in conjunction with a special microfabricated fluid mixing technique (*4*), where the spatial resolution is defined by $2\pi/q_{max}$ ($q_{max}$ is a maximum measured value). Despite the use of high-brilliant SR X-ray radiation, their experiment would suggest some difficulty in folding kinetics studies in obtaining internal structural information about proteins in solutions. This difficulty could be attributed to an intrinsic feature of structural factors of proteins, namely, to ∼100-fold damping of scattering intensities at high-angle scattering regions. Alternatively, even by using high-brilliant X-ray radiation in protein folding studies, the experimental demand for the time resolution still competes with that for the spatial resolution. On the other hand, under equilibrium conditions, high-spatial resolution scattering data of proteins in solutions can be measured by using X-rays from a third-generation SR source. We have recently shown that SR wide-angle X-ray scattering (SR-WAXS) curves of several typical globular proteins in solutions characterize well not only overall structures (molecular size and shape) but also intramolecular structures such as intradomain structures, domain−domain distance correlation, and secondary structures (*5*). We have also shown that the full-range experimental SR-WAXS curves of the proteins were mostly comparable with the theoretical ones calculated from the atomic coordinates of the proteins.

In this report, we treat the thermal structural transition of hen egg white lysozyme (HEWL). Calorimetric studies

* To whom correspondence should be addressed: Department of Physics, Gunma University, 4-2 Aramaki, Maebashi 371-8510, Japan. Telephone: +81-27-220-7554. Fax: +81-27-220-7551. E-mail: mhirai@fs.aramaki.gunma-u.ac.jp.

‡ Gunma University.

§ Japan Synchrotron Radiation Research Institute.

[1] Abbreviations: HEWL, hen egg white lysozyme; SAXS, small-angle X-ray scattering; WAXS, wide-angle X-ray scattering; SR, synchrotron radiation; JASRI, Japan Synchrotron Radiation Research Institute; DSC, differential scanning calorimetry; PDB, Protein Data Bank; $R_g$, radius of gyration; CD, circular dichroism; NMR, nuclear magnetic resonance; FTIR, Fourier transform infrared; SCOP, Structure Classification Of Proteins; SH, structural hierarchy.

10.1021/bi0499664 CCC: $27.50 
Published on Web 06/25/2004

clarified the thermodynamic basis of the conformational stability of proteins at equilibrium states. They showed that the thermal transitions of proteins, especially single-domain proteins, are essentially two-state where only fully folded and unfolded states are populated (*6*−*10*). Those studies suggested that HEWL is a typical protein showing a cooperative two-state thermal transition, whereas later spectroscopic studies showed that HEWL consists of the two structural domains which differ in the folding pathway and kinetics (*11*−*16*). In the previous studies using small- and medium-angle X-ray scattering methods combined with the calorimetric measurements, we showed that the thermal unfolding transition of HEWL greatly depends on pH and on its structural hierarchy (*17*−*20*); however, the spatial resolution was greater than ∼12 Å. For a detailed understanding of the thermal unfolding−refolding transition of HEWL, it is necessary to obtain X-ray scattering data with much higher statistical and spatial resolution which covers the whole structural hierarchy under equilibrium conditions. Then we have carried out the SR-WAXS experiments on the thermal unfolding−refolding process of HEWL at various pH values under the stepwise heating and cooling conditions. We have successfully observed the high-statistic and high-resolution WAXS data of the thermal unfolding−refolding transition of HEWL with a high reversibility. These observed WAXS data covering the spatial resolution from ∼2 to ∼125 Å have directly afforded us several features of the thermal unfolding−refolding structural transition for various hierarchical levels and for the transition concurrence and cooperativity among those levels at equilibrium. In addition, we propose a new hierarchical map for analysis of unfolding−refolding transitions of proteins, which will show further applicability of the SR-WAXS method in protein folding studies at equilibrium.

## MATERIALS AND METHODS

*Materials*. Hen egg white lysozyme (crystallized three times) was purchased from Sigma Chemical Co. and was used for WAXS measurements. In the unfolding and/or refolding process of proteins, some aggregation tends to occur in many cases. As reported previously (*17*−*21*), to obtain high-statistic WAXS data of unfolding−refolding structural transitions of proteins while avoiding some aggregation, repulsive interparticle interaction between proteins through Coulomb potential can be productively used. Then we dissolved the protein in a low-ionic strength water solvent whose pH was adjusted by adding HCl before the dissolution. The final pH values of the protein solutions were determined by using a digital HM-60V pH meter from TOA Electronics Ltd. The pH values of the protein solutions used for the thermal transition measurements were 2.2, 3.1, 3.6, and 4.5. The concentrations of the samples were 5% (w/v) for the thermal unfolding−refolding measurements and 1% (w/v) for the pH dependence measurements.

*X-ray Scattering Measurements*. Wide-angle X-ray scattering experiments were carried out by using an X-ray scattering spectrometer installed at BL-40B2 of the 8 GeV synchrotron radiation source in the Japan Synchrotron Radiation Research Institute (JASRI, Harima, Japan). The sample-to-detector distance was 41 cm. The X-ray wavelength was 0.729 Å. The beam size at the sample position was ∼0.1 mm × 0.1 mm in the full width at half-maximum. X-ray scattering intensity was detected with an imaging plate (30 cm × 30 cm in area) system from RIGAKU (R-AXIS IV). Incident and transmitted X-ray beam intensities were monitored with a pair of ionization chambers. The details of the spectrometer were described previously (*22*). The sample solutions were contained in the sample cell with a path length of 1 mm and a pair of thin quartz windows. In the thermal unfolding−refolding measurements, the temperature of the sample was changed stepwise from ∼13 to ∼84 °C. The time of exposure to X-rays at each temperature was 30 s. The sample solution was slightly moved in the cell for each measurement. This oscillatory movement of the solution was enough to avoid some radiation damage on the sample due to the small beam size (∼0.1 mm × 0.1 mm). After a beam center had been determined using a standard sample, silver behenate (*23*), two-dimensional scattering data recorded on IP were transformed to a one-dimensional scattering intensity profile $I(q)$ by using a circle-averaging program. The $q$ values were calibrated with the diffraction data of the standard sample described above.

*Scattering Data Analysis*. In WAXS data analyses of diluted solutions containing such proteins, the correct subtraction of a solvent scattering from a solution one is very important in obtaining a net scattering from solute particles. The structure factor of water increases above a $q$ of ∼0.8 Å$^{-1}$ and shows a broad correlation peak at ∼2.0 Å$^{-1}$ (*24*). In the experiments described herein, the position of the correlation peak shifted from 1.95 Å$^{-1}$ at 13 °C to 2.05 Å$^{-1}$ at 84 °C. Then in the background correction of WAXS data, we took into account both the excluded volume effect of protein molecules and the correlation peak of water molecules at each temperature. The following equation can be used to obtain the scattering $I(q)$ from the proteins (*5*).

$$I(q) = \frac{1}{I_{\mathrm{peak}}^{\mathrm{sol}}}\left[\frac{1}{B_{\mathrm{sol}}T_{\mathrm{sol}}}I_{\mathrm{sol}}(q) - \frac{1}{B_{\mathrm{celll}}T_{\mathrm{celll}}}I_{\mathrm{cell}}(q)\right] - \frac{(1 - c\upsilon_{\mathrm{a}})}{I_{\mathrm{peak}}^{\mathrm{solvl}}}\left[\frac{1}{B_{\mathrm{solv}}T_{\mathrm{solv}}}I_{\mathrm{solv}}(q) - \frac{1}{B_{\mathrm{cell}}T_{\mathrm{cell}}}I_{\mathrm{cell}}(q)\right] \quad (1)$$

where $I_{\mathrm{sol}}(q)$, $I_{\mathrm{solv}}(q)$, $I_{\mathrm{cell}}(q)$, $B_{\mathrm{sol}}$, $B_{\mathrm{solv}}$, $B_{\mathrm{celll}}$, $T_{\mathrm{sol}}$, $T_{\mathrm{solv}}$, and $T_{\mathrm{celll}}$ are the observed scattering intensities, the incident beam intensities, and the transmissions of the solution and solvent samples at each temperature and the sample cell, respectively. $I_{\mathrm{peak}}^{\mathrm{sol}}$ and $I_{\mathrm{peak}}^{\mathrm{solvl}}$ are the water correlation peak intensities of the solution and solvent samples and the solvent, respectively. $c$ and $\upsilon_{\mathrm{a}}$ are the concentration of the protein molecule and its partial specific volume, respectively. A $\upsilon_{\mathrm{a}}$ value of 0.712 mL/g was used for lysozyme (*25*). After these data corrections had been executed, the radius of gyration $R_{\mathrm{g}}$ was determined by using the following equation (*26*).

$$R_{\mathrm{g}}^{2} = \frac{\int_0^{D_{\max}} p(r) r^2 \, \mathrm{d}r}{2\int_0^{D_{\max}} p(r) \, \mathrm{d}r} \quad (2)$$

where the $p(r)$ is the distance distribution function calculated by the Fourier inversion of the scattering curve $I(q)$ as

$$p(r) = \frac{2}{\pi}\int_0^{\infty} rqI(q) \sin(rq) \, \mathrm{d}q \quad (3)$$

and $D_{max}$ is the maximum diameter of the solute particle determined by the $p(r)$ function satisfying the condition $p(r) = 0$ when $r > D_{max}$. In the estimation of $R_g$, the use of eq 2 is known to reduce some inherent systematic errors caused by concentration or aggregation effects in comparison with the use of the Guinier approximation (*27*). In previous reports, we showed that the following method, called transition-multiplicity analysis (TMA) (*18*), can be applied well in analyzing a hierarchical feature of a structural transition of proteins.

$$\Delta_{ij} = \sum_{q=q_i}^{q_j} \left| \frac{I(q,T)}{\sum_{q=q_i}^{q_j} I(q,T)} - \left[ \alpha \frac{I(q,T_N)}{\sum_{q=q_i}^{q_j} I(q,T_N)} + (1-\alpha) \frac{I(q,T_U)}{\sum_{q=q_i}^{q_j} I(q,T_U)} \right] \right| \quad (4)$$

where $I(q,T_N)$, $I(q,T_U)$, and $I(q,T)$ are the scattering profiles at the initial, final, and intermediate temperatures in a defined $q$ range of $q_i-q_j$ Å$^{-1}$, respectively. In this case, for each pH the scattering curve $I(q,T)$ in a defined $q$ range at an intermediate temperature was fitted by using a linear combination of the $\alpha I(q,T_N)$ and $(1-\alpha)I(q,T_U)$ at initial and final temperatures. The factor $\alpha$ is determined by minimizing the $\Delta_{ij}$ value in eq 4. Therefore, the determined $\alpha$ and $1-\alpha$ values at the intermediate temperature $T$ correspond to the molar fractions of the spatial conformations of the proteins at initial and final temperature states in the defined $q$ range of $q_i-q_j$. Alternatively, we assume that at equilibrium the structural transition of the proteins at the transition intermediate can be described by a population of the proteins at folded (N-state) and unfolded (U-state) states; therefore, the molar fraction $\alpha$ was 1 at the initial temperature and 0 at the final temperature. As previously discussed in detail with the SAXS and differential scanning calorimetry (DSC) experiments (*18*), the use of eq 4 is comparable with the use of the van't Hoff equation in DSC analyses that clarify thermodynamic functions and the fractions of the thermodynamic microstates as an ensemble of microstates at equilibrium (*28*−*30*).

*Theoretical Scattering Curve Calculation.* To simplify the explanation of WAXS curves of proteins, we will omit the description of the presence of solvent that gives a solute particle an excess average scattering density, so-called contrast (*31*). Then, the theoretical WAXS curve of a protein can be calculated by using the following Debye equation (*32*) and the atomic coordinates of the protein in the Protein Data Bank (PDB).

$$I(q) = \sum_{i}^{n}\sum_{j}^{n} f_i f_j \frac{\sin(qr_{ij})}{qr_{ij}} \quad (5)$$

where $f_i$ and $f_j$ are the atomic scattering factors of $i$th and $j$th atoms, respectively, in a molecule composed of $n$ atoms and $r_{ij}$ is the interatomic distance between the $i$th and $j$th atoms. According to the convolution theorem, the scattering amplitude $A(\mathbf{q})$ of a protein composed of two domains ($\alpha$-domain and $\beta$-domain in a protein such as HEWL) is given as

$$\begin{aligned} A(\mathbf{q}) &= \int \{ [\sum_{\alpha\text{-domain}} \rho_i * \delta(\mathbf{r}-\mathbf{r}_i)] * \delta(\mathbf{r}-\mathbf{r}_\alpha) + \\ &\quad [\sum_{\beta\text{-domain}} \rho_i * \delta(\mathbf{r}-\mathbf{r}_i)] * \delta(\mathbf{r}-\mathbf{r}_\beta) \} \exp(i\mathbf{q}\cdot\mathbf{r})\, d\mathbf{r} \\ &= [\sum_{\alpha\text{-domain}} f_i \exp(i\mathbf{q}\cdot\mathbf{r}_i)] \exp(i\mathbf{q}\cdot\mathbf{r}_\alpha) + \\ &\quad [\sum_{\beta\text{-domain}} f_i \exp(i\mathbf{q}\cdot\mathbf{r}_i)] \exp(i\mathbf{q}\cdot\mathbf{r}_\beta) \quad (6) \end{aligned}$$

where $*$ is the convolution integral, $\rho_i$ and $\mathbf{r}_i$ represent the electron density distribution function and its center coordinate of the $i$th atom in the protein, respectively, and $\mathbf{r}_\alpha$ and $\mathbf{r}_\beta$ are the electron density center coordinates of the $\alpha$- and $\beta$-domains, respectively. For proteins randomly oriented in solutions, an observed scattering curve is obtained by averaging the scattering intensity $I(\mathbf{q})$ over the solid angle in reciprocal space as follows.

$$\begin{aligned} I(\mathbf{q}) = \langle I(\mathbf{q}) \rangle &= \langle |A(\mathbf{q})|^2 \rangle = \langle |A_\alpha(\mathbf{q})|^2 \rangle + \langle |A_\beta(\mathbf{q})|^2 \rangle + \\ &\quad 2\langle \mathcal{R}\{A_\alpha(\mathbf{q}) A_\beta * (\mathbf{q}) \exp[i\mathbf{q}\cdot(\mathbf{r}_\alpha - \mathbf{r}_\beta)]\} \rangle \quad (7) \end{aligned}$$

where the broken brackets indicate the spherical average integral and $A_\alpha(\mathbf{q})$ and $A_\beta(\mathbf{q})$ are the scattering amplitudes of the two domains. The first and second terms in eq 7 are also given by the following Debye equations for two domains.

$$\langle |A_\alpha(\mathbf{q})|^2 \rangle = \sum_{\alpha\text{-domain}} \sum f_i f_j \frac{\sin(qr_{ij})}{qr_{ij}}, \langle |A_\beta(\mathbf{q})|^2 \rangle = \sum_{\beta\text{-domain}} \sum f_i f_j \frac{\sin(qr_{ij})}{qr_{ij}} \quad (8)$$

The third term in eq 7 represents the correlation scattering function between the domains where $r_{\alpha\beta}$ ($=|\mathbf{r}_\alpha - \mathbf{r}_\beta|$) is the distance between the electron density centers of the $\alpha$- and $\beta$-domains. This term is an oscillation function with both positive and negative values. On the basis of a theoretical scattering curve calculation using the formulations described above, we are able to discuss qualitatively characteristics of WAXS curves depending on structural hierarchy, especially for small proteins that are composed of a few domains structured separately such as HEWL. It should be mentioned that in many cases theoretical scattering curves using the Debye equation cannot quantitatively reproduce experimental ones for small- and medium-angle scattering regions ($q < \sim 0.5$ Å$^{-1}$). This is because the PDB data contain the atomic coordinates of a limited number of hydrated waters and because the Debye equation is hard to take into account for both contrasts of a protein and its hydration shell in an appropriate way. Experimental scattering curves are most appropriately reproduced by theoretical ones using a well-known program called CRYSOL (*33*). This program using the multipole expansion method (*31*) takes into account of the existence of a solvent and hydration shell based on the PDB atomic coordinates of a protein (*33*, *34*).

## RESULTS

*WAXS Curves and Hierarchical Structures of Proteins.* As reported previously for several typical proteins in

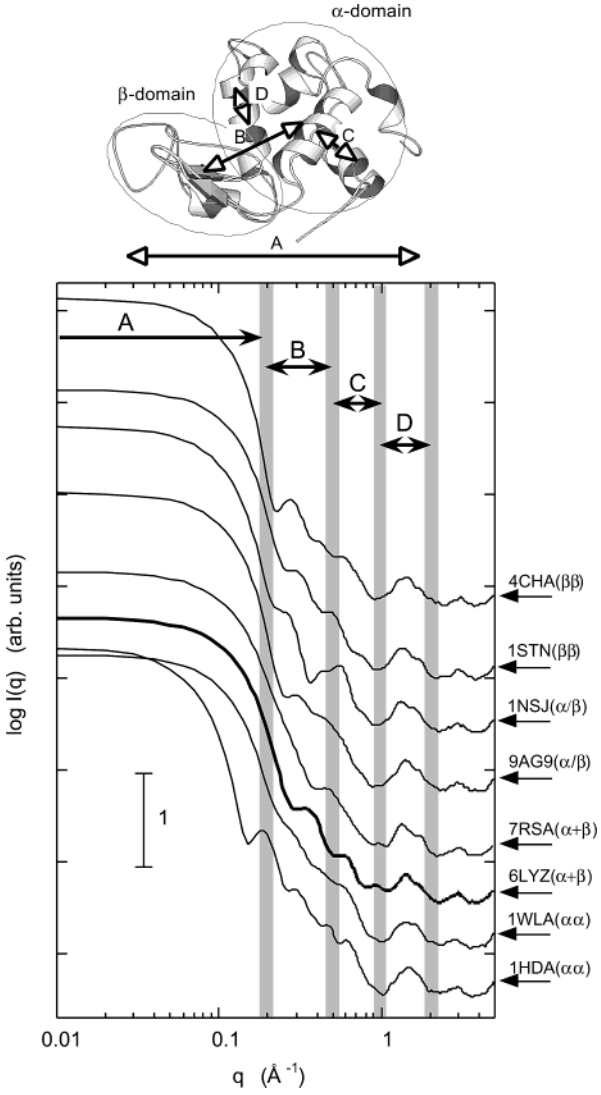


FIGURE 1: Theoretical WAXS curves, obtained by the Debye equation, of several proteins in different structural classification categories. Typical intramolecular distances within the HEWL molecule are shown schematically in its three-dimensional rendering using MolScript (*68*). The scattering curves in the regions of parts A ($q < \sim 0.2$ Å$^{-1}$), B ($\sim 0.25$ Å$^{-1} < q < \sim 0.5$ Å$^{-1}$), C ($\sim 0.5$ Å$^{-1} < q < \sim 0.8$ Å$^{-1}$), and D ($\sim 1.1$ Å$^{-1} < q < \sim 1.9$ Å$^{-1}$) mostly correspond to different hierarchical structure levels, that is, to the quaternary and tertiary structures, the interdomain correlation and intradomain structure, and the secondary structures, including the closely packed side chains, respectively. The schematic picture and the WAXS curve of HEWL are reproduced from our previous report (*5*).

solutions (*5*), a WAXS curve of a protein mostly reflects the characteristics of hierarchical structural levels within the protein. Figure 1 shows the theoretical WAXS curves of several globular proteins with the HEWL schematic three-dimensional structure. The WAXS were simply calculated by using eq 5 and the atomic coordinates of the proteins in the PDB. The proteins were selected according to the database of Structure Classification Of Proteins called SCOP (*36*). Namely, hemoglobin (1HDA) (*37*) and myoglobin (1WLA) (*38*) are classified as all-α proteins, α-chymotrypsin (4CHA) (*39*) and staphylococcal nuclease (1STN) (*40*) as all-$\beta$ proteins, phoshoribosyl anthranilate isomerase (1NSJ) (*41*) and flavodoxins (1AG9) (*42*) as α/$\beta$ proteins, and lysozyme (6LYZ) (*43*) and ribonuclease A (7RSA) (*44*) as

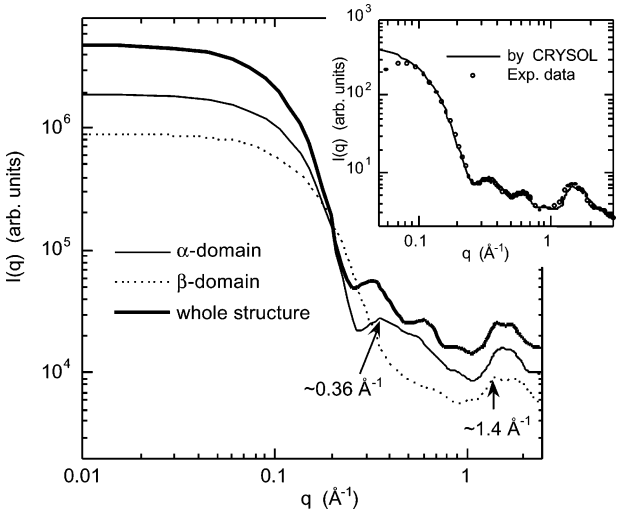


FIGURE 2: Theoretical scattering curves of the HEWL whole structure (thick line), the α-domain (thin line), and the $\beta$-domain (dotted line) which were obtained by using CRYSOL. The inset shows the observed WAXS curve [5% (w/v) HEWL at pH 4.5 and 24 °C] with the theoretical one calculated by the program fitting the experimental data: experimental data (○) and theoretical curve (—).

α+$\beta$ proteins (PDB entries used for the theoretical calculations). The previous experimental WAXS curves (*5*) and the theoretical ones in Figure 1 suggest that the scattering curves in the ranges of $\sim 0.8-1.1$ and $>1.9$ Å$^{-1}$ are less sensitive to the characteristics of the internal structures of proteins. The regions in the theoretical scattering curves are roughly assigned to the typical intramolecular distances of the different hierarchical structural levels in the HEWL three-dimensional structure. In the case of HEWL, the scattering curves in the $q$ range of less than $\sim 0.2$ Å$^{-1}$, in the $q$ range of $\sim 0.25-0.8$ Å$^{-1}$, and in the $q$ range of $\sim 1.1-1.9$ Å$^{-1}$ mostly correspond to the different hierarchical levels, namely, to the tertiary structure, the interdomain correlation and intradomain structure, and the secondary structures and the closely packed side chains in the hydrophobic cores, respectively.

For many proteins, the theoretical scattering curves given by the CRYSOL program are in agreement with the experimental WAXS ones over a wide $q$ range (*5*) since this program takes into account contrast and hydration shell (*33*, *34*). Figure 2 shows the theoretical scattering curves of HEWL using the CRYSOL program. In the program, we used 50 spherical harmonics and other default parameters, and PDB entry 6LYZ of HEWL was used. The inset in Figure 2 shows good agreement of the experimental WAXS curve [5% (w/v) HEWL at pH 4.5 and 24 °C] with the theoretical one using this program, where only a limited number of experimental data points are plotted. This would ensure the reasonability of the following hierarchical analyses for HEWL structure and stability. According to the SCOP database (*36*), HEWL is in the α+$\beta$ group of proteins and is known to be composed of two structural domains, that is, the α-domain with exclusively α-helical structure and the $\beta$-domain with predominantly $\beta$-structure. The theoretical scattering curves in Figure 2 were calculated by using the atomic coordinates of the HEWL whole structure, the α-domain, and the $\beta$-domain. In the scattering curve of the α-domain, the damped trapezoidal profile in the $q$ range of

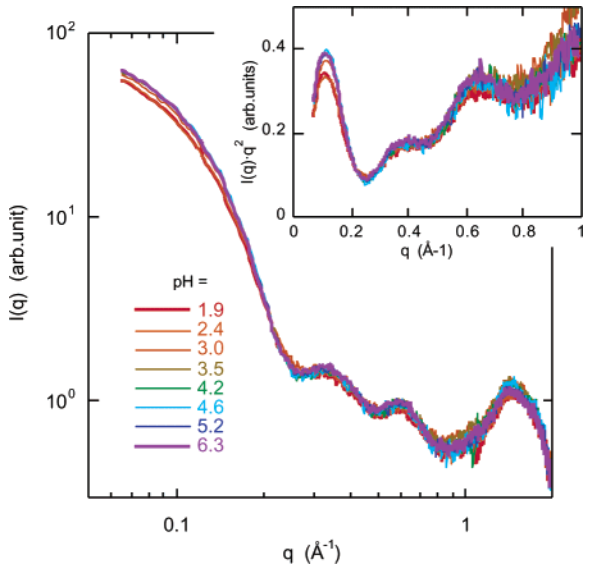


FIGURE 3: pH dependence of the WAXS curve $I(q)$ of 1% (w/v) HEWL at 25 °C. The pH was varied from 1.9 to 6.3, and the inset shows the Kratky plots of $I(q)$.

$\sim$0.27−0.7 Å$^{-1}$ with a maximum around $\sim$0.36 Å$^{-1}$ characterizes the internal structure such as the α-helix packing. The scattering curve of the α-domain in the $q$ range of $\sim$1.2−2.0 Å$^{-1}$ has a round hump, which can be seen in several proteins with exclusively α-helical structures, such as in all-α proteins (*5*). On the other hand, the scattering curve of the $\beta$-domain is simply dumped in the $q$ range of $\sim$0.27−0.7 Å$^{-1}$ without any correlation peak. This would reflect the fact that the $\beta$-domain is a small domain with less secondary structure packing correlation. In the $q$ range of $\sim$1.2−2.0 Å$^{-1}$, the scattering curve of the $\beta$-domain has a broad hump with a maximum around $\sim$1.4 Å$^{-1}$. Similar maxima are found in several all-$\beta$ proteins (examples in Figure 1). The scattering curves in Figure 2 can be explained theoretically by using eqs 6−8 based on the convolution theorem. Thus, the scattering curves of the α-domain and $\beta$-domain in Figure 2 correspond to the first and second terms in eq 7 (namely, in eq 8), respectively. The scattering curve of the whole molecule in Figure 2 corresponds to the summation of all terms in eq 7. Noticeably, the scattering curve of the HEWL whole structure in this $q$ region is mostly given by the simple summation of the scattering curves of the α- and $\beta$-domains, and the maximum $q$ characterizing the $\beta$-domain is $\sim$1.4 Å$^{-1}$. The maximum $q$ of $\sim$1.4 Å$^{-1}$ is actually observed in the experimental scattering curve in the inset. This is because the interdomain correlation function of the third term in eq 7 has a factor of $\sin(qr_{\alpha\beta})/qr_{\alpha\beta}$ which damps faster against $q$ than the first and second terms in eq 7 given by eq 8. From the theoretical scattering curves in the $q$ range of 1.2−2.0 Å$^{-1}$, we can estimate separately the contributions of three scattering terms (α-domain, $\beta$-domain, and the correlation between them) in eq 7 to the total scattering intensity in this region. Those values are 61.8% for the α-domain, 38.3% for the $\beta$-domain, and −0.1% for the correlation. It would also explain the simple summation. The above characterizations of the WAXS curve of HEWL are important in the following analyses.

*pH Dependence of HEWL Structure*. Figure 3 shows the pH dependence of the WAXS curve $I(q)$ of 1% (w/v) HEWL at 25 °C and pH 1.9−6.3, where the inset shows the Kratky plots of $I(q)$. HEWL consists of two structural domains whose average distance is $\sim$18 Å (*35*). As described above, the broad humps around 0.34 and 0.62 Å$^{-1}$ mostly reflect the domain−domain correlation and the internal structures of the domains. Despite the change in pH, in Figure 3 the $I(q)$ profiles at all pH values mostly agree with each other, indicating that the pH variation from 6.3 to 1.9 at 25 °C slightly affects the whole HEWL structure. There is a minor change in the scattering curve depending on pH. At pH 4.6−1.9, the slope $I(q)$ when $q < \sim$0.15 Å$^{-1}$ becomes smaller with a slight decrease in the zero-angle scattering intensity $I(0)$. $I(0)$ is proportional to $n(\bar{\rho}V)^2$, where $n$ is the number of solute particles, $\bar{\rho}$ the contrast of solute particles, and $V$ the particle volume (*31*). The radius of gyration $R_g$ decreases from 15.6 ± 0.4 to 15.1 ± 0.3 Å, accompanying the $\sim$14% decrease in $I(0)$. As is well-known (*31*), the $R_g$ value depends not only on the molecular size and shape but also on the scattering density distribution within the molecule. In addition, a hydration shell density of a protein also affects both $R_g$ and $I(0)$ values (*34*). As reported previously (*20*), an interparticle interaction does not affect the scattering curve when $q > \sim$0.05 Å$^{-1}$ at <1% (w/v) HEWL. Then the observed changes in the $R_g$ and $I(0)$ values could be attributed to a change in the tertiary structure, the hydration shell, or both. In the Kratky plots in the inset, a slight decrease in the peak height when $q \sim$ 0.12 Å$^{-1}$ from pH 4.6 to 3.5 might indicate a slight change in the compactness of the protein structure (*45*, *46*). This accompanies a slight damping of the ripple profiles when $q \sim$ 0.2−0.7 Å$^{-1}$, suggesting that the pH variation induces a minor change both in the interdomain correlation and the internal structures of the domains. The scattering intensity when $q \sim$ 1.5 Å$^{-1}$ tends to decrease below pH 3.5, which suggests some change in the secondary structure and the closely packed side chains. However, these changes depending on pH are relatively small at pH >2, which results from an intrinsic characteristic of HEWL, in contrast to the cases of other acid-denatured proteins at low pH (*47*−*49*). Thus, under the experimental conditions presented here, the observed WAXS curves show that the HEWL retains mostly its native-like structure even at low pH.

*Thermal Unfolding and Refolding of HEWL Hierarchical Structure*. Protein solutions with low concentrations are usually preferable for avoiding some effect of interparticle interaction on scattering data, especially for the small-angle scattering region. However, in Figure 3 the signal-to-noise (S/N) ratio of the WAXS data when $q >$ 1 Å$^{-1}$ is not sufficiently high at 1% (w/v) HEWL. To examine the detailed features of the structural transition of HEWL using the WAXS method, scattering data covering the secondary structural level should be collected at a higher S/N ratio under unaggregative conditions. Then we used 5% (w/v) HEWL solutions for the thermal transition measurements. Figure 4 shows the temperature dependence of $I(q)$ at different pH values, where labels A−C correspond to pH 2.2, 3.6, and 4.5 and labels a and b correspond to the heating (unfolding) and cooling (refolding) processes, respectively. The broad peak at $q \sim$ 0.08 Å$^{-1}$ is ascribed to the repulsive interparticle interaction among HEWL molecules. As shown experimentally (*20*), this correlation peak hardly affects the WAXS profile above $q = \sim$0.1 Å$^{-1}$ for 5% (w/v) HEWL solutions. At pH >5, the complete refolding of HEWL in the cooling

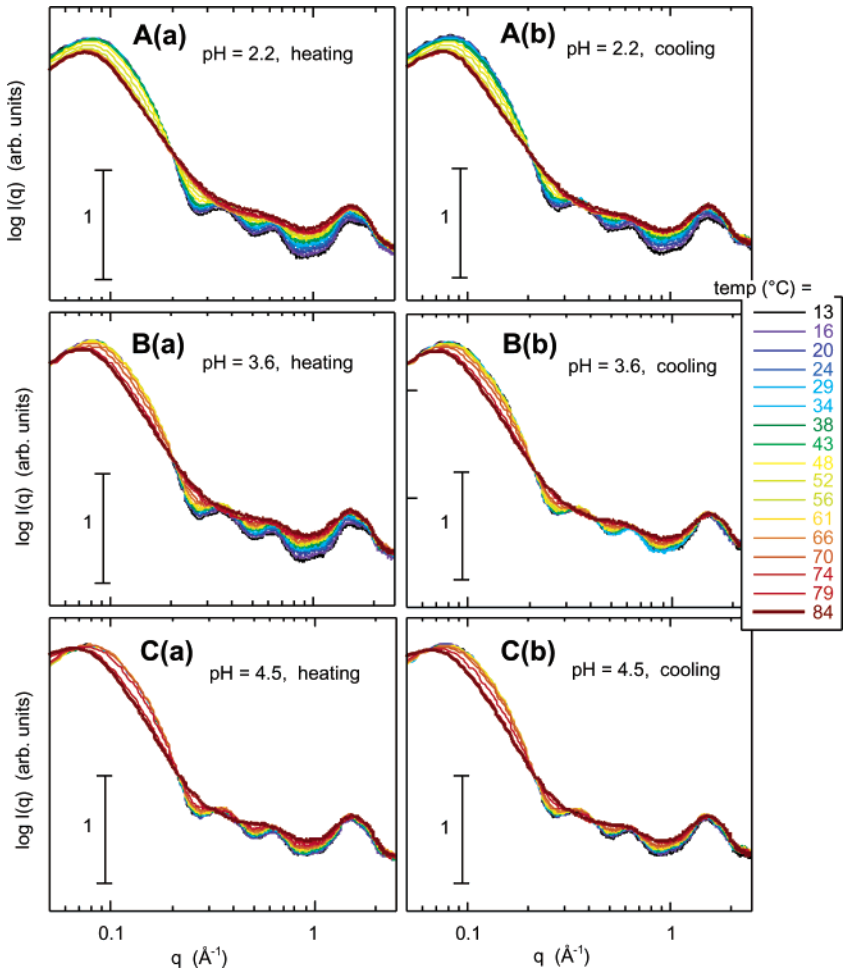


FIGURE 4: Temperature dependence of the WAXS curve $I(q)$ of 5% (w/v) HEWL at different pH values. Labels A−C correspond to pH 2.2, 3.6, and 4.5 and labels a and b to the heating (unfolding) and cooling (refolding) processes, respectively.

process was disturbed by the creation of some aggregates of unfolded HEWL molecules at the initial temperature (~84 °C) (*20*). At pH <5, the scattering profile of HEWL shows a high thermal reversibility between the heating and cooling processes, as shown in Figure 4. On the other hand, we can recognize a significant difference in the changing manners of the $I(q)$ curves at different pH values. Thus, when the pH is reduced, the $I(q)$ changes rather gradually with temperature in comparison with that at pH 4.5.

To consult this feature in detail, panels A−C of Figures 5 (heating process) and 6 (cooling process) display separately the $I(q)$ curves at pH 2.2 and 4.5 in three different scattering regions: $q < \sim 0.2$ Å$^{-1}$, $\sim 0.25$ Å$^{-1} < q < \sim 0.7$ Å$^{-1}$, and $\sim 1$ Å$^{-1} < q < \sim 3$ Å$^{-1}$, respectively. These selected $q$ regions mostly correspond to the different hierarchical structure levels of HEWL as mentioned above. In Figures 5 and 6, one $I(q)$ curve at 13 °C and five $I(q)$ curves around the transition midpoint temperatures are plotted. Two $I(q)$ curves plotted in thick lines show the maximum temperature differential value defined by $\Delta I(q)/\Delta T$. Here we tentatively define the intermediate value between the temperatures of the two curves as the transition midpoint temperature for each $q$ region. In the heating process, at pH 4.5 the transition midpoint temperatures for all different regions are the same (~77 °C) and at pH 2.2 the transition midpoint temperatures depend on the regions: ~54 °C for $q < \sim 0.2$ Å$^{-1}$, ~46 °C for $\sim 0.25$ Å$^{-1} < q < \sim 0.7$ Å$^{-1}$, and ~22 °C for $\sim 1$ Å$^{-1} < q < \sim 3$ Å$^{-1}$, respectively. At all pH values, the transition midpoint temperatures in the cooling process for different regions are mostly the same in the heating process. At pH 3.1 (data not shown), the transition midpoint temperatures in the heating and cooling processes are as follows: ~68 °C for $q < \sim 0.2$ Å$^{-1}$, ~64 °C for $\sim 0.25$ Å$^{-1} < q < \sim 0.7$ Å$^{-1}$, and ~27 °C for $\sim 1$ Å$^{-1} < q < \sim 3$ Å$^{-1}$; at pH 3.6 (data not shown), the transition midpoint temperatures are mostly the same (~74 °C) for all $q$ regions. In other words, Figures 5 and 6 suggest that the thermal transition for each $q$ region, namely, for each hierarchical level, proceeds reversibly at all pH values, and that the transition cooperativity or concurrency among the different hierarchical levels tends to be lost with a reduced pH.

*Radius of Gyration and Transition Temperature of the Tertiary Structure.* Figure 7 shows the temperature dependence of the radii of gyration $R_g$ at different pH values obtained from eq 2. An $R_g$ obtained by SAXS is usually used to estimate a compactness of tertiary structure of proteins in folding studies (*45*, *46*). From Figure 7, we determined the apparent values of the onset temperature $T_o$ and the midpoint temperature $T_m$ of the tertiary structure transition. The $T_o$ value was estimated from the temperature around a first increase in $R_g$ beyond the experimental error by sight. The $T_m$ value was determined by the temperature at the intermediate value between the average $R_g$ values at the lowest and highest temperatures. As reported previously (*17*, *18*,

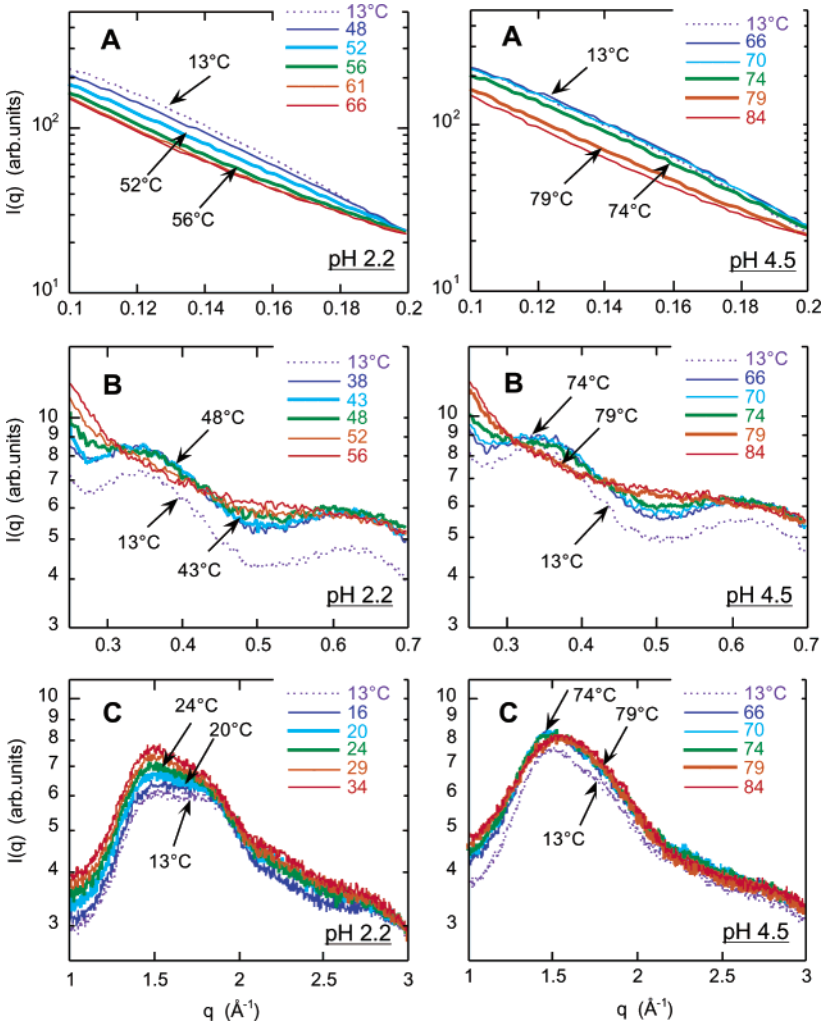


FIGURE 5: Selected WAXS curves $I(q)$ at pH 2.2 and 4.5 around the transition midpoint in the heating (unfolding) process. The thick lines show the maximum temperature differential values $[\Delta I(q)/\Delta T]$. Labels A−C correspond to the scattering curves in the different $q$ ranges: $q < \sim 0.2$ Å$^{-1}$, $\sim 0.25$ Å$^{-1} < q < \sim 0.8$ Å$^{-1}$, and $\sim 1$ Å$^{-1} < q < \sim 3$ Å$^{-1}$, respectively.

*20*), $T_o$ and $T_m$ values determined from the change in $R_g$ are strongly coupled with those values determined from differential scanning calorimetry (DSC) thermograms. The apparent $T_o$ and $T_m$ values from Figure 7 are ∼66 and ∼76 °C at pH 4.5, ∼62 and ∼73 °C at pH 3.6, ∼59 and ∼71 °C at pH 3.1, and ∼45 and ∼53 °C at pH 2.2, respectively. The $T_m$ values mostly agree with the transition midpoint temperatures obtained from $I(q)$ for $q$ values of less than ∼0.2 Å$^{-1}$ in Figures 5 and 6. Thus, the collapse and regeneration of the HEWL tertiary structure proceed reversibly for all pH values under these conditions. The pH dependence of $T_o$ and $T_m$ suggests that the intramolecular interactions stabilizing the tertiary structure are weakened with a reduced pH, which would result in a loss of a cooperativity among the hierarchical levels in the unfolding−refolding transition as mentioned above.

*Polypeptide Chain Characteristic Depending on Temperature.* By using the Kratky plots $[q^2I(q)$ vs $q]$, we can see the temperature-dependent characteristics of the polypeptide chain as a polymer. The Kratky plots, often used in polymer science, are known to reflect interactions between polymer elements, local conformations, and a certain rigidity within polymer chains (*50*−*52*). Figure 8 shows the Kratky plots at pH 2.2 and 4.5 in heating and cooling processes, where labels A and B are for pH 2.2 and 4.5 and labels a and b are for heating and cooling, respectively. We can also easily recognize the high reversibility of the structural transition under these experimental conditions. The Kratky plot at $q < \sim 0.2$ Å$^{-1}$ indicates the collapse and regeneration of the tertiary structure and the retention of some compact tertiary structure at even the highest temperature (*45*, *46*). The profile of the Kratky plot at $q > \sim 0.2$ Å$^{-1}$ is altered reversibly from an oscillating curve to an asymptotic slope in heating and cooling, which is quite different from the Kratky plot of proteins in a highly denatured state (*45*, *46*, *49*). This indicates that in the thermal unfolding and refolding transitions the polypeptide conformation changes from a persistent chain with persistence of direction and curvature to a persistent chain with only persistence of direction (so-called Kratky−Porod chain) (*52*). Above ∼1 Å$^{-1}$ there still exists a broad hump as shown the inset in Figure 8, suggesting that a short-range interaction between structural elements holds in this $q$ range at even the highest temperature. It should be noted that an asymptotic slope continues up to ∼1 Å$^{-1}$. This suggests that the Kratky plot analysis for polypeptide chains in proteins would be available up to ∼1 Å$^{-1}$. Polypeptide chains in proteins are typical heteropolymers containing various amino acid residues that have quite different physicochemical characteristics and average radii ranging from ∼2.5 to ∼3.8 Å (*53*). Due to this fact and to

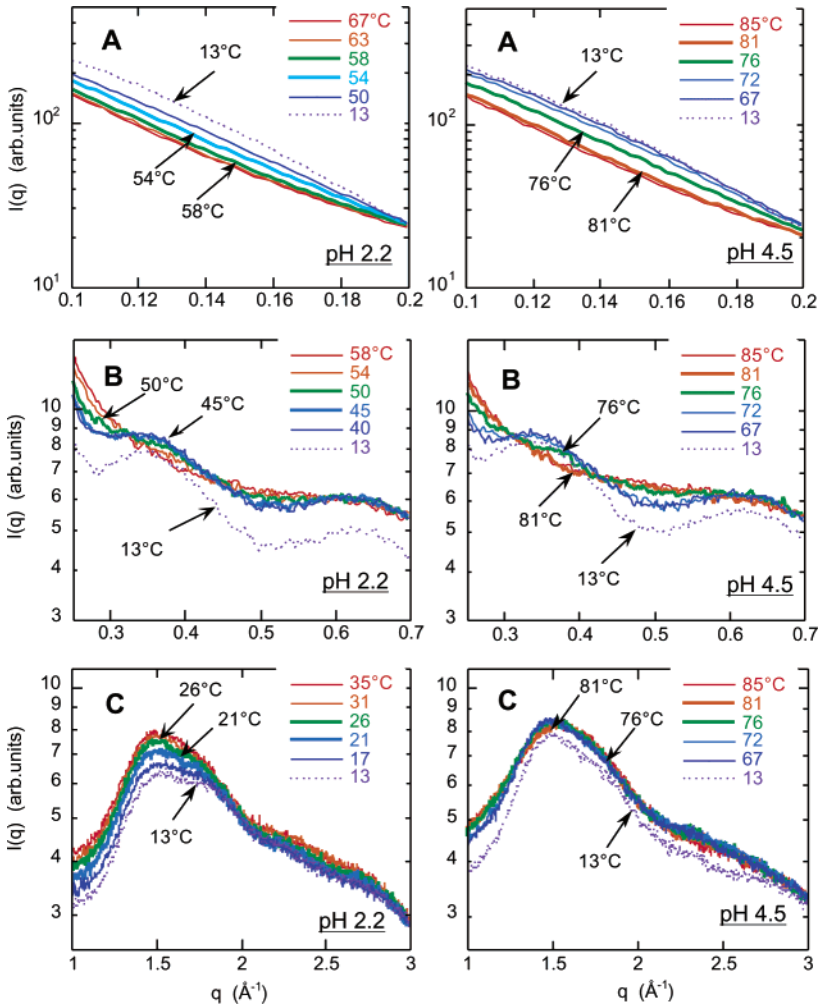


FIGURE 6: Selected WAXS curves $I(q)$ at pH 2.2 and 4.5 around the transition midpoint in the cooling (refolding) process. The thick lines show the maximum temperature differential values [$\Delta I(q)/\Delta T$]. Labels A−C correspond to the scattering curves in the different $q$ ranges as in Figure 5.

the solvation effect, there is no general equation of WAXS representing polypeptide chains as perturbed chains. Therefore, we must await a general formulation treating actual polypeptide chains on a theoretical basis. However, we can tentatively say that the plot of ~0.2− 1 Å$^{-1}$ would reflect the changes in interactions between the local structural elements which are larger than polypeptide units.

*Molar Fractions of Native-like Structure at Different Hierarchical Levels.* By using eq 4, we estimated the molar fractions α of native-like HEWL structure at an intermediate temperature for different $q$ ranges. Figure 9 shows the temperature dependence of α values for all pH values in the heating (unfolding) process. In Figure 9, labels A−D correspond to the α values estimated from the different $q$ ranges: 0.1 Å$^{-1}$ < $q$ < ~0.2 Å$^{-1}$, 0.25 Å$^{-1}$ < $q$ < 0.8 Å$^{-1}$, 1.2 Å$^{-1}$ < $q$ < 1.9 Å$^{-1}$, and 1.35 Å$^{-1}$ < $q$ < 1.45 Å$^{-1}$, respectively. The former three regions were selected because they are assumed to be representative regions of the different hierarchical levels, that is, the tertiary structure, the inter- and intradomain structures (α- and $\beta$-domains), and the secondary structures in two domains, including the closely packed side chains, as discussed in the legends of Figures 1 and 2. The fourth region was selected because it is assumed to be sensitive to a change in the $\beta$-domain, as discussed in Figure 2.

As shown previously, the molar fractions determined from the scattering curves in the $q$ range of 0.1−0.2 Å$^{-1}$ agree well with those determined from DSC thermograms (*18*). The transition midpoint temperatures $T_m$ of the α values in this $q$ range, determined by the temperatures when α = 0.5 in Figure 9A, are ~77 °C at pH 4.5, ~74 °C at pH 3.6, ~71 °C at pH 3.1, and ~53 °C at pH 2.2. These values are quite close to the $T_m$ values determined from the $R_g$ values in Figure 7, and mostly agree with those determined from Figure 7. The onset temperatures $T_o$, determined from the initial temperatures with a large decrease in α, are 70 °C at pH 4.5, 65 °C at pH 3.6, ~60 °C at pH 3.1, and ~43 °C at pH 2.2, respectively. These values mostly agree with the $T_o$ values determined by sight in Figure 7. At pH 2.2 in Figure 9A, we can see that before the main transition a minor transition already starts from ~25 °C and continues to ~43 °C. It suggests the existence of a pretransition accompanying a minor change in the tertiary structure. Thus, the α values in this $q$ range are assumed to reflect well the transition features in the tertiary structure level. In Figure 9B, the $T_m$ values in the $q$ range of 0.25−0.8 Å$^{-1}$, mainly corresponding to the domain regions, are also well-defined as ~75 °C at pH 4.5, ~72 °C at pH 3.6, ~68 °C at pH 3.1, and ~49 °C at pH 2.2. The $T_m$ values in Figure 9C, corresponding to both the secondary structures and the closely packed side

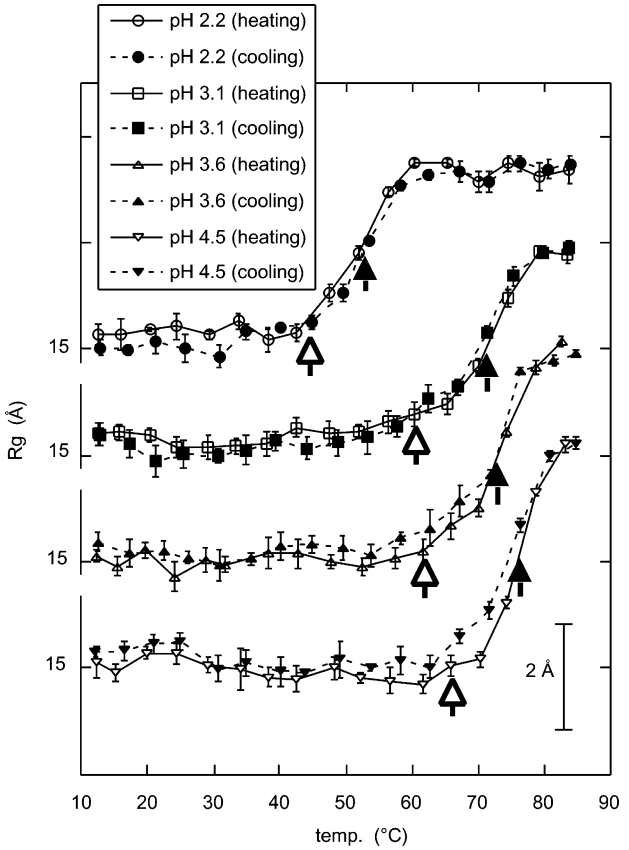


FIGURE 7: Temperature dependence of the radii of gyration $R_g$ at different pH values in heating (unfolding) and cooling (refolding) processes. Empty symbols with solid lines correspond to heating and filled marks with broken lines to cooling: (○ and ●) pH 2.2, (□ and ■) pH 3.1, (△ and ▲) pH 3.6, and (▽ and ▼) pH 4.5. The empty and filled arrows indicate the apparent temperatures of the onset and transition midpoint in $R_g$, respectively.

chains, are estimated to be ~75 °C at pH 4.5, ~71 °C at pH 3.6, ~50 °C at pH 3.1, and ~40 °C at pH 2.2. Thus, the reduction of the pH lowers the transition midpoint temperatures for all hierarchical levels. At pH 4.5 and 3.6, the $T_m$ values corresponding to the domain and secondary structure levels are mostly compatible with each other, and are slightly lower by ~2−3 °C than those of the tertiary structure. It indicates that the thermal stability is primarily the same for all hierarchical structure levels and that the structural transition proceeds concurrently with high cooperativity among them, especially between the domain and secondary structure levels. On the other hand, the $T_m$ values at pH 3.1 and 2.2 in Figure 9C descend much more and significantly deviate from those in panels A and B. It suggests that the structural transition cooperativity among the different hierarchical levels becomes weaker, especially for the cooperativity between the domain and secondary structure levels.

As explained in Figure 2, for HEWL at the native state the scattering curve in the $q$ region of 1.2−1.9 Å$^{-1}$ is mostly given by the simple summation of the scattering curves of α- and β-domains. For this reason, it might be possible to compare the thermal stabilities of the two domains to a certain extent. Thus, the α values, estimated from the $q$ range of 1.2−1.9 Å$^{-1}$ in Figure 9C, would reflect the total fraction of the native-like structures at the secondary structure level which include α-helices (29%), β-sheets (16%), and irregulars (55%) (*54*). At pH 3.1 and 2.2, the α values in this $q$ range begin to decrease gradually from lower temperature in comparison with those at pH 3.6 and 4.5. On the other hand, the α values in Figure 9D were estimated from the $q$ range of 1.35−1.45 Å$^{-1}$. As shown in Figure 2, this $q$ range includes in particular the characteristic maximum of the β-domain at ~1.4 Å$^{-1}$. Thus, the α values in Figure 9D are assumed to be more sensitive to a change in the secondary structure of the β-domain. Although the estimated α values have large errors and evidently include the contributions from both domains, the α values seem to decrease from a much lower temperature at all pH values in comparison to the cases in Figure 9C. It might suggest that the thermal structural transition of the β-domain proceeds prior to that of the α-domain, especially at low pH. It should be noted again that the α values in panels C and D essentially reflect the native-like molar fractions of all structures at the secondary structure level, not only α-helices and β-sheets but also irregulars, including the closely packed side chains in the whole HEWL structure.

*Hierarchical Unfolding and Refolding Maps under Thermal Equilibrium*. Figure 10 summarizes the α values depending on both temperature and $q$ range in heating and cooling processes at pH 2.2, 3.1, 3.6, and 4.5, where we display the α values with rainbow colors and contour lines. In other words, Figure 10 represents the temperature dependence of the molar fraction of native-like structure at different hierarchical levels in the unfolding−refolding transition, which we call structural hierarchy (SH) maps from now on. As shown in Figure 9, the contour line at $\alpha = 0.5$ would mostly correspond to the transition midpoint temperature for each hierarchical level. Via a comparison of the SH map at heating with that at cooling, the reversibility between the unfolding and refolding transitions at each pH mostly holds over all the hierarchical levels. In addition, the SH maps clearly show that the pH variation alters the unfolding−refolding features significantly, such as the transition midpoint temperatures for different hierarchical levels.

From the SH map representation, we are able to gain further information about the structural transitions of proteins. As shown in Figures 1 and 2, the scattering curves in the $q$ ranges of ~0.8−1.1 and >1.9 Å$^{-1}$ are less sensitive to the characteristics of the internal structures of proteins. In the case presented here, the contour lines in these $q$ ranges tend to be broadened with a reduced pH; however, the α values take medium values in the wide temperature range at all pH values and are mostly independent of both temperature and pH. On the other hand, the SH maps in the $q$ ranges of <~0.2, ~0.25−0.8 Å$^{-1}$, and ~1.2−1.9 Å$^{-1}$ significantly depend on hierarchical level, pH, and temperature. The transition midpoint temperatures for different hierarchical levels can be obtained from the intersections of the contour lines of $\alpha = 0.5$ in different $q$ ranges against the temperature axis. As shown in Figure 9A, both the transition midpoint and onset temperatures of the tertiary structure transition are in good agreement with those values determined by the $R_g$ values in Figure 7. The contour lines at $q < \sim 0.2$ Å$^{-1}$ become closer with an increase in pH, indicating that the tertiary structural transition becomes sharper and more cooperative in a case such as a two-state transition. In the same way, we can estimate the transition midpoint temperatures for the respective $q$ ranges, namely, for the respective hierarchical levels.

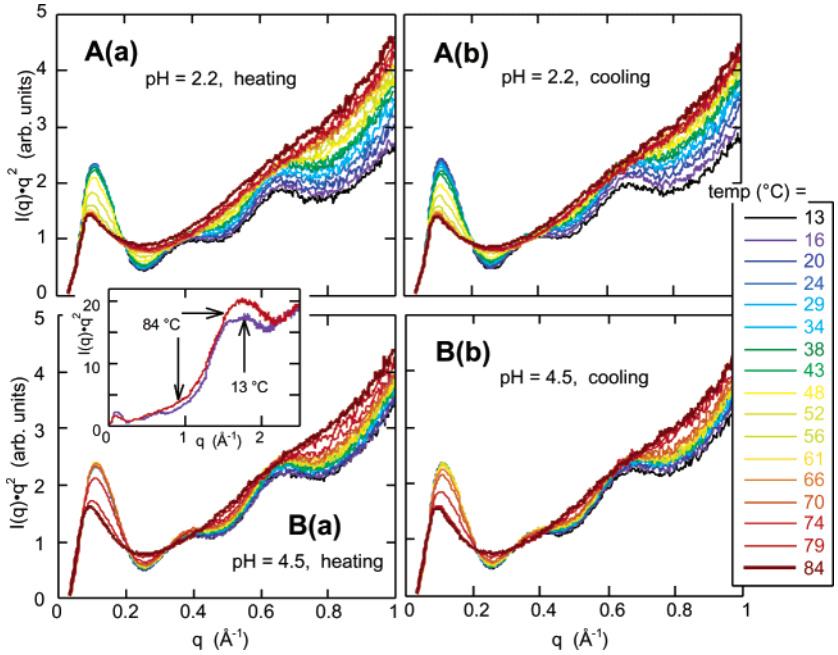


FIGURE 8: Kratky plots at pH 2.2 and 4.5 in heating and cooling processes, where labels A and B correspond to pH 2.2 and 4.5 and labels a and b to heating and cooling, respectively. The inset shows the plots up to $q = 2.5$ Å$^{-1}$ in panel B(a).

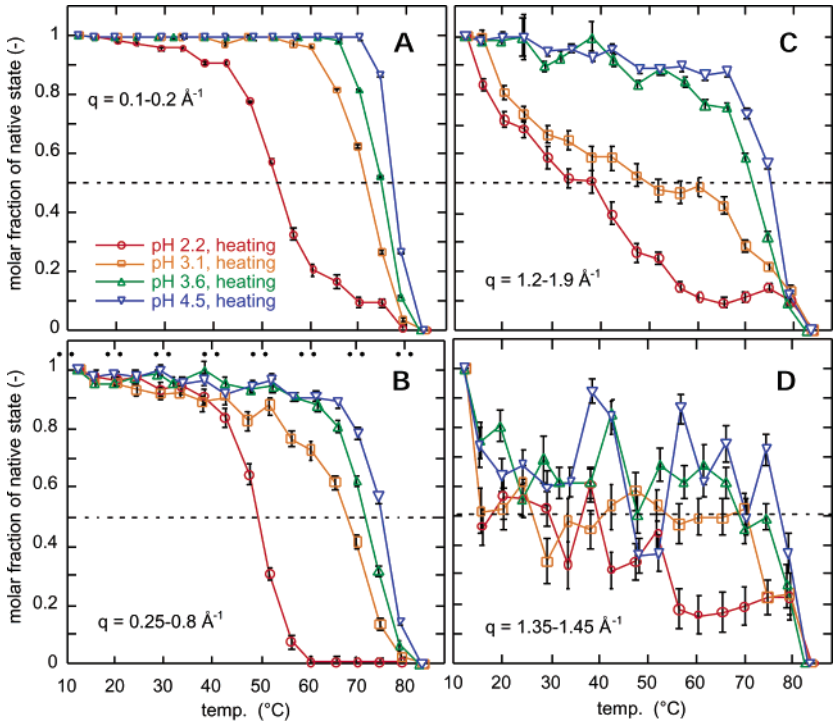


FIGURE 9: Molar fractions α (α ≤ 1) of native-like HEWL structures at intermediate temperatures in the heating process, which were determined from the WAXS curves in different $q$ ranges by using eq 4. Labels A−D correspond to the α values estimated from the different $q$ ranges; 0.1 Å$^{-1}$ < $q$ < ∼0.2 Å$^{-1}$, 0.25 Å$^{-1}$ < $q$ < 0.8 Å$^{-1}$, 1.2 Å$^{-1}$ < $q$ < 1.9 Å$^{-1}$, and 1.35 Å$^{-1}$ < $q$ < 1.45 Å$^{-1}$, respectively. The dashed line indicates α = 0.5.

The extent to which transition cooperativity or concurrency exists among the hierarchical levels would be also appreciable from the SH maps. At pH 4.5, the contour lines below $q \sim 0.8$ Å$^{-1}$ and in the $q$ range of ∼1.3−1.7 Å$^{-1}$ are densely packed around the same transition midpoint temperature (α = 0.5) and have relatively small slopes and flat profiles against $q$ axis. As also shown in Figure 9A−C, this indicates that at pH 4.5 the thermal stability for different hierarchical levels is mostly of the same magnitude, suggesting that the unfolding−refolding transitions for all hierarchical levels proceed rather cooperatively or concurrently. Such transition cooperativity over all the hierarchical levels tends to be gradually missing with a decrease in pH. At pH 2.2, the contour lines for the different $q$ ranges deviate significantly, suggesting that the thermal stability differs for each hierarchical level and that the transition over the whole protein structure proceeds gradually in a case such as a multistate transition (*18*, *20*).

As shown in Figures 5 and 6, in the unfolding−folding transition the WAXS profile in the $q$ range of ∼1.2−2.0 Å$^{-1}$

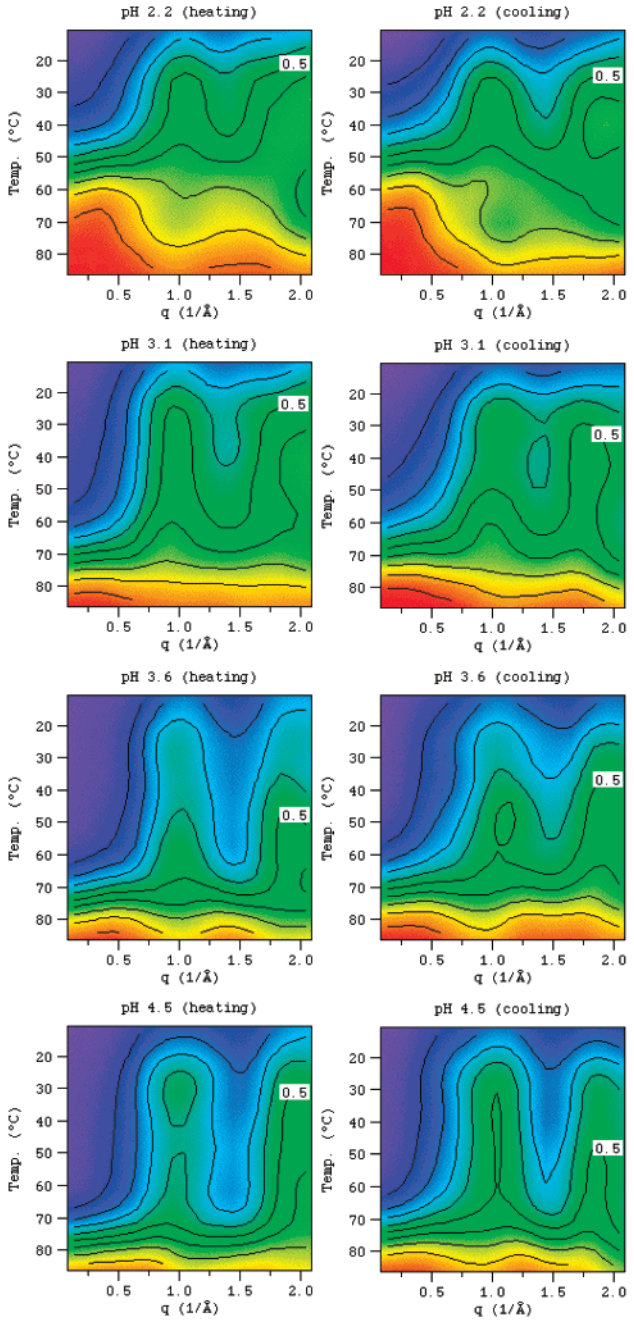


FIGURE 10: Two-dimensional hierarchical maps of unfolding−refolding transitions of the HEWL structure against temperature and $q$ value, termed SH maps, which are obtained from the molar fractions α of native-like structures for different $q$ ranges ($q$ intervals of 0.05 Å$^{-1}$) in the WAXS curves in Figure 4. The α values are displayed with rainbow colors and contour lines. The $\alpha = 0.5$ contour line mostly corresponds to the transition midpoint for each hierarchical structure.

changes most evidently around $q \sim 1.4$ Å$^{-1}$, which occurs very gradually from lower temperatures at pH 2.2. In the SH maps at high pH values, the contour lines have deep concave regions around $q \sim 1.4$ Å$^{-1}$, which become shallower with a decrease in pH. With an increase in temperature, the maximum around $q \sim 1.4$ Å$^{-1}$ changes significantly at low pH values, suggesting the collapse of the $\beta$-domain prior to that of the α-domain. Alternatively, two domains have *intrinsically* different thermal stabilities, which tend to be enhanced at low pH values. This understanding would be consistent with the previous result of the presence of two domains which differ significantly in folding and unfolding kinetics (*11*−*14*, *55*). The recent studies using CD, $^1$H NMR, and mass spectroscopies indicate that the α-domain of HEWL is an independent folding domain at equilibrium (*15*, *16*), which also agrees with the discrepancy in thermal stability between α- and $\beta$-domains seen in the SH maps. From early studies (*54*, *56*), CD and Fourier transform infrared (FTIR) are recognized well as powerful tools for determining secondary structures of proteins. They have also been used in many studies of lysozyme folding and unfolding as well as other spectrometric methods (*57*). FTIR studies of the thermal structural transition of HEWL show evidently denaturation of the $\beta$-sheet even under the different solvent conditions (*56*, *58*). A study of equilibrium unfolding of HEWL, under the presence of 0−8 M urea at pH 2.7, explained consistently the transitions in both tertiary and secondary structures (*59*). In this study, the authors measured SAXS curves ranging from ∼0.02 to 0.25 Å$^{-1}$ and CD spectra, and they defined the existence of unfolding intermediates by using the singular-value decomposition method. They concluded that the initial unfolding in a urea solution would be correlated with a breakdown of intersubunit interactions between the α-domain and the $\beta$-domain due to the unfolding of the $\beta$-domain. Despite the quite different denaturation condition, the unfolding scheme would resemble the results presented here. However, at this stage, we cannot quantify separately the contents of native-like secondary structures in the α- and $\beta$-domains. This is because we have not observed WAXS curves in the fully unfolded state, compared with the cases of CD and FTIR that are now well established for determining helix and sheet contents (*1*).

## DISCUSSION

In many SAXS studies of protein folding, radii of gyration $R_g$ obtained from high-statistic scattering data at small-angle regions are occasionally used to characterize overall structures of proteins in native, unfolded, and intermediate states. However, according to the definition of $R_g$ (*31*), even for a monodispersion system, the $R_g$ value of a solute particle depends not only on the particle shape and size but also on the internal scattering density distribution of the particle. Actually, in some cases, the structural changes in solute particles occurred simultaneously in both the shape and the internal scattering density distribution (*60*, *61*). For example, when a structural change in a protein occurs in both the size and the internal scattering density distribution such as packing geometry of secondary structures, we cannot distinguish between them. Therefore, for understanding a detailed feature of the structural transition of a protein, it is essentially important to obtain high-statistical scattering data with as wide an angle as possible. In this report, we have observed the thermal unfolding−refolding transition of HEWL by using the high-resolution and high-statistic WAXS measurements at the third-generation SR source. The $q$ range measured was from ∼0.05 to ∼3 Å$^{-1}$. This range corresponds to the spatial distance $d$, defined by $d = 2\pi/q$, from ∼125 to ∼2 Å, which covers all hierarchical structures of HEWL from the tertiary structure to the secondary one. In the full $q$ range, the experimental WAXS curve of HEWL in the native state agrees well with the theoretical one using the CRYSOL program and the atomic coordinates from the

PDB. This would ensure the reasonability of the WAXS data analysis presented here.

The WAXS data observed have enabled us to characterize the hierarchical features of the thermal unfolding−refolding transition of HEWL, whereupon we have presented a new map, termed the SH map. This map can help us to understand the unfolding−refolding transition of a protein depending on various structural hierarchical levels. In this case, the SH map is represented by the contour lines of the temperature dependence of the molar fractions $\alpha$ ($\alpha < 1$) of the native-like structure as a function of $q$ range, namely, as a function of hierarchical level or resolution. The contour lines of $\alpha = 0.5$ in the SH maps for different $q$ ranges correspond to the transition midpoints for different hierarchical levels. The density and slope of the contour lines depict the thermal stability and transition feature for each hierarchical level, which would relate to the transition cooperativity among different hierarchical levels. At all pH values from 4.5 to 2.2, the SH maps of HEWL in the unfolding (heating) and refolding (cooling) processes show high reversibility for each hierarchical level. At pH 4.5, the contour lines for all hierarchical levels are crowed and parallel around the same transition midpoint temperature, suggesting that the collapse and regeneration of the native structure proceed rather concurrently or cooperatively for all hierarchical levels in such a two-state transition. For all hierarchical levels, the transition midpoint temperatures $T_m$ at pH 4.5 are ~76 °C, which agree well with those determined by the $R_g$ values and the previous DSC measurements (*17*, *18*, *20*). Such a cooperativity among the different hierarchical levels tends to be gradually lost with a decrease in pH from 4.5 to 2.2. The SH map at pH 2.2 shows that the transition midpoints for each hierarchical level greatly differ from one another, suggesting that the transition proceeds in a multistate fashion.

A WAXS pattern $I(q)$ of a particle in solutions is given by the spherical average in reciprocal space ($q$ space) of the square of the Fourier transform of the scattering density distribution. Then, in general, a simple assignment of specific regions of $I(q)$ to a particular distance is dangerous, unless one is dealing with diffraction peaks. However, for known proteins, we can discuss qualitatively the extent to which a particular level of the structural hierarchy dominates scattering in a given range of $q$ values, which has major implications for the SH maps. In addition, previous thermodynamic considerations showed that a collapse or a regeneration of a protein tertiary structure induces a significant change in the water-accessible area of the protein surface between the native and denatured states with an accompanying large heat capacity change (*62*). As shown previously (*18*), the $\alpha$ value estimated below $q \sim 0.2$ Å$^{-1}$ agrees well with the molar fraction of protein in the native state obtained by DSC. Therefore, the tertiary structure transition seen in the SH map below $q \sim 0.2$ Å$^{-1}$ would correspond to a significant change in the Gibbs free energy. The intramolecular transitions seen in the SH map above $q \sim 0.2$ Å$^{-1}$ would correspond to minor changes in the Gibbs free energy. Alternatively, we assume that SH maps for various $q$ ranges might be related to a hierarchical fine structure of a "folding funnel" or "energy landscape" of a protein (*1*−*3*). Nevertheless, the exact meaning of the SH maps will then be left somewhat fuzzy and will await sharpening by additional works. Although there is such an inherent limitation in the structural determination using the solution scattering method, WAXS data of proteins using a high-intensity third-generation SR source would be sufficiently powerful to clarify a detailed transition feature over all the hierarchical levels at equilibrium.

Finally, we should make a brief comment about another theoretical and experimental basis for understanding SH maps. As is well-known, the scattering curve of a solute particle, including proteins, is given by the summation of three basic scattering curves which depend on the particle shape, the fluctuation of the scattering density of the particle around its mean value, and those cross terms (*31*). The basic scattering curves of a protein for X-ray were separated theoretically and compared with the experimental ones (*63*, *64*). They showed that above $q \sim 0.3$ Å$^{-1}$ the scattering intensity from the intramolecular scattering density fluctuation is several times higher than those from the shape and the cross term. Below $q \sim 0.2$ Å$^{-1}$, the scattering intensity of the basic scattering curve of the molecular shape becomes dominant in the scattering curve. The basic scattering curves of HEWL for neutron were obtained experimentally (*65*), which also shows that the scattering curves below $q \sim 0.2$ Å$^{-1}$ and above $q \sim 0.3$ Å$^{-1}$ mostly reflect the molecular shape and the intramolecular structure, respectively. This is also the case for many other globular proteins, and the theoretical separation of the simulated wide-angle scattering curves of ~60 proteins into those basic scattering curves would support this fact (*66*). We should mention again that the SH map analysis using WAXS would be available for elucidation of hierarchical transition features for known proteins whose WAXS curves can be characterized to some extent based on the crystallographic data. Recently, an advanced approach to classifying unknown protein structures using experimental WAXS data and the structure classification database has been proposed (*67*), which would also suggest another prominent potentiality of high-statistic and high-resolution SR-WAXS studies of proteins in solutions.

## ACKNOWLEDGMENT

This work was performed under the approval of the Program Advisory Committee of the Japan Synchrotron Radiation Research Institute (Proposals 2001A0001-NL and 2001B0455-NL).

## REFERENCES

1. Pain, R. H., Ed. (2000) *Mechanisms of Protein Folding*, Oxford University Press, New York.
2. Dill, K. A., and Chan, H. S. (1997) From Levinthal to pathways to funnels, *Nat. Struct. Biol. 4*, 10−19.
3. Dinner, A. R., Sali, A., Smith, L. J., Dobson, C. M., and Karplus, M. (2000) Understanding protein folding via free-energy surfaces from theory and experiment, *Trends Biochem. Sci. 25*, 331−339.
4. Pollack, L., Tate, M. W., Finnefrock, A. C., Kalidas, C., Trotter, S., Darnton, N. C., Lurio, L., Austin, R. H., Batt, C. A., Gruner, S. M., and Mochrie, S. G. J. (2001) Time resolved collapse of a folding protein observed with small-angle X-ray scattering, *Phys. Rev. Lett. 86*, 4962−4965.
5. 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.
6. Tanford, C. (1968) Protein denaturation, *Adv. Protein Chem. 23*, 121−282.
7. Tanford, C. (1970) Protein denaturation, *Adv. Protein Chem. 24*, 1−95.
8. Pfeil, W., and Privalov, P. L. (1976) Thermodynamic investigation of proteins. I. Standard functions for proteins with lysozyme as an example, *Biophys. Chem. 4*, 23−32.

9. Pfeil, W., and Privalov, P. L. (1976) Thermodynamic investigation of proteins. III. Thermodynamic description of lysozyme, *Biophys. Chem. 4*, 41−50.
10. Privalov, P. L. (1989) Thermodynamic problems of protein structure, *Annu. Rev. Biophys. Chem. 18*, 47−69.
11. Miranker, A., Radford, S. E., Karplus, M., and Dobson, C. M. (1991) Demonstration by NMR of folding domains in lysozyme, *Nature 349*, 633−636.
12. Radford, S. E., Dobson, C. M., and Evans, P. A. (1992) The folding of hen lysozyme involves partially structured intermediates and multiple pathways, *Nature 358*, 302−307.
13. Buck, M., Radford, S. E., and Dobson, C. M. (1993) A partially folded state of hen egg-white lysozyme in trifluoroethanol: Structural characterization and implications for protein folding, *Biochemistry 32*, 669−678.
14. Matagne, A., Chung, E., Ball, L. J., Radford, S. E., Robinson, C. V., and Dobson, C. M. (1998) The origin of the α-domain intermediate in the folding of hen lysozyme, *J. Mol. Biol. 277*, 997−1005.
15. Tachibana, H., Oka, T., and Akasaka, K. (2001) Native-like tertiary structure formation in the α-domain of hen lysozyme two-disulfide variant, *J. Mol. Biol. 314*, 311−320.
16. Bai, P., and Peng, Z. (2001) Cooperative folding of the isolated α-helical domain of hen egg-white lysozyme, *J. Mol. Biol. 314*, 321−329.
17. Hirai, M., Arai, S., Iwase, H., and Takizawa, T. (1998) Small-angle X-ray scattering and calorimetric studies of thermal conformational change of lysozyme depending on pH, *J. Phys. Chem. B 102*, 1308−1313.
18. Hirai, M., Arai, S., and Iwase, H. (1999) Complementary analysis of thermal transition multiplicity of hen egg-white lysozyme at low pH using X-ray scattering and scanning calorimetry, *J. Phys. Chem. B 103*, 549−556.
19. Hirai, M., Arai, S., and Iwase, H. (2000) Concentration dependence of thermal structural transition of hen egg-white lysozyme under constant heating rate studied by time-resolved SAXS, *Thermochim. Acta 344*, 95−102.
20. Arai, S., and Hirai, M. (1999) Reversibility and hierarchy of thermal transition of hen egg-white lysozyme studied by small-angle X-ray scattering, *Biophys. J. 76*, 2192−2197.
21. Arai, S., and Hirai, M. (1998) Synchrotron radiation small-angle X-ray scattering study of structural transition and gelation of hen egg-white lysozyme depending on pH and ionic strength under isothermal heating, *Kobunshi Ronbunshu 55*, 653−660.
22. Miura, K., Kawamoto, M., Inoue, K., Yamamoto, M., Kumasaka, T., Sugiura, M., Yamano, A., and Moriyama, H. (2000) Commissioning for wide-angle routine proteomix beam line BL40B2: protein crystallography and small-angle scattering, *SPring-8 User Experiment Report 4*, 168.
23. Huang, T. C., Toraya, H., Blanton, T. N., and Wu, Y. (1993) X-ray powder diffraction analysis of silver behenate, a possible low-angle diffraction standard, *J. Appl. Crystallogr. 26*, 180−184.
24. Bosio, L., Teixeira, J., and Bellissentfunel, M. C. (1989) Enhanced density-fluctuations in water analyzed by neutron-scattering, *Phys. Rev. A 39*, 6612−6613.
25. Gekko, K., and Hasegawa, Y. (1986) Compressibility-structure relationship of globular proteins, *25*, 6563−6571.
26. Glatter, O. (1982) in *Small-Angle X-ray Scattering* (Glatter, O., and Kratky, O., Eds.) pp 119−196, Academic Press, London.
27. Feigin, L. A., and Svergun, D. I. (1987) in *Structure Analysis by Small-Angle X-ray and Neutron Scattering* (Taylor, G. W., Ed.) pp 68−73, Plenum Press, New York.
28. Privalov, P. L., and Khechinashvili, N. N. (1974) A thermodynamic approach to the problem of stabilization of globular protein structure: A calorimetric study, *J. Mol. Biol. 86*, 665−684.
29. Filimonov, V. V., Potekhin, S. A., Matveev, S. V., and Privalov, P. L. (1982) Thermodynamic analysis of scanning microcalorimetric data, *Mol. Biol. 16*, 435−444.
30. Kidokoro, S., and Wada, A. (1987) Determination of thermodynamic functions from scanning calorimetry data, *Biopolymers 26*, 213−229.
31. Stuhrmann, H. B., and Miller, A. (1978) Small-angle scattering of biological structures, *J. Appl. Crystallogr. 11*, 325−345.
32. Debye, P. (1915) Zerstreuung von röntgenstrahlen, *Ann. Phys. 46*, 809−823.
33. Svergun, D. I., Barberato, C., and Koch, M. H. J. (1995) CRYSOL: a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates, *J. Appl. Crystallogr. 28*, 768−773.
34. Svergun, D. I., Richard, S., Koch, M. H. J., Sayers, S., Kuprin, S., and Zaccai, G. (1998) Protein hydration in solution: Experimental observation by X-ray and neutron scattering, *Proc. Natl. Acad. Sci. U.S.A. 95*, 2267−2272.
35. Iwase, H., Hirai, M., Arai, S., Mitsuya, S., Furusaka, M., Otomo, T., and Shimizu, S. (1999) Comparison of DMSO-induced denaturation of hen egg-white lysozyme and bovine α-lactalbumin, *J. Phys. Chem. Solids 60*, 1379−1381.
36. Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) SCOP: A structural classification of proteins database for the investigation of sequences and structures, *J. Mol. Biol. 247*, 536−540.
37. Perutz, M. F., Fermi, G., Poyart, C., Pagnier, J., and Kister, J. (1993) A novel allosteric mechanism in haemoglobin. Structure of bovine deoxyhaemoglobin, absence of specific chloride-binding sites and origin of the chloride-linked Bohr effect in bovine and human haemoglobin, *J. Mol. Biol. 233*, 536−545.
38. Maurus, R., Overall, C. M., Bogumil, R., Luo, Y., Mauk, A. G., Smith, M., and Brayer, G. D. (1997) A myoglobin variant with a polar substitution in a conserved hydrophobic cluster in the heme binding pocket, *Biochim. Biophys. Acta 1341*, 1−13.
39. Tsukada, H., and Blow, D. M. (1985) Structure of α-chymotrypsin refined at 1.68 Å resolution, *J. Mol. Biol. 184*, 703−711.
40. Hynes, T. R., and Fox, R. O. (1991) The crystal structure of staphylococcal nuclease refined at 1.7 Å resolution, *Proteins 10*, 92−105.
41. Hennig, M., Sterner, R., Kirschner, K., and Jansonius, J. N. (1997) Crystal structure at 2.0 Å resolution of phosphoribosyl anthranilate isomerase from the hyperthermophile *Thermotoga maritima*: possible determinants of protein stability, *Biochemistry 36*, 6009−6016.
42. Hoover, D. M., and Ludwig, M. L. (1997) A flavodoxin that is required for enzyme activation: the structure of oxidized flavodoxin from *Escherichia coli* at 1.8 Å resolution, *Protein Sci. 6*, 2525−2537.
43. Diamond, R. (1974) Real-space refinement of the structure of hen egg-white lysozyme, *J. Mol. Biol. 82*, 371−391.
44. Wlodawer, A., Svensson, L. A., Sjolin, L., and Gilliland, G. L. (1988) Structure of phosphate-free ribonuclease A refined at 1.26 Å, *Biochemistry 27*, 2705−2717.
45. Kataoka, M., Nishi, 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.
46. Kataoka, M., and Goto, Y. (1996) X-ray solution scattering studies of protein folding, *Folding Des. 1*, R107−R112.
47. 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.
48. Kuwajima, K. (1989) The molten globule state as a clue for understanding of the folding and cooperativity of globular-protein structure, *Proteins: Struct., Funct., Genet. 6*, 87−103.
49. Kataoka, M., Hagihara, Y., Mihara, K., and Goto, Y. (1993) Molten globule of cytochrome *c* studied by small-angle X-ray scattering, *J. Mol. Biol. 229*, 591−596.
50. Kratky, O., and Porod, G. (1949) Röntgenuntersuchung gelöster fadenmoleküle, *Recl. Trav. Chim. Pays-Bas 68*, 1106−1122.
51. Kriste, R. G. (1967) Neue Vorstellungen über statistische fadenknäuel, *Makromol. Chem. 101*, 91−103.
52. Kriste, R. G., and Oberthür, R. C. (1982) Synthetic polymers in solution, in *Small-Angle X-ray Scattering* (Glatter, O., and Kratky, O., Eds.) pp 387−431, Academic Press, London.
53. Zamyatnin, A. A. (1972) Protein volume in solution, *Rep. Prog. Phys. 24*, 108−123.
54. Chen, Y.-H., Yang, J. T., and Martinez, H. M. (1972) Determination of the secondary structures of proteins by circular dichroism and optical rotatory dispersion, *Biochemistry 11*, 4120−4131.
55. Radford, S. E., Buch, M., Topping, K. D., Dobson, C. M., and Evans, P. A. (1992) Hydrogen exchange in native and denatured states of hen egg-white lysozyme, *Proteins: Struct., Funct., Genet. 14*, 237−248.
56. Anderle, G., and Mendelsohn, R. (1987) Thermal denaturation of globular proteins: Fourier transform-infrared studies of the amide III spectral region, *Biophys. J. 52*, 69−74.
57. Matagne, A., and Dobson, C. M. (1998) The folding process of hen lysozyme: a perspective from the 'new view', *CMLS, Cell. Mol. Life Sci. 54*, 363−371.
58. van Stokkum, I. H. I., Linsdell, H., Hadden, J. M., Haris, P. I., Chapman, D., and Bloemendal, M. (1995) Temperature-induced

changes in protein structures studied by Fourier transform infrared spectroscopy and global analysis, *Biochemistry 34*, 10508−10518.
59. Chen, L., Hodgson, K. O., and Doniach, S. (1996) A lysozyme folding intermediate revealed by solution X-ray scattering, *J. Mol. Biol. 261*, 658−671.
60. Hirai, M., Takizawa, T., Yabuki, S., Nakata, Y., and Hayashi, K. (1996) Thermotropic phase behavior and stability of monosialoganglioside micelles in aqueous solution, *Biophys. J. 70*, 1761−1768.
61. Hayakawa, T., and Hirai, M. (2002) Hydration and thermal reversibility of glycolipids depending on sugar-chains, *Eur. Biophys. J. 31*, 62−72.
62. Privalov, P. L., and Makhatadze, G. I. (1990) Heat capacity of proteins. II. Partial molar heat capacity of the unfolded polypeptide chain of proteins: Protein unfolding effects, *J. Mol. Biol. 213*, 385−391.
63. Stuhrmann, H. B. (1973) Comparison of the three basic scattering functions of myoglobin in solution with those from the known structure in crystalline state, *J. Mol. Biol. 77*, 363−369.
64. Ibel, K., and Stuhrmann, H. B. (1975) Comparison of neutron and X-ray scattering of dilute myoglobin solutions, *J. Mol. Biol. 93*, 255−265.
65. Stuhrmann, H. B., and Fuess, H. (1976) A neutron small-angle scattering study of hen egg-white lysozyme, *Acta Crystallogr. A32*, 67−74.
66. Iwase, H., and Hirai, M. (2001) Simulations of neutron solvent contrast variation of proteins and intramolecular structures, *J. Phys. Soc. Jpn. 70*, 420−422.
67. Sokolova, A. V., Volkov, V. V., and Svergun, D. I. (2003) Prototype of database for rapid protein classification based on solution scattering data, *J. Appl. Crystallogr. 36*, 865−868.
68. Kraulis, P. J. (1991) MolScript: A program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr. 24*, 946−950.

BI0499664