Refolding of lysozyme by quasistatic and direct dilution reaction paths: A first-order-like state transition

Chia-Ching Chang,[1,*] Xu-Cheng Yeh,[1,2] Hui-Ting Lee,[1] Po-Yen Lin,[1,2] and Lou-Sing Kan[2,†]
[1]*Department of Physics, National Dong Hwa University, Hualien, Taiwan 97401*
[2]*Institute of Chemistry, Academia Sinica, Nankang, Taipei, Taiwan 11529*
(Received 6 November 2003; published 7 July 2004)

A first-order-like state transition model is considered to be a global reaction mechanism to directly folded proteins from an unfolded state to its native form. In order to verify the general applicability of this mechanism, we used lysozyme as a model protein. It was fully unfolded by 4.5 M urea, 0.1 M dithiothreitol (DTT) in *p*H 3 and refolded to its native form by way of an overcritical reaction path (a quasistatic process) or directly crossing transition boundary path (a directly dilution process). In addition to the two states coexisting in the direct folding path, lyzosyme might be trapped in a glassy state. However, it can escape from the glassy state by concentration twice. This indicates the existence of a state transition line or boundary in the direct folding reaction. However, lysozyme can continuously fold from unfolded to native by an overcritical reaction path. During the overcritical path, four stable folding intermediates and native lysozyme were obtained. The secondary structures, particle size distributions, thermal stabilities, and oxidation state of disulfide bonds of folding intermediates were analyzed by circular dichroism spectra, dynamic light scattering, differential scanning calorimetry, and Raman spectra, respectively. According to the data, the intermediates of both the overcritical reaction and the direct crossing transition boundary paths can be described by a common concept pertaining to a model that undergoes collapse, sequential, and first-order-like state transition. This indicated that protein folding by way of different reaction paths might follow a similar folding mechanism—i.e., a mechanism of overcritical folding of intermediates. A protein folding reaction diagram is postulated and discussed. In spite of a global interaction mechanism the α-helix is formed prior to the β-sheet, which may indicate that protein folding is initiated by local interactions.

DOI: 10.1103/PhysRevE.70.011904 PACS number(s): 87.14.Ee, 87.15.Cc, 87.15.By, 87.15.He

I. INTRODUCTION

A protein can exist in two distinct states, a folded (native) state and an unfolded state, analogous to the two phases of a material. The two states may also coexist, just as many materials have coexisting phases. In between the two states, a protein may possess many intermediate states [1,2]. In order to distinguish from the commonly accepted "first-order phase transition" for general materials, we adopt the language "first-order-like state transition" when describing the protein folding-unfolding state changes and their models [1]. A first-order phase-state transition is a transition that has a limited "state transition line." A phase-state change can occur by crossing the state transition line. It can also occur by routing through an undefined "overcritical region"—i.e., reaching another state but keeping away from defined state transition line and critical point. This path is called an "overcritical reaction path" [2]. The aim of this paper is to study the mechanism of protein folding and the applicability of a global model that we previously proposed [1–3] by extending it to and verifying it by yet another common protein, lysozyme, in an acidic environment.

Protein folding may follow a spontaneous process [4] or a reaction-path-directed process (an overcritical process) [5], determined by the various folding transition boundaries intrinsic to the specific protein. When the transition line or region of a protein is so limited that the unfolded protein will become folded to its native form without going through the transitional states, then this process is called a spontaneous folding. We previously proposed that a general first-order-like state transition model could encompass both processes without any conflicts [1]. In this general model part of the protein aggregates by crossing a state transition line or boundary directly. However, if the transition boundary is finite and limited, we may design an overcritical folding path steered away from such boundary and fold the protein without trapping it in an aggregated state [1–3,6]. Therefore, the folding path may not be unique. We also found that some of the intermediates, following an overcritical path, have characteristics consistent with those of Ptitsyn's postulates of a "molten globule state" [1–3,6,7], that they possess secondary structures, have melting temperatures when heated, and are structurally different from the native and unfolded proteins. Due to the lack of long-range order, some proteins do not show distinct melting behavior when heated as is in the case of glasses. They are, therefore, denoted as in a glassy state. The definition of a "molten globule state" differs somewhat in various studies [7–9]. We follow that of Ptitsyn and use his term "highly ordered molten globule" [7] to describe a "nativelike fold" which contains both tertiary and secondary structures. The folding of the intermediates follows that of a sequential model in which the secondary structure is formed prior to the tertiary [10]. However, by comparing the particle size distributions of the intermediates, the unfolded and the

*Corresponding author. FAX: 886-3-863-3690. Electronic address: chiaching@mail.ndhu.edu.tw
†Corresponding author. FAX: 886-2-2788-4184. Electronic address: lskan@chem.sinica.edu.tw

native proteins, one may also see that the protein shrinks rapidly in the beginning of a folding process, consistent with a collapse model [11].

The first-order-like state transition theory was successfully applied to the refolding of growth hormone bearing α-helix and metallothionein, a metal binding protein in an alkaline environment [1–3,6]. In order to expand the generality of the above theory, we now examine the refolding of lysozyme that contains both α-helix and β-sheet [12–15] in acidic conditions.

In this study, lysozyme was unfolded completely by urea (4.5 M) in an acidic buffer. Two pairs of disulfide bonds of cysteines were reduced into SH form by DTT (0.1 *M*). When folding via an overcritical reaction path, the cysteines were gradually oxidized by air and the concentration of urea was eventually removed by a three-step dialysis. The lysozyme was refolded into its native form after the urea was removed completely and the *p*H was changed to 5. All the intermediates involved in the above path are in a rigid and highly ordered molten globule state [7]. On the other hand, when the lysozyme is folded directly across the transition boundary, it was trapped into a mixed glassy state containing reduced cysteine (SH) and oxidized cystine (SS). This intermediate therein can escape from the glassy state and change to the molten globule state by concentrating the lysozyme with a semipermeable membrane device Microcon (see experimental section) with a molecular weight cutoff (MWCO) of 10000 g/mol. According to our observations, we concluded that lysozyme folding is consistent with a first-order-like state transition model [1]. Although the intermediates from different reaction paths may present different characteristics, they are consistent with both collapse and sequential models. This indicates that the intermediates involved in either path may follow the same first-order-like state transition folding mechanism.

As for local interactions, we observed that α-helix is formed prior to the β sheet in the overcritical refolding intermediates, consistent with the zipper model proposed by Schellman [16].

II. MATERIAL AND METHODS

A. Materials and buffers

Hen egg white lysozyme was purchased from Sigma Ltd. (St. Louis, MO). All other chemicals were obtained from Merck Ltd. (Rahway, NJ). The denaturing-unfolding buffer (*p*H 3) contained 4.5 M urea with 10 mM phosphate buffer, 0.1 M dithiothreitol (DTT), 0.1% mannitol, and 0.5 mM Pefabloc. The urea will break the hydrogen bonds of protein and relax the protein from packed conformation. The lower *p*H helps stabilizing relaxed conformation of protein. DTT was used for reducing the oxidative disulfide bond (S-S) of Cys residues into the SH group in high concentration. The reduction can be examined by Raman spectroscopy. In low concentration, DTT will protect proteins from unwanted oxidation, similar to the role of mannitol. The Pefabloc is a proteinase inhibitor to protect proteins from unexpected degradation. There were five folding buffers employed in this study. Their compositions are summarized in Table I. Thus, the folding buffers dialyzed away the urea and mannitol in the denaturing-unfolding buffer. In the meantime, the *p*H value was raised from an acidic to a neutral range. The intermediates M_1, M_2, M_3, M_4, and M_5 denoted are stable intermediates in the solvent environments expressed in Table I.

TABLE I. Chemical compositions of refolding buffers.

	Phosphate buffer (mM)	*p*H	Urea (M)	DTT (mM)	Mannitol (%)	Pefabloc. (μM)
Folding buffer 1	10	3	2	0.1	0.1	0.5
Folding buffer 2	10	3	1	0.1	0.1	0.5
Folding buffer 3	10	3		0.1	0.1	0.5
Folding buffer 4	10	5		0.1	0.1	0.5
Folding buffer 5 (the native buffer)	10	5		0.1		0.5

B. Folding of lysozyme by direct dilution and quasistatic thermal equilibrium dialysis

Unfolded lysozyme (*U*) was obtained by treating natural lysozyme with denaturing-unfolding buffer to make it 10 mg/ml in concentration. This solution was left at room temperature for 1 h and then was centrifuged at 4000 g for 30 min to remove the undissolved residue. The folding processes were accomplished by two methods. One was a direct dilution of unfolded lysozyme into 100-fold native buffer, mixed well at ambient temperature. The concentration of urea, in this mixed solution, was 45 mM and the *p*H value was 5 after this direct dilution process. Lysozyme, which was trapped in a glassy state (*I*), was concentrated two fold by a semipermeable membrane device Microcon (Millipore Corp., Billerica, MA) and resulted in a molten globule state (I'). The average pore size of semipermeable membrane is quite small and protein and/or DNA can be separated by its molecular size and shape. The excess buffer can be removed by this device. Therefore, the concentration of protein can be increased. The intermediate I' can be folded further into native form (*N*) by dialysis against a native buffer.

The second folding process is described as a quasistatic procedure that involves five consecutive thermal equilibrium dialysis (TED) steps [1–3]. Each of the folding intermediates (M_{1-5}) was dialyzed against a particular folding buffer at 4 °C, as shown in Table I. In order to avoid possible deterioration of the buffer during the dialysis, all refolding buffers were replaced by fresh ones midway through each dialysis.

Step 1. The unfolded lysozyme (*U*) in the denature-unfolding buffer was dialyzed against the folding buffer 1 for 72 h to dilute the urea concentration from 4.5 to 2 M (this produces intermediate 1 or M_1).

Step 2. M_2 was obtained by dialyzing M_1 against folding buffer 2 for 24 h to dilute the urea concentration to 1 M.

Step 3. M_3, an intermediate without denaturant (urea) in solution, was then obtained by dialyzing M_2 against folding buffer 3 for 24 h.

Step 4. M_3 was further dialysed against folding buffer 4 for 24 h, and the *p*H changed from 3 to 5 to produce M_4.

Step 5. Finally, the chemical chaperon mannitol was removed by dialyzing M_4 against the native buffer for 8 h to yield M_5.

C. Analysis of the secondary structure of lysozyme by CD

CD spectra in the UV region (260–200 nm) of U, M_{1-5}, and N were recorded on a Jasco J 720C spectropolarimeter at 20 °C. A 0.1 cm-light-path cuvette was used to reduce the light scattering of the solution. Data are expressed in molar ellipticity $[\theta]$ (deg cm^2 dmol^{-1}), calculated based on the molecular weight of 14.4 KD for lysozyme. Unlike mean residue molar ellipticity $[\theta]$ (deg cm^2 dmol^{-1} residue^{-1}) [17], the scale of molar ellipticity $[\theta]$ of lysozyme is around 10^5-10^6. The CD spectrum was measured for each individual sample diluted with the corresponding folding buffer at each stage. The solvent contribution on CD was subtracted from each spectrum. The component secondary structures—i.e., α-helix, distorted α-helix, β-sheet, distorted β-sheet, turns, and unordered—were analyzed by the singular value decomposition algorithm provided by the SELCON3 program [18] which selected lysozyme as one of the proteins of its training set of this program.

D. Analysis of size distribution of U, N, and M_{1-5} by dynamic light scattering

The light scattering apparatus was equipped with an argon ion laser light source (4880 Å) (Ion Laser Tech., UT). The scattered light was collected at 90° by a goniometer from Brookhaven Instruments Corp (BIC, Holtsville, NY). The chamber temperature was controlled by a water circulator. The autocorrelation function (ACF) was computed using a digital correlator (BI9000), then analyzed with the nonnegatively constrained least-squares (NNLS) method [19,20]. The instrument performance was calibrated by measuring the dynamic light scattering (DLS) of standard suspensions of polystyrene beads (R_H=68 nm) (Polysciences, Warrington, PA). The concentration of lysozyme in each dialyzing step was 2 mg/ml, and all experiments were conducted at 20 °C.

E. Thermal stability monitored by differential scanning calorimetry

Thermal stability experiments by differential scanning calorimetry (DSC) were performed with a TA Q10 differential scanning calorimeter (New Castle, DE) interfaced to a computer equipped with an automatic data collection program. Variant folding intermediates of the lysozyme were calibrated to a concentration of 10 mg/ml in their refolding state buffer. All samples were filtered and degassed under vacuum for 10–15 min with gentle stirring prior to being loaded into the calorimetric cell. DSC experiments were performed under a constant external pressure of 1 bar in order to avoid bubble formation, and samples were heated at a constant scan rate of 10 °C/min. Unlike other conventional differential scanning calorimeters [17], the signal of this DSC (TA Q10) is relatively stable in the scan rate of 10 °C/min and the signal of this experiment is as good as previously studied. Experimental data were corrected for buffer contribution prior to data analysis. After concentration normalization, a chemical base line calculated from the progress of the unfolding transition was subtracted. The excess heat capacity was calibrated with a sapphire reference. The excess heat capacity functions were then analyzed using a program that came with the instrument. The excess heat capacity was calculated by the equation

$$\frac{\Delta Q_{sample} - \Delta Q_{baseline}}{\Delta Q_{reference} - \Delta Q_{baseline}} = \frac{C_{p_{sample}}}{C_{p_{reference}}},$$

where ΔQ's denote the heat absorbed by the sample, reference, and background. In this study, we use sapphire as the reference and blank as the background. Meanwhile, we also checked the heat capacity of the refolding solvent. Therefore, the excess C_p of folding intermediates can be determined.

F. Reduction-oxidation state of sulfur group of cys monitored by a micro-Raman spectrometer

Raman experiments were carried out with the Ranishow system 1000 micro-Raman spectrometer (Ranishow Plc, Gloucestershire, UK). Five μL lysozyme solutions of varying folding states (10 mg/ml) were lyophilized onto the surface of a platinum plate. The Raman spectra were collected around 2570 and 510 cm^{-1}. They represent the stretching frequencies of the S-H bond ($\nu_{S\text{-}H}$) of the reduced cysteine and the S-S bond ($\nu_{S\text{-}S}$) of the oxidized cystine of lysozyme, respectively [21]. The wave numbers of the Raman spectra were calibrated with silicon oxide. Each spectrum was compared with the same folding state buffer to identify the signal of protein and solvent, respectively.

III. RESULTS

A. Refolding of lysozyme by an overcritical path

1. Monitored by CD

Figure 1 shows the CD spectra of various folding states of lysozyme via the overcritical path. In the unfolding buffer, the CD of U showed no meaningful spectral features in wavelengths shorter than 230 nm (Fig. 1). A secondary structural analysis by SELCON3 [16] indicated that it contained a helical-like conformation (Fig. 2). However, their optical characteristics were restored after a 72 -h TED process, during which the urea concentration was changed gradually from 4.5 M to 2 M in the first folding step (M_1). Its secondary structures can be analyzed as also shown in Fig. 2. It contains 57% of α-helix, 22% distorted α-helix, 6% turn, and 15% unordered form. No β-sheet was detected. This composition is similar to the intermediate of the second folding step (M_2). β-sheet conformation occurred right after the urea was removed by TED process (M_3). M_3 contains 4% of β-sheet and 7% distorted β-sheet, but the other conformation has been redistributed (21 % of α-helix, 13% distorted α-helix, 24% turn, and 31% unordered form). The above composition is similar to that of the native lysozyme (Fig. 2).

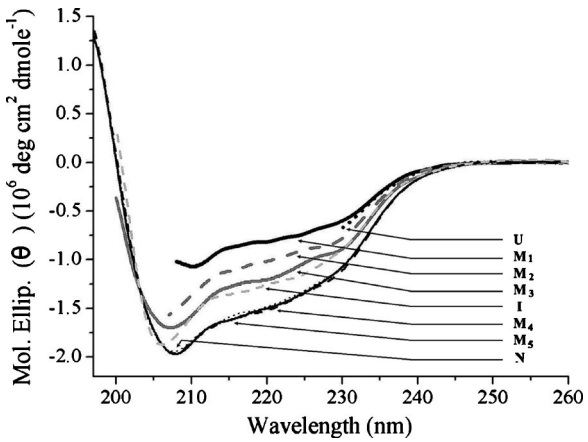

FIG. 1. Circular dichroism profiles of lysozyme at 20 °C. For the sake of clarity, the CD curves of U, M_1, M_2, M_3, M_4, M_5, N, and I are annotated in the figure. U, ■ ■ ■ ■; M_1, ■■■ M_2, ■ ■; M_3, ■■■; M_4, ■ ■ ; M_5, ——; N, ·····; I, ▬ ▬ . Due to the high absorption background of urea in UV range the signal-to-noise ratio is very low when the wavelength is lower than 230 nm of U, and is therefore not shown.

The molar ellipticities of M_4, M_5 and native lysozyme (N) are almost the same (Fig. 1). The composition of the secondary structure is also similar to that of M_3.

According to the above CD results (spectral characteristics and thermal stability), we found that the helical structure of lysozyme formed at 2 M of urea concentration. The β-sheet formed in a urea-free environment, and a preliminary tertiary structure of lysozyme was about to be formed when the *p*H was raised to 5 (see below).

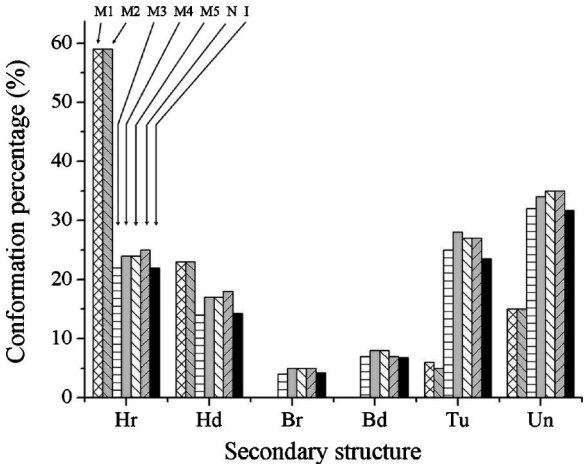

FIG. 2. The CD profiles analysis by SELCON3 of overcritical reaction path and direct folding intermediates of lysozyme. The secondary structure Hr: regular α-helix, Hd: distorted α-helix, Br: regular β-strand, Bd: distorted β-strand, Tu: turns, and Un: unordered. The % of folding intermediates distributed in each secondary structure are represented by histograms: U, mosaic with white background; M_1, slash from upper left to lower right with gray background; M_2, horizontal line with white background; M_3, filled in solid gray; M_4, slash from upper left to lower right with white background; M_5, slash from lower left to upper right with gray background; and I, filled in solid back.

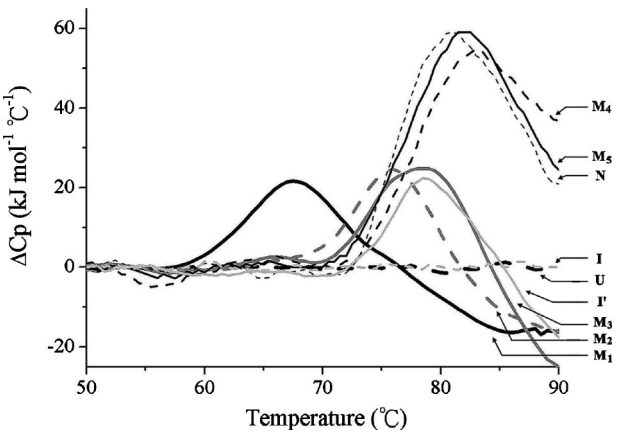

FIG. 3. Phase transition profiles of variant folding intermediates by DSC. The convention of each profile is the same as in Fig. 1. In addition, the gray fine line is represented intermediates of direct folding, I'. I'▬. Where the intermediate I' is the intermediate I which was concentrated twofold.

2. Monitored by DSC

Figure 3 shows the phase transition profiles of the various folding intermediates. A DSC profile of lysozyme in the unfolded state (U) showed no significant phase transition point during the temperature scanning. However, M_{1-5} and N had transition temperatures of 67.3, 75.9, 78.4, 83.0, 82.0, and 81.0 °C, respectively. The similarity in the transition temperatures of M_4,M_5, and N indicated that the conformations of M_4 and M_5 are restored to the native form. These observations are consistent with our recently NMR study (data not shown). We have repeated these data four times and the Tm accuracy was within 1 °C.

3. DLS studies on lysozyme intermediates (U,N,M_{1-5})

In solution, dynamic light scattering is one of the methods for measuring particle size distribution (PSD) in the submicron range [17]. Similar to our previous studies [3,5,6], the size distributions of all folding intermediates are uniform in a single distribution as measured by DLS. The uniformity and single distribution of the diameters indicated that the TED process has reached an equilibrium state and the state of the protein is in a monomerous state [1]. The hydrodynamic diameter (R_H) of U shrank from 8.55 to 5.00 nm (Table II) right after the first refolding stage. This is consistent with our previous studies [1–3,6] as well as the protein collapse model [11]. Although the effective diameter of the hen egg white lysozyme obtained from x-ray crystallography, around 4.32 nm [15], is smaller than those of all the intermediates and native lysozyme in solution, a similar study by NMR and DLS [22] also indicated that native lysozyme may have similar hydrodynamic diameters. This indicated that the conformation of lysozyme in solution is more flexible than its highly packed crystal [15]. The relaxed diameter of U was only 1.7-fold larger than the native protein. This might indicate that the unfolded lysozyme was not fully extended. It is also consistent with the CD data and previous studies [1,3,6].

TABLE II. Hydrodynamic diameter of each folding intermediate.

Folding state	Hydrodynamic diameter (nm)
U	8.55+/−0.21%
I	5.00+/−0.94%
M_1	4.85+/−0.52%
M_2	4.98+/−0.19%
M_3	4.97+/−0.59%
M_4	4.99+/−0.25%
M_5	4.99+/−0.39%
N	5.00+/−0.32%

4. Raman spectra studies of disulfide bond formation of lysozyme

The oxidation state of the cystein residue of lysozyme was studied by the Raman spectral method. The stretching vibrational frequency is around 2570 cm^{-1} for the reduced SH bond and is around 510 cm^{-1} for the oxidized form SS [21]. The unfolded lysozyme (U) showed a significant absorption at 2574.1 cm^{-1}, and an absence of absorption near 510 cm^{-1} was observed. On the contrary, at the first folding state of the overcritical folding path, no absorption around 2570 cm^{-1} was found, whereas a broad absorption at 507 cm^{-1} was observed. Similar patterns were observed for M_{2-5} and N (Fig. 4). All of these indicated that all the cysteines were in a reduced state (SH type) at the unfolded stage and that the all disulfide bonds (SS type) formed right at the first folding stage. There were no mixed SH and SS states of lysozyme found in these folding stages. The Raman shift around 540 cm^{-1} and 565 cm^{-1} can be observed only in the buffers containing high concentrations of urea ($\geqslant$1 M) and this absorption can also be observed in a buffer-only system. Therefore, this absorption is independent of the protein. Based upon the study of Frost *et al.* [23] both signals around 540 cm^{-1} and 565 cm^{-1} might be due to the weak absorption of the NCO and NCN bending of urea. Therefore, they can only be observed under high concentrations.

B. Refolding of lysozyme by direct dilution

Figure 1 also shows the CD profiles of the intermediates (U, I, and N') obtained by way of direct dilution. The CD profiles N' are identical to M_5 (of the overcritical reaction path). Therefore, its CD curve did not show in Fig. 1. However, the CD intensity of the intermediate (I) is similar to that of M_3 (21 % of α-helix, 14% distorted α-helix, 4% β-sheet, 7% distorted β-sheet, 23% turn, and 31% unordered form, Figs. 1 and 2). Thus, the secondary structure of intermediate

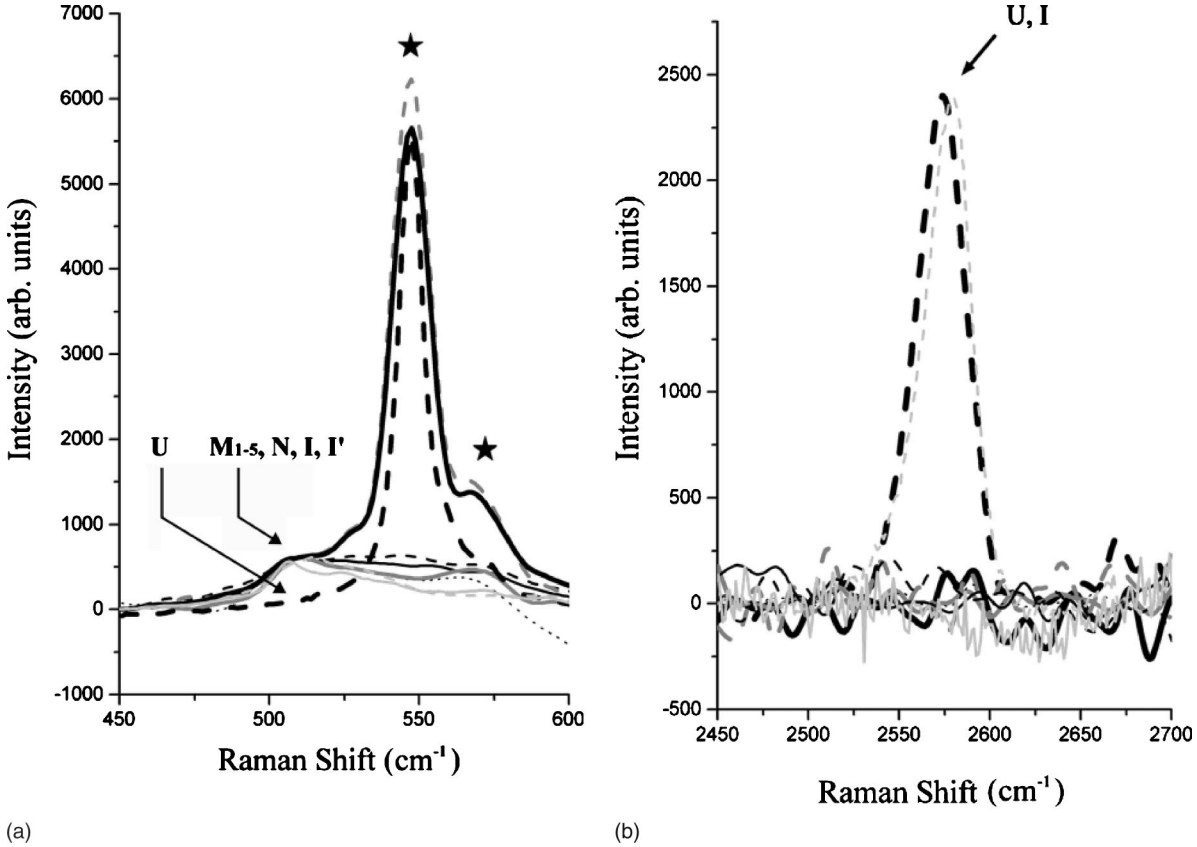

FIG. 4. Raman shift of lysozyme folding intermediates. (a) The S-S ($\nu_{S\text{-}S}$) stretching region. (b) The S-H ($\nu_{S\text{-}H}$) stretching region. The arrow indicates the S-S stretching frequency. Stars denote the NCN and NCO bending signals around 546 and 560 cm^{-1} of high-concentration urea, respectively. All profiles, excluding the denature lysozyme, in (a) were normalized with their intensity and shifted at 510 cm^{-1}, magnified to normal count. "Arb. units" means arbitrary unit of intensity.

I is similar to the native structure. Likewise, intermediate I' escaped from I and contained an identical CD profile (data not shown). A DSC study indicated that the intermediate I contained no phase transition point. However, by concentrating with Microcon, the intermediate I' contained a significant phase transition point at 78.5 °C, which is similar to M_3 (Fig. 3). A DLS study indicated that the hydrodynamic diameter of I/I' is 5.00 nm $(+/-0.94\%)$, very close to its size in the native form (Table II). Raman spectra studies indicated that intermediate I contained both reduced ν_{SH} and ν_{SS} signals around 2570 and 510 cm^{-1}. These indicated that I is in a mixed state. However, I' contained only ν_{SS} signal around 510 cm^{-1}.

IV. DISCUSSION

A. Lysozyme folding is consistent with both collapse and sequential framework models

The Raman spectra of unfolded and various folding intermediates indicated that all the disulfide bonds of unfolded intermediate were in reduced states (Fig. 4). A DSC study also indicated that the unfolded state has no structure (Fig. 3). However, according to the DLS study, the unfolded lysozyme did not extend fully, and the effective gyration diameter was around 1.7-fold of the size of the native structure (Table II). Meanwhile, a CD profile also indicated that the unfolded lysozyme might be in a helixlike conformation. However, the effective gyration diameter of the lysozyme shrank to a nativelike diameter at the first folding stage. Therefore, this conformation may not be a stable structure. This is consistent with the collapse model. Although the concentration, examined by CD (0.2 mg/ml) and DLS (2 mg/ml), was lower than the refolding condition (10 mg/ml), there was no aggregation observed in the high-concentration condition and this was consistent with previous NMR studies [24]. Therefore, the single-particle size distribution and small deviation in these data indicated that these intermediates had reached a monomerous and stationary state. According to previous studies with CD and DSC, all overcritical folding intermediates contained stable structures, and these intermediates (M_{1-4}) were in molten globule states. These intermediates can further fold into the native form via a quasistatic folding process [3–5]. Therefore, the reaction path can be expressed as $U \rightarrow M_1 \rightarrow M_2 \rightarrow M_3 \rightarrow M_4 \rightarrow M_5$ (N). This is also consistent with the sequential framework model [10]. Although intermediate I was trapped in a glassy state, it can transition to M_3-like intermediates by perturbing its solvent environment and it can further fold to a form identical to M_5, indicating that it is the true folding intermediate. Therefore, the overall reaction paths of protein folding may be expressed as $U \rightarrow I \rightarrow M \rightarrow N$ or $U \rightarrow M \rightarrow N$, where M denotes the overcritical reaction intermediates and I denotes the glassy state intermediate.

B. Lysozyme folding is consistent with the first-order-like state transition model

A thermodynamic experiment using DSC revealed the structurally related thermal stability potential. The peak of excess heat capacity change is the "first-order phase transition" [25] reaction. The integrated area under the peak is the energy (enthalpy ΔH) to melt the secondary or tertiary structures of the protein. However, some proteins [26–28] did not represent a "first-order phase transition" but a "second-order phase transition" thermal denaturation behavior of their heat capacities change [29]. Namely, these proteins may have no long-range structural order and they melt continuously as glasses do. Therefore, we named the protein in such a state a glassy state [30].

In denaturant-controlled protein folding processes, two protein folding paths were studied. We found that the protein can be folded from the same initial point and end up with the same native form via different reaction paths. We folded them continuously by a stepwise quasistatic process so that each intermediate during solvent environment exchange could be obtained and analyzed. Evidence of the secondary structure and the tertiary interactions indicates that those intermediates are in between a folded and an unfolded state. Since there is no third state present in a two-state transition, the overcritical state still remains undefined.

However, for a direct folding path, the protein was trapped in a state, which contained secondary structures but with the same Tm of tertiary structures. By analyzing the oxidation states of their disulfide bonds we found that both the reduced state (SH) and the oxidized state (SS) signals were present. This intermediate seemed to contain both the unfolded and folded states.

We can draw some similarities of the two protein folding paths with the gas-liquid phase transition of water. A stepwise refolding path is similar to a path detoured from the critical point, whereas a direct folding path is analogous to the phenomenon of being trapped in a phase transition line with two coexisting phases. Therefore, we use a similar term "first-order-like state transition" to model the reaction of protein folding.

Accordingly, we propose a global reaction state diagram for protein folding processes, as shown in Fig. 5. We use the term "protein folding state" to represent the possible states, whether they are well defined or not. In Fig. 5, $\Phi(n_1, n_2, \ldots) = 1$ indicates a folded state or native state (N) and $\Phi = 0$ indicated an unfolded state. In general, the intensity of CD profiles at 222 nm may represent the relative folding state [17] of those folding intermediates (i.e., M_{1-5}), although this may need further justification by the experimental results of DSC [17] or NMR studies (data not shown). According to the CD profiles of lysozyme (Fig. 1), we can see that M_4, M_5, and N have similar intensities at 222 nm, indicating that the two folding intermediates M_4 and M_5 may very well be identical to N. These data are consistent with ^{1}H-NMR studies of all intermediates and native lysozyme (data not show). The n_1 and n_2 denote the variables of a folding state such as the concentration of urea, *p*H, and temperature of the solution. I', a perturbed folding state of I which has an identical CD profile of I, is very similar to M_3, as examined by both DSC (Fig. 3) and CD (Fig. 1). The only difference between I' and M_3 was the solvent environment: 45 mM urea and *p*H 5 (for diluting unfolded protein into a 100-fold native solution) for the former and no urea and *pH* 3 for the latter. This indicated that the effects of these variables, such as concentration of urea and *p*H, can be com-

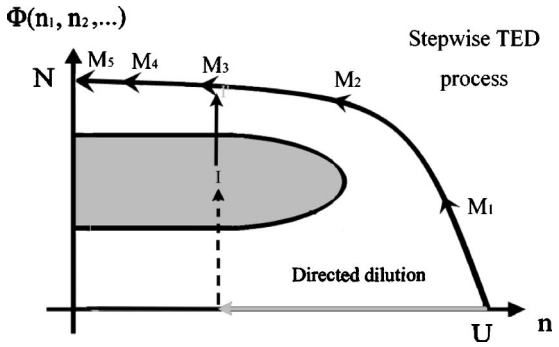

FIG. 5. Proposed first-order-like state transition model of the protein folding reaction diagram. Φ $(n_1,n_2,\ldots)$ denotes the folding state of protein. The parameters of n,n_1, and n_2 denote the folding state variables such as temperature, concentration of denaturants, etc.

bined and are interchangeable. Therefore, the variables $n_1,n_2\ldots$ may be degenerated into an effective variable n on the proposed folding diagram as indicated in Fig. 5. From the results presented in this paper together with our previous studies [1–3,6], the "first-order-like state transition" model is well suited for a protein going through multiple aggregation or a glassy state that may have a finite boundary on a protein folding reaction diagram.

C. All reaction folding intermediates may follow the same folding mechanism

Although the secondary structure composition and the hydrodynamic diameter of intermediate *I* via direct folding are similar to the native lysozyme, DSC studies did not show any significant phase transition profile during the temperature scanning. This indicated that *I* is in a glassy state [30], and there may be a folding transition barrier in which the protein was trapped. It is consistent with an interrupted refolding kinetic study by Kiefhaber [12]. However, by concentrating this intermediate *I* with Microcon, a stable conformation was observed by DSC (Fig. 3), indicating that the intermediate *I* has escaped from a glassy state into a molten globule state, in which the state of the protein contained a stable structure but different from a native state. These implied that intermediates in any folding path might have similar and corresponding intermediates in another path as discussed previously.

The intermediates we obtained from direct folding or sequential folding are all consistent with the collapse model and framework sequential model where the secondary structure is formed prior to any tertiary interactions. Therefore, we may propose that protein folding in any reaction path should follow the same thermal dynamic rules and same folding mechanism. Namely, the mechanism revealed by stepwise folding intermediates may represent the true mechanism of protein folding.

D. Local interactions initiate protein folding

When a protein reaches a thermal equilibrium or stationary state, it is usually at its lowest potential energy and is relatively more stable than other conformations in similar environments. Therefore, if the folding process mimics a nativelike reaction, the folding intermediates may represent the possible intermediates of the real reaction path. A similar argument has been supported by protein denaturation experiments [31,32]. In their studies, the transient structure intermediates of cytochrome C [31] and apo-myoglobin [32] have been identified as being in a molten globule state [7]. A similar observation is reported here in this study. All reaction intermediates may follow the same folding mechanism. The initiation of protein folding has been modeled after a framework [10] or collapse [11] mechanism. Based on our previous experiments [1–3,6], both the framework and collapse models fit satisfactorily. In this study, we found that the unfolded lysozyme was not fully extended (Table II) but contained some helical-like conformation (Fig. 1). Consequently, a stable and complete helical intermediate was obtained at a later folding stage. After the helix formation, part of the helix was transformed into a sheet structure and then to an unordered structure after the urea was removed at the third refolding state. These indicated that under a restrained environment, the helix formation was relatively easy, and it may well be the initiation step of the protein folding. As indicated by the zipper model of helix formation [16], local interactions play the most important role. Therefore, protein folding may be initiated from a local interaction such as the β-sheet formation we have observed here.

Similarly to our previous studies, the growth hormone and metallothionein could be unfolded and refolded reversibly by our TED process [1–3,6]. Through similar processes, the lysozyme could also be refolded to the native protein, as suggested by our CD, DSC, DLS, and Raman data. It appeared that this model protein could also be folded in a stepwise manner. However, unlike the growth hormone, the lysozyme intermediate of direct folding did not precipitate but was trapped in a glassy state without any stable conformation. We have demonstrated that lysozyme can be folded following a first-order-like state transition process by either overcritical or direct pathways. Thus, the folding path of a protein is not unique; i.e., if a proper path is carefully chosen, a protein can always be folded correctly. Namely, an unfolded protein which refolds back through a carefully designed stepwise quasistatic process may experience a gradual change in its environment and thus will not be trapped in a glassy state or subsequently form an aggregate. In addition to the overall protein folding reaction, the helical-sheet transition between folding intermediates also indicated that a local interaction mechanism (helix formation) might initiate protein folding.

ACKNOWLEDGMENTS

We thank Dr. Eric To for providing the instrument of DLS for protein size determination, Dr. Charlie C. L. Cheng for providing the instrument of the micro-Raman spectrometer, Dr. Ching-Nien Chen of National Institutes of Health, Bethesda, MD for valuable discussion, and W. L. Chiu of Lihyuan Enterprise Co., LTD. for assisting with the DSC measurements. This study was supported in part by Grant Nos. NSC 92-2112-M-259-016 (for C.C.C.) and NSC 92-2113-M-001-030 (for L.S.K.) and Academia Sinica.

[1] C. C. Chang, Y. C. Su, M. S. Cheng, and L. S. Kan, Phys. Rev. E **66**, 021903 (2002).
[2] Y. L. Liu, H. T. Lee, C. C. Chang, and L. S. Kan, Biochem. Biophys. Res. Commun. **306**, 59 (2003).
[3] C. C. Chang, M. S. Cheng, Y. C. Su, and L. S. Kan, J. Biomol. Struct. Dyn. **21**, 247 (2003).
[4] C. B. Anfinsen, Science **181**, 223 (1973).
[5] B. Honig, A. Ray, and C. Levinthal, Proc. Natl. Acad. Sci. U.S.A. **73**, 1974 (1976).
[6] C. C. Chang, C. T. Tsai, and C. Y. Chang, Protein Eng. **5**, 437 (2002).
[7] P. B. Ptitsyn, Adv. Protein Chem. **47**, 83 (1995).
[8] H. S. Chan and K. A. Dill, Annu. Rev. Biophys. Biophys. Chem. **20**, 447 (1991).
[9] A. V. Finkelstein and E. I. Shaknovich, Biopolymers **28**, 1681 (1989).
[10] P. S. Kim and R. L. Baldwin, Annu. Rev. Biochem. **51**, 459 (1982).
[11] K. A. Dill, Biochemistry **24**, 1501 (1985).
[12] T. Kiefhaber, Proc. Natl. Acad. Sci. U.S.A. **92**, 9029 (1995).
[13] P. Roux, M. Delepierre, M. E. Glodberg, and A. F. Chaffotte, J. Biol. Chem. **272**, 2483 (1997).
[14] K. Sasahara, M. Demura, and K. Nitta, Proteins **49**, 472 (2002).
[15] S. Datta, B. K. Biswal, and M. Vijayan, Acta Crystallogr., Sect. D: Biol. Crystallogr. **57**, 1614 (2001).
[16] J. A. Schellman, J. Phys. Chem. **62**, 1485 (1958).
[17] S. Yadav and F. Ahmad, Anal. Biochem. **283**, 207 (2000).
[18] N. Sreerama, S. Y. Venyaminov, and R. W. Woody, Protein Sci. **8**, 370 (1999).
[19] D. F. Nicoli, K. Hasapidis, P. O'Hagan, D. C. McKenzie, J. S. Wu, Y. J. Chang, and B. E. H. Schade, in *Particle Size Distribution III: Assessment and Characterization*, edited by T. Provder (ASC Publications, Columbus, OH, 1998), p. 52.
[20] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems* (Prentice-Hall, Englewood Cliffs, NJ, 1974).
[21] S. K. Freeman, *Applications of Laser Raman Spectroscopy* (Wiley-Intersciences, New York, 1974).
[22] J. Poznanski, Y. Georgalis, L. Wehr, W. Saenger, and P. Zielenkiewicz, Biophys. Chem. **104**, 605 (2003).
[23] R. L. Frost, J. Kristof, L. Rintoul, and J. T. Kloprogge, Spectrochim. Acta, Part A **56**, 1681 (2000).
[24] N. Niccolai, O. Spiga, A. Bernini, M. Scarselli, A. Ciutti, I. Fiaschi, S. Chiellini, H. Molinari, and P. A. Temussi, J. Mol. Biol. **332**, 437 (2003).
[25] P. L. Privalov and N. N. Khechinashvili, J. Mol. Biol. **86**, 665 (1974).
[26] J. D. Bryngelson and P. G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **84**, 7524 (1987).
[27] E. I. Shakhnovich and A. V. Finkelstein, Biopolymers **28**, 1667 (1989).
[28] A. V. Finkelstein and E. I. Shakhnovich, Biopolymers **28**, 1681 (1989).
[29] P. L. Privalov, J. Mol. Biol. **258**, 707 (1996).
[30] K. Kuwajima and F. X. Schmid, Adv. Biophys. **18**, 43 (1984).
[31] T. R. Sosnick, L. Mayne, and S. W. Englander, Proteins **24**, 413 (1996).
[32] S. N. Loh, M. S. Kay, and R. L. Baldwin, Proc. Natl. Acad. Sci. U.S.A. **92**, 5446 (1995).