Kinetic Evidence of Microscopic States in Protein Folding†

Hueih M. Chen, Vladislav S. Markin,‡ and Tian Yow Tsong*

Department of Biochemistry, University of Minnesota College of Biological Sciences, St. Paul, Minnesota 55108

Received November 20, 1991; Revised Manuscript Received October 6, 1992

ABSTRACT: Staphylococcal nuclease unfolds at acidic pHs and refolds at neutral pH. Previous kinetic analysis based on both the direct pH jump and the sequential pH jump, from a native condition (pH 7.0) to pHs beyond unfolding transition zones (pH 3.0 and pH 12), and vice versa, supports the mechanism, $D_3 \rightleftharpoons D_2 \rightleftharpoons D_1 \rightleftharpoons N_0$, in which N_0 is the native state and D's are the three substates of the denatured form [Chen, H. M., You, J. L., Markin, V. S., & Tsong, T. Y. (1990) *J. Mol. Biol. 220*, 771–778; Chen, H. M., Markin, V. S., & Tsong, T. Y. (1992) *Biochemistry 31*, 1483–1491]. Here we show that both the single- and the double-pH jump kinetics of folding and unfolding to the intermediate pHs (3.4–5.0, i.e., in the transition zone), in which both the native and the denatured states coexist, are not compatible with this simple sequential model. At 25 °C, log τ_1^{-1} (for the $D_1 \rightleftharpoons N_0$ step) and log τ_2^{-1} (for the $D_2 \rightleftharpoons D_1$ step) vs pH show a $\sqrt{\ }$-shaped dependence on the final pH, with minimal values (τ_1^{-1} of 0.56 s^{-1} and τ_2^{-1} of around pH 3.9. The third relaxation τ_3 (for the $D_3 \rightleftharpoons D_2$ step, 35 s) was independent of pH in the range 3.4–8.5. The $\sqrt{\ }$-shaped dependence on pH of log τ_1^{-1} and log τ_2^{-1} cannot be reproduced by the above but can be accounted for if each of N_0, D_1, and D_2 is composed of many microscopic states in rapid equilibrium. These microscopic states are designated as two subpopulations, α_i ($i = 0, 1$, and 2), from which unfolding can take place, and β_i ($i = 1$, 2, and 3), from which folding can take place. Analysis of kinetic data indicates that $\alpha_1 = 1$ and $\alpha_2 = 1$ for the whole pH range. However, $\beta_1 = 1$ only for pH > 5.5 and decreases to 0 for pH < 3 (pK_a around 4.5), and $\beta_2 = 1$ only for pH < 2 and decreases to small values for pH > 3 (pK_a around 1.2). Analysis also shows that equilibrium unfolding of the protein is triggered by the absorption of 2.6 ± 0.3 protons, of which 0.8 ± 0.3 occurs in the N_0 to D_1 step and 1.7 ± 0.3 occurs in the D_1 to D_2 step. We conclude that acidic unfolding can take place from the whole population of microscopic states while folding at pH 7 can take place only through those subpopulations which are folding-permitting. pH-dependent equilibrium transitions of all kinetically distinctive intermediates in the folding pathway are also determined and shown to be consistent with the above interpretation.

Many kinetic and equilibrium studies of protein folding have provided evidence for the existence of partially folded intermediate states (Kim & Baldwin, 1990; Jaenicke, 1987). Characterization of these intermediates is an essential step toward understanding mechanisms of protein folding. For staphylococcal nuclease (SNase), several folding intermediates have been detected and characterized. Using the stopped-flow kinetic method, Schechter et al. (1970) and Epstein et al. (1971) have detected a 50-ms reaction and a 550-ms reaction by folding from acidic conditions. They assigned the faster one to the nucleation event and the slower one to the folding reaction. Davis et al. (1979) compared the folding kinetics of two isoenzymes (SNase A and B) and detected an additional reaction at 35 s. Nakano and Fink (1990) studied effects of low temperature and cosolvents on the three kinetic phases in folding. Sugawara et al. (1991) studied folding from the urea-denatured form by circular dichroism stopped-flow kinetics and detected a fast reaction (<15 ms) which acquired nearly 30% of the helical content of the native protein.

The complex nature of the folding/unfolding of SNase has been shown to be due to the heterogeneity of the native and the unfolded states by the NMR spectroscopy (Markley et al., 1970; Fox et al., 1986). Fox et al. (1986) have employed magnetization transfer–NMR to reveal two distinct forms of the native SNase in equilibrium. Subsequently, Evans et al. (1987, 1989) have identified the two isomeric forms of the native state to be due to the cis/trans isomerism at the bond Lys116/Pro117. A mutant in which Pro117 was replaced with Gly lacked a minor form of the native state (in which Pro117 presumably is in the trans configuration) (Evans et al., 1987). These authors have proposed a cyclic kinetic scheme consisting of two isomeric forms of native state (N and N*) and two isomeric forms of denatured states (D and D*) (Evans et al., 1989). Kinetic and equilibrium constants for the cyclic scheme have been derived on the basis of the NMR data. Following these studies, Alexandrescu et al. (1989) have reported that an additional conformational substate can be detected in the native state of SNase. This latter form (N″) appears at higher nuclease concentrations and may correspond to a dimerized enzyme.

Our study of the acid and alkaline induced folding/unfolding, by direct pH jump and sequential pH jump methods (Brandts et al., 1975; Hagerman & Baldwin, 1976; Hagerman et al., 1979), in the time range 2 ms to 300 s, was consistent with the sequential mechanism in Scheme I, in which N_0, D_1, D_2, and D_3 are the native state and the three substates of the denatured state, respectively (Chen et al., 1991, 1992):

Scheme I

$$D_3 \rightleftharpoons D_2 \rightleftharpoons D_1 \rightleftharpoons N_0$$

Free energy of conversion between any two neighboring D substates was less than 0.5 kcal mol^{-1}, although they were separated by considerable kinetic barriers (ca. 7 kcal mol^{-1} between D_1 and D_2 and 13.5 kcal mol^{-1} between D_2 and D_3).

† This work was supported by NIH Grant GM 37304.

‡ V. S. Markin was a visting scholar from Frumkin Institute of Electrochemistry, Academy of Sciences of the USSR, Moscow, USSR 117071.

0006-2960/92/0431-12369$03.00/0

The three D's are isoenergetic, and a unique pathway cannot be determined by the difference in their thermodynamic stability. We concluded that kinetic barriers rather than thermodynamic stabilities of intermediate states would determine the pathway in the early stage of protein folding (Chen et al., 1991, 1992).

Additional evidence supporting Scheme I is the finding that the unfolding kinetics were monophasic while the folding kinetics were triphasic (Chen et al., 1991, 1992; see also Discussion). A native state composed of two or three isomeric forms would have given rise to complex kinetics in unfolding, which was not detected by the Trp140 fluorescence (Chen et al., 1991). Eftink et al. (1989) have reported that Trp140 fluorescence may not be sensitive to the N to N* transition of Evans et al. (1989). In addition, the concentration of nuclease and salts and the temperature in these experiments were different.

Another aspect of our previous experiments is that kinetics were performed at the baseline regions of the native and the pH-denatured states [see Figure 1a of Chen et al. (1991)], and it has been known that any partially folded intermediates in folding would not accumulate under such conditions (Kim & Baldwin, 1990). Furthermore, although in the equilibrium measurements unfolding was triggered by the protonation of 2.6 titratable groups, kinetic experiments detected only the absorption of 0.8 proton in the N_0 to D_1 transition. A total of 1.8 protons remained unaccounted for in the kinetic experiments. These titratable groups all have pK_a values of approximately 3.8.

MATERIALS AND METHODS

Enzyme. Wild-type SNase was purified from *Escherichia coli* strain AR120 which carries the recombinant plasmid pL9 with the nuclease gene (gift of Dr. David Shortle of the Johns Hopkins University) by the procedure of Shortle and Meeker (1989). The purity of the protein was judged to be greater than 98% by SDS–polyacrylamide gel electrophoresis. The concentration of the protein was determined by a Perkin-Elmer Lambda 5 spectrophotometer using the optical density at 280 nm (0.93 at 1 mg mL^{-1}) (Fuchs et al., 1967).

Stopped-Flow Kinetic Measurements. Kinetics were measured with a Hi-Tech Model PQ/SF-53 stopped-flow instrument with a multiple mixing attachment, as described previously (Chen et al., 1991, 1992). This instrument has a nominal mixing time of 1.2 ms for an observation chamber with a 10-mm optical path length. The baseline of the instrument is stable in the time range of a few milliseconds to several hundred seconds (see Figure 1). Thus, reactions which occur in 2 ms to 500 s can be studied with great confidence. The fluorescence of Trp 140 was used to monitor the kinetics by setting the excitation beam at 285 nm and following the emission integrated above 300 nm. Signals were digitized to 12-bit resolution and analyzed with an IBM-compatible PC. As was mentioned previously (Chen et al., 1992), the fluorescence intensity of the solution was linear with the protein concentration only up to 70 μM (1.18 mg mL^{-1}). Beyond this concentration, there was aggregation of protein or other trivial quenching of fluorescence. Thus, all measurements with SNase were done below this concentration. The buffer contained 0.10 M NaCl and 0.05 M NaAc, 0.05 M sodium phosphate, 0.05 M glycine, and 20 μM EDTA. The pH was adjusted by adding HCl or NaOH. In the direct pH jump experiment, a protein solution was mixed with an equal volume of buffer to give a desired final pH. In the double jump experiment, a protein solution at pH 7.0 was mixed with a buffer to a desired pH. The mixture was kept at that pH for a period of time called the delay time (t_D). t_D in the range 10 ms to 10 s was used. The mixture was then mixed with a second buffer to bring the final mixture back to pH 7. Kinetics in both the forward jump and the back jump were analyzed for time constants and amplitudes of different kinetic phases [see Chen et al. (1992)]. With the analysis already presented (Chen et al., 1992), a complete set of kinetic constants for Scheme I can be derived.

Previous data show that there is a slow reaction in the 30-s time range (τ_3) which is at least 10-fold slower than the next reaction. Therefore, a complete reaction was recorded in two separate time ranges, one from 2 ms to 5 s and the other from 5 s to 250 s. Deconvolution was done to each record in the fast and the slow time ranges by a nonlinear least-squares program supplied by Hi-Tech, based on the algorithm of Marquardt (1963). Data were fit into the equation

$$\Delta F(t)/\Delta F_\infty = A_0 + \sum A_i \exp(-\lambda_i t) \qquad (1)$$

where $\Delta F(t)$, ΔF_∞, A_0, A_i, and λ_i are, respectively, fluorescence change up to time t, fluorescence change when t = 250 s, a constant, amplitude of the ith kinetic phase, and time constant of the ith kinetic phase. A fit was considered good if the Chi-square deviation and the number of kinetic phases, i, were small, and the decay time constants were separated by at least 3-fold. If two decay constants had similar values, i was reduced by 1. If the Chi-square deviation did not worsen, the fit was accepted. A few examples of least-squares deconvolution are given in Figure 1. For fitting kinetic data versus pH to a set of kinetic equations (see Figure 2), we used the Mathematica program in a Macintosh computer.

RESULTS

Effects of pH on Folding/Unfolding Kinetics. So far, our kinetic measurements have been done at strongly native (e.g., pH 7) or strongly denaturing (e.g., pH 3) conditions (Chen et al., 1991, 1992). Under these conditions, the folding was triphasic and the unfolding was monophasic. When folding or unfolding was measured inside the folding/unfolding transition zone, i.e., pH 3.5–4.5, folding remained triphasic but so did the unfolding. Because of the inconsistency observed for the unfolding reactions, a rigorous analysis of the kinetic records was done for the unfolding at pH 2.2 and 3.5. In Figure 1A, the kinetic record for the unfolding at pH 3.5 is given in curve A for the fast time range (2 ms to 5 s) and that at pH 2.2 in curve B. Residuals for both curves after the deconvolution are also shown. The perfectly flat residual curves assure that there was no systematic error in the least-squares fits. As explained in the legend, the best fit for curve A gave two time constants, 0.81 s^{-1} and 0.20 s^{-1}, and the best fit for curve B gave one time constant, 6.0 s^{-1}. The same reactions recorded at the slow time range (2 ms to 250 s) are shown in Figure 1B. Curve A (unfolding at pH 3.5) gave an additional time constant, 0.039 s^{-1}, in the time range 5 s to 250 s, but curve B (unfolding at pH 2.2) showed no detectable signals in this slow time range. Thus, unfolding to pH 2.2 remains monophasic, as was reported, and to pH 3.5 it was triphasic.

A complete set of kinetic constants for the folding/unfolding in the pH range 1.5–8.5 is shown in Figure 2. Data in the open squares were obtained by direct pH jump from pH 7.0 to acidic pHs, and they resemble those reported previously for unfolding to strongly denaturing conditions (Chen et al., 1991). Data in the open diamonds (for the fast unfolding phase)

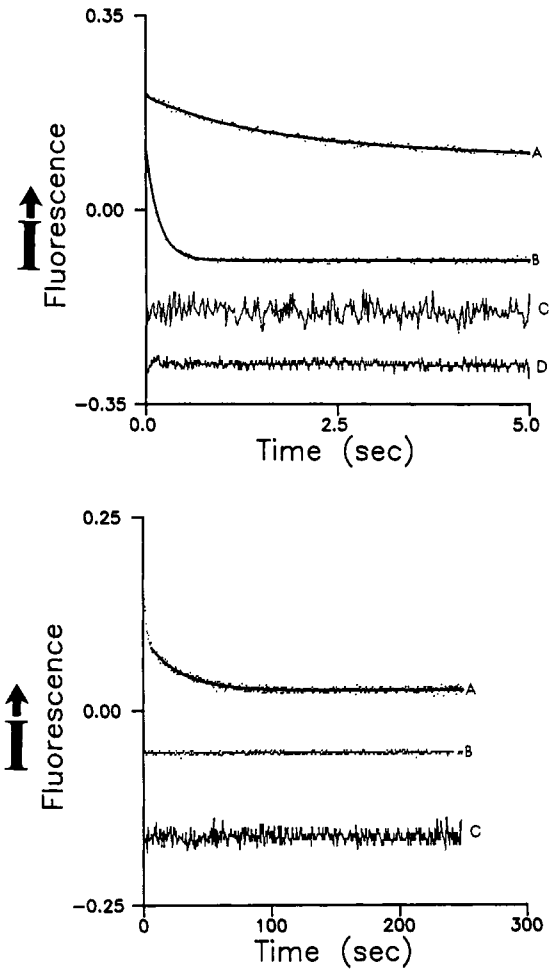


FIGURE 1: Comparison of kinetics of unfolding of SNase in strongly unfolding (pH 2.2) and a weakly unfolding (pH 3.5) conditions. Reactions which occurred in 2 ms to 5 s are recorded in panel A (top) and those which occurred in 2 ms of 250 s are recorded in panel B (bottom). Deconvolution showed that unfolding to pH 3.5 (curves A) gave three kinetic phases [two exponential decay constants, 0.81 s^{-1} and 0.2 s^{-1}, in panel A; and one, 0.039 s^{-1}, in panel B]. However, unfolding to pH 2.2 (curves B) gave a single kinetic phase [one exponential decay constant, 6.0 s^{-1}, in panel A and no detectable signals in panel B]. Deconvolutions were done with the nonlinear least-squares program supplied by the manufacturer of the instrument, which was based on the numerical procedure of Marquardt (1963). The fast reaction was assigned the τ_1 reaction, the middle speed one, the τ_2 reaction, and the slow one, the τ_3 reaction. Fitting these kinetic records into more than three relaxation times did not improve the chi-square value. Instead, it resulted in the degeneracy of relaxation times. In panel A, the residual signals for the kinetic records are shown in curves C and D, respectively, for curves A and B. In panel B, curve C is residual signal for curve A. No systematic deviations are discernible from these residual signals. Numerical values of the deconvolutions are listed below. Accepted values are indicated in boldface. The uncertainty in these experiments was ±5% for the amplitude and ±8% for the relaxation time. All rates reported here were independent of protein concentrations in the range 5–30 μM. The temperature was 25 °C. The following data represent deconvolution of kinetic records into eq 1 for i = 1, 2, or 3. Panel A (top): time range 2 ms to 5 s. Curve A: for i = 1, k_1 = 0.56, A_0 = 0.098, A_1 = 0.13, chi sq = 1.16; for i = 2, **k_1 = 0.81, k_2 = 0.20, A_0 = 0.081, A_1 = 0.07, A_2 = 0.06, chi sq = 1.11**; for i = 3, k_1 = 2.50, k_2 = 0.40, k_3 = 0.47, A_0 = 0.095, A_1 = 0.01, A_2 = –0.03; A_3 = 0.15, chi sq = 1.17. Curve B: for i = 1, **k_1 = 6.0, A_0 = –0.091, A_1 = 0.2, chi sq = 0.756**; for i = 2, k_1 = 7.2, k_2 = 4.9, A_0 = 0.91, A_1 = 0.11, A_2 = 0.09, chi sq = 0.758; for i = 3, k_1 = 6.5, k_2 = 6.6, k_3 = 3.8, A_0 = –0.091, A_1 = 0.09, A_2 = 0.08, A_3 = 0.03, chi sq = 0.761. Panel B (bottom): time range 5 s to 250 s. Curve A: for i = 1, **k_1 = 0.039, A_0 = 0.025, A_1 = 0.07, chi sq = 1.06**; for i = 2, k_1 = 0.049, k_2 = 0.037, A_0 = 0.025, A_1 = 0.014, A_2 = 0.056, chi sq = 1.07; for i = 3, k_1 = –9.6, k_2 = 0.043, k_3 = 0.037, A_0 = 0.025, A_1 = 0.007, A_2 = 0.032, A_3 = 0.031, chi sq = 1.06. Curve B: no signals.

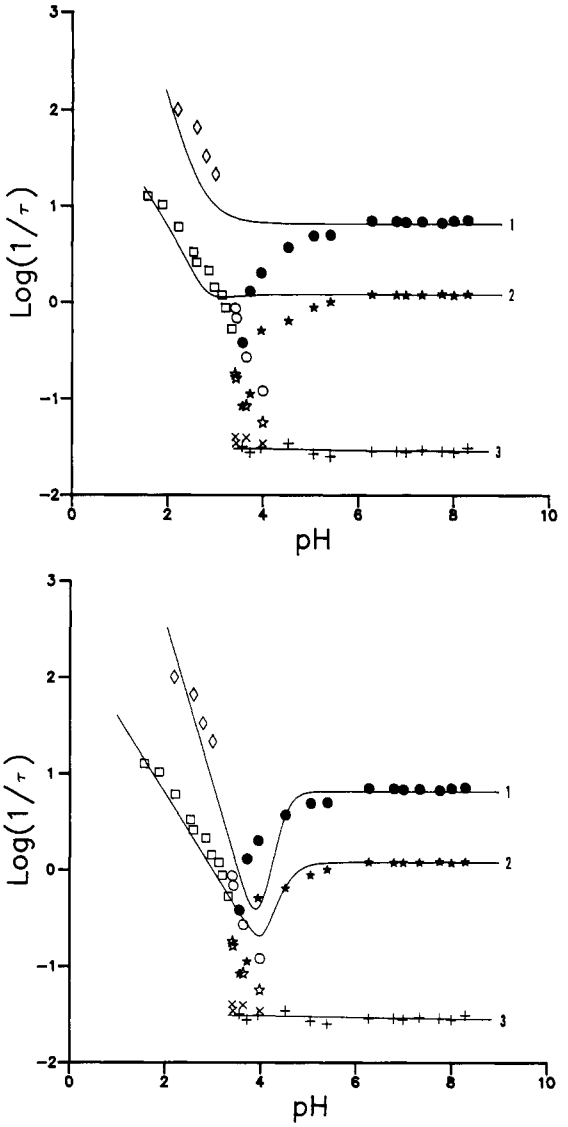


FIGURE 2: (A, top) Dependence of log τ_i^{-1} (i = 1, 2, or 3) on pH for the folding/unfolding of SNase. Data in the filled circles, filled stars, and plus signs were obtained by direct pH jump from pH 3.1 (folding). Data in open circles, open stars, and multiplication signs were obtained by direct pH jump from pH 7.0 (unfolding). Data in open diamonds were obtained by the double pH jump measurements. Curves 1 and 2 were drawn, respectively, for the τ_1 and τ_2 reactions (eqs 2 and 3), according to kinetic Scheme II, using the parameters of eq 7. (B, bottom) The same set of data were analyzed according to Scheme III. Curves 1 and 2 were drawn for the τ_1 and τ_2 reactions using eqs 5 and 6. The parameters used are shown in eqs 7 and 8. Experimental conditions are similar to those of Figure 1.

were obtained by the sequential jump, i.e., unfolding from pH 7.0 to an acidic pH followed by a second jump back to pH 7.0, after a delay time (t_D). Each data point was derived from several measurements with different t_D values. One interesting feature of these kinetic constants is the linkage of data for the unfolding and refolding reactions when they were measured at the same pH values. Another is the $\sqrt{\ }$-shaped dependence on pH for the two faster kinetic constants (λ_1 and λ_2). The rate of the slowest reaction λ_3 showed no dependence on pH down to pH 3. Analysis of these results is given below.

pH-Dependent Equilibrium Melting of Different Kinetic Species. As was shown previously (Chen et al., 1991, 1992), the amplitudes of the three kinetic phases in refolding to pH 7 are approximately equal to the equilibrium distribution of protein in the three substates of the unfolded form, D_1, D_2, and D_3, at the initial pH. The kinetic method simply provides a signal for monitoring the concentration of each species. By

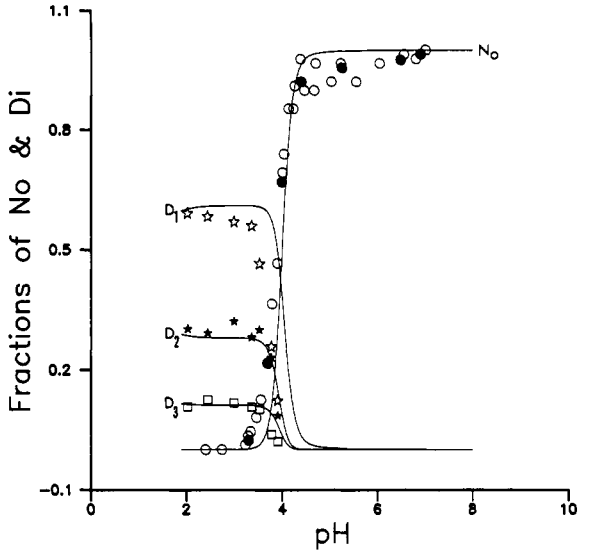


FIGURE 3: pH dependence of the equilibrium distribution of SNase in N_0 (open and closed circles), D_1 (open stars), D_2 (filled stars), and D_3 (open squares) forms. The data in the open circles were obtained with equilibrium pH titration from 7.0 to 2.0, monitoring fluorescence change of Trp140, and those in the closed circles were obtained by back-titration. The coincidence of the two titration curves ensures the reversibility of the folding/unfolding transition. Data points for D_1, D_2, and D_3 were obtained from the amplitudes of the τ_1, τ_2, and τ_3 reactions, respectively. These amplitudes have been shown to be approximately equal to the equilibrium distribution of the three D's at a given pH (Chen et al., 1991, 1992). Solid curves were obtained by the kinetic analysis outlined in the text according to eq 10, using the parameters described in eqs 7 and 8. The experiment was done at 25 °C. The final solutions contained 0.12 M NaCl, 0.05 M sodium acetate, 0.06 M sodium phosphate, 0.05 M glycine, and 20 μM EDTA, at the designated pH. The concentration of SNase in the initial medium was 60 μM. The final concentration was 30 μM.

starting from different pHs to refold to pH 7.0, the populations of the three D's as a function of pH were determined. The result is shown in Figure 3. The transition curve of N_0 was obtained by the equilibrium pH titration monitoring the Trp140 fluorescence. The forward titration is given in the open circles, and the back titration is given in the filled circles. The coincidence of the forward titration and the backward titration assures that the folding/unfolding was fully reversible. The amplitudes for the fast reaction (D_1), the mid-range reaction (D_2), and the slow reaction (D_3) all exhibit a sharp transition around pH 3.8. However, the ratios of the three D's at different pHs remained rather constant. Analysis of these transition curves is given below.

Kinetic Evidence of Rapidly Converting Microscopic States. The $\sqrt{}$-shaped dependence of log λ_1 and log λ_2 versus pH (Figure 2) cannot be explained by Scheme I. For clarity and simplification of analysis, we will consider only a subset of Scheme I, Scheme II for the analysis of the two faster reactions, τ_1 and τ_2, shown in Figure 2:

Scheme II

$$D_2 \underset{k_{12}}{\overset{k_{21}}{\rightleftharpoons}} D_1 \underset{k_{01}}{\overset{k_{10}}{\rightleftharpoons}} N_0$$

This choice is justified because the D_3 to D_2 transition (λ_3) was 10–30-fold slower than the two faster reactions in the entire pH range studied (Figure 1) and λ_3 was also independent of pH. Thus, the omission of D_3 to D_2 from Scheme I would not have strong effect on the behavior of Scheme II, as we shall see later. Scheme II is described by a set of three differential equations from which the two rate ($1/\tau$) equations were derived [see e.g., Szabo (1969), Rodiguin and Rodiguina (1964), and Bernasconi (1976)]:

$$\tau_1^{-1} = 0.5(k_{01} + k_{12} + k_{10} + k_{21}) + 0.5[(k_{01} - k_{12} + k_{10} - k_{21})^2 + 4k_{10}k_{12}]^{1/2} \quad (2)$$

$$\tau_2^{-1} = 0.5(k_{01} + k_{12} + k_{10} + k_{21}) - 0.5[(k_{01} - k_{12} + k_{10} - k_{21})^2 + 4k_{10}k_{12}]^{1/2} \quad (3)$$

Equations 2 and 3 contain four rate constants. One can experimentally measure τ_1 and τ_2 at different pH values and determine the four rate constants by the regression analysis. Two pieces of information had to be taken into account. First, our previous pH refolding experiments, in the range pH 5.8–9.0, have given $k_{10} = 6.5\ s^{-1}$ and $k_{21} = 6940\ s^{-1}$ (Chen et al., 1991) and shown that these two rate constants were independent of pH. Second, data in Figure 2, as well as our previous single- and double-pH jump experiments (Chen et al., 1991, 1992), in the pH range 2.0–3.4, have shown that k_{01} and k_{12} were pH dependent. Regression analysis was performed to fit the data in this pH range, for the two faster reactions in Figure 2A with eqs 2 and 3. We obtained $k_{12} = k^0_{12}[H^+]^{1.7}$ and $k_{01} = k^0_{01}[H^+]^{0.8}$ where k^0_{12} and k^0_{01} denote pH-independent parts of k_{12} and k_{01}, respectively. However, attempts to fit the complete data sets for the τ_1 and the τ_2 reactions with eqs 2 and 3 were unsuccessful. The $\sqrt{}$-shaped dependence of log τ_1^{-1} vs pH and of log τ_2^{-1} vs pH could not be reproduced. The best we could do was to simulate the upward trend of these two sets of data at pH below 3. The two solid curves drawn through data points in Figure 2A were produced with these relationships for the four rate constants.

To fit the data of Figure 2 to a kinetic model, it was necessary to modify Scheme II into a more complex form. Scheme III incorporates the concept of the microscopic states, where α_i and β_i are designated the unfolding-permitting subpopulations and the folding-permitting subpopulations, respectively, from the ith species.

Scheme III

$$\begin{array}{c} D_2(\alpha_2 \ldots \beta_2) \\ \Big\| \\ D_1(\alpha_1 \ldots \beta_1) \\ \Big\| \\ N_0(\alpha_0 \ldots \beta_0) \end{array}$$

(k_{21}, k_{12} between D_2 and D_1; k_{10}, k_{01} between D_1 and N_0)

Three differential equations describing Scheme III are

$$\left.\begin{aligned} d[N_0]/dt &= -k_{01}\alpha_0[N_0] + k_{10}\beta_1[D_1] \\ d[D_1]/dt &= k_{01}\alpha_0[N_0] - (k_{10}\beta_1 + k_{12}\alpha_1)[D_1] + k_{21}\alpha_2[D_2] \\ d[D_2]/dt &= k_{12}\alpha_1[D_1] - k_{21}\beta_2[D_2] \end{aligned}\right\} \quad (4)$$

Solving eq 4 gave the rate equations for the τ_1 and τ_2 reactions:

$$\tau_1^{-1} = 0.5\,(\alpha_0 k_{01} + \alpha_1 k_{12} + \beta_1 k_{10} + \beta_2 k_{21}) + 0.5[(\alpha_0 k_{01} - \alpha_1 k_{12} + \beta_1 k_{10} - \beta_2 k_{21})^2 + 4\alpha_1 k_{12}\beta_1 k_{10}]^{1/2} \quad (5)$$

$$\tau_2^{-1} = 0.5\,(\alpha_0 k_{01} + \alpha_1 k_{12} + \beta_1 k_{10} + \beta_2 k_{21}) - 0.5\,[(\alpha_0 k_{01} - \alpha_1 k_{12} + \beta_1 k_{10} - \beta_2 k_{21})^2 + 4\alpha_1 k_{12}\beta_1 k_{10}]^{1/2} \quad (6)$$

Regression analysis of the $\sqrt{\ }$-shaped data on the τ_1 and τ_2 reactions in Figure 2 (solid curves drawn across the data points in Figure 2B) gave the following relationships for the four rate constants.

$$\left.\begin{aligned} k_{01} &= 10^{(2.4-0.8\mathrm{pH})}\ \mathrm{s}^{-1};\quad k_{10} = 6.5\ \mathrm{s}^{-1} \\ k_{12} &= 10^{(5.5-1.7\mathrm{pH})}\ \mathrm{s}^{-1};\quad k_{21} = 6940\ \mathrm{s}^{-1} \end{aligned}\right\} \quad (7)$$

One should mention that in the pH around 4, the time separations of the three relaxation times are not great. The use of Scheme III instead of Scheme IV for analysis would be subject to larger uncertainty. This is reflected in the greater departure of the regression curve from data points around pH 4 in Figure 2B. However, this uncertainty would not change the conclusion drawn in this communication.

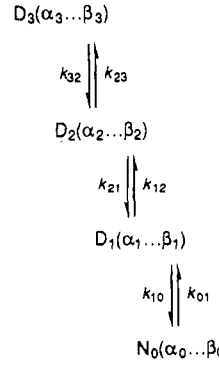
The regression analysis of the experimental data (Figure 2B) also produced the relationships for α_i and β_i.

$$\left.\begin{aligned} \alpha_0 &= 1.0;\quad \beta_1 = [1 + 10^{(13.5-3.0\mathrm{pH})}]^{-1} \\ \alpha_1 &= 1.0;\quad \beta_2 = (1.73 \times 10^{-4})[1.0 + 10^{(9.0-2.0\mathrm{pH})}]^{-1} + [1.0 + 10^{(1.7\mathrm{pH}-2.0)}]^{-1} \end{aligned}\right\} \quad (8)$$

These quantities of eq 8 are plotted vs pH in Figure 4. Both β_1 and β_2 show sigmoidal dependence on pH. β_1 exhibits a transition from 1 at higher pHs to small values at lower pHs (midpoint around pH 4.5), whereas β_2 exhibits a transition in the opposite direction (midpoint around pH 1.2). The meaning of these results is discussed below.

pH-Dependent Equilibrium Distributions of the Folding Intermediates. If the kinetic analysis described above is a good approximation of the pH-induced folding/unfolding of SNase, the relationships given in eqs 7 and 8 for the rate constants, α_i and β_i, should be able to reproduce the pH-dependent equilibrium melting curves of Figure 3. For this purpose, we have considered the complete kinetic Scheme IV:

Scheme IV



A set of four differential equations similar to eq 4 were solved for the three rates. The first rate equation (τ_1^{-1}) is identical to eq 5 and the second rate equation (τ_2^{-1}) to eq 6. The third rate equation is

$$\tau_3^{-1} = \beta_3 k_{32} + \alpha_2 k_{23} B/(1 + A + B) \quad (9)$$

and the normalized concentrations of the four molecular species at equilibrium (i.e., $t \rightarrow \infty$) in Scheme IV are

$$\left.\begin{aligned} \mathrm{N_0} &= 1/(1 + A + B + C) \\ \mathrm{D_1} &= A/(1 + A + B + C) \\ \mathrm{D_2} &= B/(1 + A + B + C) \\ \mathrm{D_3} &= C/(1 + A + B + C) \end{aligned}\right\} \quad (10)$$

In eqs 9 and 10

$$\left.\begin{aligned} A &= \alpha_0 k_{01}/(\beta_1 k_{10}) \\ B &= \alpha_0\alpha_1 k_{01} k_{12}/(\beta_1\beta_2 k_{10} k_{21}) \\ C &= \alpha_0\alpha_1\alpha_2 k_{01} k_{12} k_{23}/(\beta_1\beta_2\beta_3 k_{10} k_{21} k_{32}) \end{aligned}\right\} \quad (11)$$

A and B in eq 11 were calculated from eqs 7 and 8, and C was assigned to be equal to $0.36B$. This assignment was based on the following facts. First, from eq 11, $C/B = \alpha_2 k_{23}/\beta_3 k_{32}$. $\alpha_2 k_{23}$ and $\beta_3 k_{32}$ are the apparent rate constants for unfolding and refolding of D_2, respectively, and have been determined to be 0.033 and 0.091 s^{-1}, by the double jump experiment (Chen et al., 1992). Thus, their ratio is 0.36. And since $k_{23}/k_{32} = [D_3]/[D_2]$, which was determined to be 0.37 (0.11/0.30) from the data in Figure 3 and from our previous experiments (Chen et al., 1991, 1992), it follows then that $\alpha_2/\beta_2 \approx 1$. A, B, and C were then determined at different pH values, and the pH-dependent equilibrium folding/unfolding transition curves of N_0, D_1, D_2, and D_3 were calculated, as shown with the solid lines drawn across the data points in Figure 3. Experimentally, the fraction of protein in N_0 was

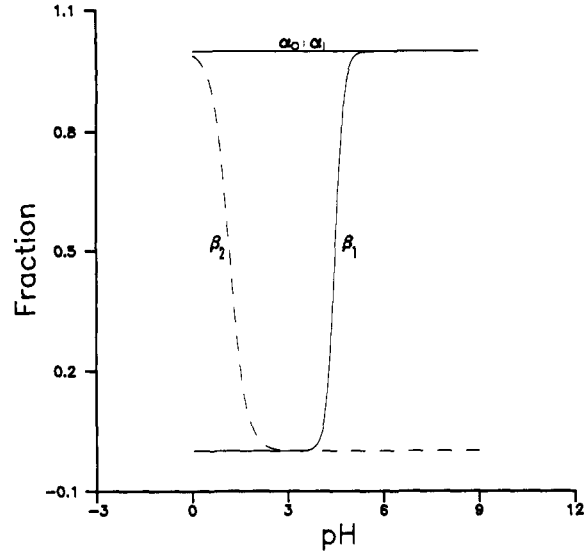


FIGURE 4: pH-dependent distribution of microscopic states of Scheme III. α_0 and α_1 are the subpopulations of microscopic states, from which unfolding can take place from N_0 and from D_1, respectively. β_1 and β_2 are the subpopulations of microscopic states from which folding can take place from D_1 and D_2, respectively. These curves were drawn from eq 8. Both α_3 and β_3 are unity in the pH range 3–8.5. See text for details.

determined by the pH titration monitoring the Trp140 fluorescence and those for the three D's were determined by the amplitudes of the three kinetic phases in the folding from pH 2.2 to pH 7.0 [see Chen et al. (1991, 1992)]. All four calculated curves agreed well with the experimental data. One should mention that kinetic and equilibrium data of Figure 2 inside the transition zone, i.e., in the range pH 3.3–4.5, were subject to large uncertainties (±10–±20%) because, in this pH range, only fractions of the fluorescence signals for the folding/unfolding transition were available. Thus, the theoretical fits in Figures 2 and 3 must be considered excellent.

DISCUSSION

The failure of the Scheme II to produce the experimental observation of Figure 2 cannot be due to the omission of the third kinetic step, i.e., the $D_2 \rightleftharpoons D_3$ transition in our analysis. We have shown that this third step does not depend on pH and its rate is at least 1 order of magnitude slower than that of the two faster reaction in the entire pH range studied, except in the narrow range of pH 3.6–4.0. If the log τ^{-1} versus pH plots were to assume a simple V-shape, it would be possible to reproduce data with Scheme II, with the assumption that all four rate constants were dependent on pH. The V-shaped dependence of log τ^{-1} versus θ, the order parameter for a protein, has indeed been reported frequently in the literature [see, e.g., Brandts et al. (1977), Hagerman and Baldwin (1976), Ikai et al. (1973), Kim and Baldwin (1982), Tsong (1976), and Sugawara et al. (1991)], and analysis of these data has been straightforward. However, because of the improved sensitivity of the stopped-flow kinetic instrument, we were able to resolve, with confidence, the more complex feature of the pH-induced folding/unfolding of SNase. That is, in the pH range 5.5–9.0, the three folding reactions were independent of pH. The presence of this flat regions of the τ_1 and τ_2 reactions made the log τ^{-1} versus pH plots $\sqrt{\ }$-shaped, and these data could not be reproduced with Scheme II. Unquestionably, there will be a number of ways to resolve this difficulty. Our experiments and analysis of data would suggest that each of the discrete kinetic species, N_0 and three D's (separated by an activation barrier of 6–14 kcal mol^{-1}), is composed of microscopic states in rapid equilibrium.

The concept of microscopic states is apparent in the treatment of the helix–coil transitions of DNA and polypeptides (Schwartz, 1965). The cluster merging kinetics for protein folding also encompass a similar concept (Kanehisa & Tsong, 1978). Notwithstanding, the separation of microscopic states into two subpopulations, one folding-permitting and the other unfolding-permitting, is unique. The present analysis also shows that unfolding by direct pH jump to 3 detected the τ_2 reaction rather than the τ_1 reaction (Figure 2). We should caution that our analysis is tentative and future study should attempt to prove directly by experiments the existence of these microscopic states.

The nature of the three D states has been discussed (Chen et al., 1991, 1992). Basically, they have similar free energies (a difference of less than 0.5 kcal mol^{-1} between any two states) but are separated by activation barriers of 6–14 kcal mol^{-1}. It was suggested that the D_3 to D_2 transition belongs to the cis/trans isomerization of one or more of the six proline residues because the time constant (30 s) and the activation energy (13.6 kcal mol^{-1}) fall in the range for the proline isomerization (Kim & Baldwin, 1982, 1990). This interpretation would be consistent with the observation of Evans et al. (1987, 1989). However, our preliminary data on the Pro117Ala mutant indicate that Pro117 may not be directly responsible for the slow fluoresence signals (unpublished results). The amplitude of the slow phase increased substantially (instead of decreasing), from 11% for the wild type to 60% for the Pro117Ala mutant. Other prolines would have to be considered. The D_2 to D_1 transition may have detected the condensation of an extended unfolded form to a compact unfolded form because the reaction was dependent on the solvent viscosity (to be published). From the three pH-dependent transition curves of the D states in Figure 3, free energies of conversion between any two of these D states can be calculated for the pH range 1–3.8. Apparently, these free energies are small (<0.5 kcal mol^{-1}) and do not vary substantially with pH.

The multiple isomeric forms of the native state detected by the NMR studies (Fox et al., 1986; Evans et al., 1987, 1989; Alexandrescu et al., 1989) are not necessarily contradictory to kinetic Scheme IV. The N″ form was present only at high concentration of SNase (Alexandrescu et al., 1989). In the range of concentration we used, no N″ form should be detectable. The conversion of N to N* (Fox et al., 1986) may be slow at neutral pH but becomes too fast to be resolved by our stopped-flow measurement under the unfolding pH range.

Our data also show that unfolding from N_0 to D_1 was catalyzed by 0.8 proton and that from D_1 to D_2 was catalyzed by 1.7 proton. The pH-dependent equilibrium unfolding curve given in Figure 3 has been shown to involve protonation of 2.6 ionizable groups (Chen et al., 1991). Thus, electrostatic interactions play a role in the unfolding of N_0 and D_1. However, folding was likely driven by hydrophobic interactions rather than electrostatic interactions. One should mention that the numbers 0.8, 1.7, and 2.6 represent only the average numbers of groups protonated in these events.

From the above analysis, one may derive some information on the physical properties of the microscopic states. First, these microscopic states are separated neither by large free energy differences nor by large activation barriers. Second, the conversion among them is much faster than milliseconds. Whether the physical events of conversion involve formation/disintegration of local structures, such as those fast kinetics detected by the CD (Sugawara et al., 1991), or motions of the peptide chain (molecular dynamics) remains unclear.

ACKNOWLEDGMENT

We thank Carol J. Gross for help with the manuscript.

REFERENCES

Alexandrescu, A. T., Ulrich, E. L., & Markley, J. L. (1989) *Biochemistry 28*, 204–211.
Bernasconi, C. F. (1976) *Relaxation Kinetics*, Academic Press, New York.
Brandts, J. F., Halvorson, H. R., & Brennan, M. (1975) *Biochemistry 14*, 4953–5963.
Brandts, J. F., Brennan, M., & Lin, L.-N. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4178–4181.
Chen, H. M., You, J. L., Markin, V. S., & Tsong, T. Y. (1991) *J. Mol. Biol. 220*, 771–778.
Chen, H. M., Markin, V. S., & Tsong, T. Y. (1992) *Biochemistry 31*, 1483–1491.
Davis, A., Parr, G. R., & Taniuchi, H. (1979) *Biochim. Biophys. Acta 578*, 505–510.
Eftink, M. R., Ghiron, C. A., Kautz, R. A., & Fox, R. O. (1989) *Biophys. J. 55*, 575–579.
Epstein, H. F., Schechter, A. N., Chen, R. F., & Anfinsen, C. B. (1971) *J. Mol. Biol. 60*, 499–508.

Evans, P. A., Dobson, C. M., Kautz, Hatfull, G., & Fox, R. O. (1987) *Nature (London) 329*, 266–268.
Evans, P. A., Kautz, R. A., Fox, R. O., & Dobson, C. M. (1989) *Biochemistry 28*, 362–370.
Fox, R. O., Evans, P. A., & Dobson, C. M. (1986) *Nature (London) 320*, 192–194.
Fuchs, S., Cuatrecasas, P., & Anfinsen, C. B. (1967) *J. Biol. Chem. 242*, 4768–4770.
Hagerman, P. J., & Baldwin, R. L. (1976) *Biochemistry 15*, 1462–1473.
Hagerman, P. J., Schmid, F. X., & Baldwin, R. L. (1979) *Biochemistry 18*, 293–297.
Ikai, A., Fish, W. W., & Tanford, C. (1973) *J. Mol. Biol. 73*, 145–163.
Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol. 49*, 117–237.
Kanehisa, M. I., & Tsong, T. Y. (1978) *J. Mol. Biol. 124*, 177–194.
Kim, P. S., & Baldwin, R. L. (1982) *Annu. Rev. Biochem. 51*, 459–489.
Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631–660.
Marquardt, D. W. (1963) *J. Soc. Indust. Appl. Math. 11*, 431–441.
Nakano, T., & Fink, A. L. (1990) *J. Biol. Chem. 265*, 12356–12362.
Rodiguin, N. M., & Rodiguina, E. N. (1964) *Consecutive Chemical Reaction*, D. Van Nostrand Co., Inc., Princeton, NJ.
Schechter, A. N., Chen, R. F., & Anfinsen, C. B. (1970) *Science (Washington, D.C.) 167*, 886–887.
Schwartz, G. (1965) *J. Mol. Biol. 11*, 64–77 (1965).
Shortle, D., & Meeker, A. K. (1989) *Biochemistry 28*, 936–944.
Sugawara, T., Kuwajima, K., & Sugai, S. (1991) *Biochemistry 30*, 2698–2706.
Szabo, Z. G. (1969) in *Comprehensive Chemical Kinetics* (Bamford, C. H., & Tipper, C. F. H., Eds.) Chapter 1, Vol. 2, Elsevier, Amsterdam.
Tsong, T. Y. (1976) *Biochemistry 15*, 5467–5473.