A Quantitative Treatment of the Kinetics of the Folding Transition of Ribonuclease A†

Paul J. Hagerman‡ and Robert L. Baldwin*

ABSTRACT: New experimental data and a quantitative theoretical treatment are given for the kinetics of the thermal folding transition of ribonuclease A at pH 3.0. A three-species mechanism is used as a starting point for the analysis: U_1 (slow) $\rightleftharpoons$ U_2 (fast) $\rightleftharpoons$ N, where U_1 and U_2 are two forms of the unfolded enzyme with markedly different rates of refolding and N is the native enzyme. This mechanism is based on certain facts established in previous studies of refolding. The kinetics of unfolding and of refolding show two phases, a fast phase and a slow phase, over a range of temperatures extending above the transition midpoint, T_m. The three-species mechanism can be used in this range. At higher temperatures a new, much faster, kinetic phase is also observed, corresponding to the transient formation of a new intermediate (I). Although the general solution for a four-species mechanism is complex, it is not difficult to extend the three-species analysis for the special case found here, in which the fast reaction (I $\rightleftharpoons$ N) is well separated from the other two reactions. At temperatures below the transition zone the slow phase of refolding becomes kinetically complex. No attempt has been made to extend the analysis to include this effect. The basic test of the three-state analysis is the prediction as a function of temperature of α_2, the relative amplitude of the fast phase, both for unfolding and refolding. At temperatures above T_m, for which the three-state analysis must be extended to include the new intermediate I, a corresponding quantity α_2(cor) is predicted and compared with measured values. Data used in the three-state prediction are values of τ_2 and τ_1, the time constants of the fast and slow kinetic phases, plus a single value of α_2 measured when τ_2 and τ_1 are well separated. The observed and predicted values of α_2 agree within experimental error. The analysis predicts correctly that, for these experiments, α_2 should have the same value in unfolding as in refolding in the same final conditions. The analysis also predicts satisfactorily the equilibrium transition curve from kinetic data alone. Four striking properties of the kinetics are explained or correlated by the analysis: (a) the drop in α_2 to a minimum near T_m as well as the delayed rise in α_2 above T_m; (b) the vanishing of α_1 above the transition zone; (c) the sharp drop in τ_1 inside the transition zone followed by a partial leveling off outside this zone; and (d) the passage of τ_2 through a maximum near T_m. Through a comparison of observed and predicted values of α_2, the analysis also rules out the alternative three-species mechanism U_1 (slow) $\rightleftharpoons$ N (fast) $\rightleftharpoons$ U_2. Finally, the temperature dependence of the amplitude for the fast reaction (I $\rightleftharpoons$ N) is discussed: the behavior of I is like that of U_2, and I may be an unfolded species populated at equilibrium. If so, I accounts for only 2% of the total unfolded enzyme and would not be detected in refolding experiments below T_m. Possible molecular interpretations of the $U_1 \rightleftharpoons U_2 \rightleftharpoons$ I $\rightleftharpoons$ N mechanism are discussed briefly.

Temperature-jump and stopped-flow kinetic studies demonstrate the presence of intermediates in the kinetically reversible, thermal folding transition[1] of RNase A.[2] The observation of a fast kinetic process[3] in the millisecond time range, in addition to a slow process in the second time range (Tsong et al., 1971, 1972a), shows that at least three molecular species participate in the folding transition. Some facts about the mechanism of refolding are now known (Garel and Baldwin, 1973, 1975a). (1) Both fast and slow processes in refolding yield native enzyme possessing catalytic activity and inhibitor binding capability. (2) The fast-refolding species is a stable intermediate existing in a pH-dependent equilibrium with the slow-refolding species. (3) The fast- and slow-refolding species are nearly equivalent optically, as measured by exposure of tyrosine groups that are buried in the native protein. (4) In the initial conditions, the fast-refolding species is present at equilibrium above the transition zone, and its ratio to the slow-refolding species is independent of temperature. This last observation implies the existence of at least two distinct forms of heat-unfolded protein, a result confirmed recently by Brandts et al. (1975).

The model proposed (Garel and Baldwin, 1973, 1975a,b) to account for these features of the folding process is a linear three-state model

$$\underset{\text{slow}}{U_1} \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} \underset{\text{fast}}{U_2} \underset{k_{32}}{\overset{k_{23}}{\rightleftharpoons}} N \qquad \text{(a)}$$

in which U_1 and U_2 are present at equilibrium in and above

† From the Department of Biochemistry, Stanford Medical School, Stanford, California, 94305. *Received July 22, 1975.* This research has been supported by grants from the National Science Foundation (BMS72-02262 A02) and National Institutes of Health (5 RO1 GM 19988-15), and by a Medical Scientist Training Program Grant from the National Institutes of Health, GM 1922.

‡ Predoctoral fellow, Medical Sciences Training Program.

* To whom reprint requests should be addressed.

[1] The term "protein folding/unfolding" is defined operationally as the reversible passage of a protein (S-S bonds intact) through the thermally induced transition zone. No specific structural changes are implied.

[2] Abbreviations used: RNase A, bovine pancreatic ribonuclease A; T_m, temperature midpoint of the transition curve for thermal unfolding; τ, time constant of a reaction or kinetic phase (reciprocal of the apparent rate constant); α, relative amplitude of a kinetic phase (fraction of the total amplitude).

[3] The term "fast phase" refers to the process associated with τ_2 and is to be distinguished from a much faster process (τ_3) observed above T_m.

the transition zone and N (native enzyme) is present in and below the transition zone. No attempt has been made previously to distinguish this from a second mechanism, (b), which has the same three species but a different sequence of reactions.

$$U_1 \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} N \underset{k_{32}}{\overset{k_{23}}{\rightleftharpoons}} U_2 \qquad \text{slow} \quad \text{fast} \tag{b}$$

The structural basis of the difference in refolding rates between U_1 and U_2 is not known, but it seems likely that it is a difference in configuration of the unfolded chain since the U_2^0/U_1^0 ratio is independent of temperature (Garel and Baldwin, 1975a). Brandts et al. (1975) have suggested that slow cis–trans isomerizations of prolyl residues are responsible: i.e., that rapid refolding ($U_2 \rightleftharpoons N$) of a heat-unfolded protein can occur only when each prolyl residue has the same conformation in the unfolded and native states.

Some kinetic properties of the RNase A folding transition are puzzling: why should the relative amplitude of the fast phase drop sharply as the temperature approaches T_m (Tsong et al. 1972a)? Other properties become puzzling when the simple three-species mechanism (a) or (b) is used to explain the kinetic amplitudes. For example, these mechanisms predict that the amplitude of the slow phase should vanish at a temperature slightly above the transition zone, in the pH range 2–3, and yet the slow phase is reported to persist to higher temperatures in studies of unfolding at pH 2.0 (Tsong et al., 1972a).

The purpose of this work is to test the three-species mechanism (a) by a quantitative analysis of the kinetic amplitudes, making use of the fact that the slow and fast refolding reactions are well resolved in certain conditions. This allows the principle of separability to be used. Use is also made of the known optical equivalence of U_1 and U_2. The basic result of the analysis is the prediction of amplitudes, both for unfolding and refolding. The puzzling kinetic properties mentioned above are explained, and the results distinguish between mechanisms (a) and (b).

In two different temperature ranges, the kinetics become more complex. At temperatures below the transition zone, the slow refolding reaction becomes kinetically complex. We continue to use the amplitude analysis based on the three-species mechanism with the assumption that the predicted amplitude for the slow phase can be compared with the sum of the experimental amplitude changes in this phase. At high temperatures, above the transition midpoint, a new intermediate (I) becomes populated. It is the first species formed in unfolding from N. Because this new reaction is well separated from the other two, it becomes a simple matter to extend the amplitude analysis to include this species explicitly, and the model is expanded to the four-species mechanism $U_1 \rightleftharpoons U_2 \rightleftharpoons I \rightleftharpoons N$ at temperatures above T_m.

A test for abortive intermediates in refolding has been devised by Ikai and Tanford (1971, 1973), based on the relative signs of the fast and slow phases in unfolding and refolding. This test does not apply here because it is valid only when there is a single native state N and a single unfolded state U. The relation between our present analysis and the one worked out by Ikai and Tanford (1971, 1973) will be discussed elsewhere (Hagerman, in preparation).

Theory

General Considerations. In the theoretical analysis of complex kinetic systems such as the linear sequential model

$$A_1 \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} A_2 \rightleftharpoons \cdots \underset{k_{N,N-1}}{\overset{k_{N-1,N}}{\rightleftharpoons}} A_N$$

used to describe helix–coil transitions for various biopolymers, the concentration of each species (c_i) can be expressed as a sum of terms relating to each characteristic rate process. Here

$$c_i(t) = \sum_{j=0}^{N} C_{ij} \exp(-t/\tau_j) \tag{1}$$

where the C_{ij} and τ_j contain information regarding the elementary rate constants of the system. The overall change in any measured property, such as absorbance, can be expressed in terms of contributions from each species as

$$\mathrm{ABS}(t) = \sum_{i=1}^{N} \epsilon_i c_i(t) \tag{2}$$

where the ϵ_i represent, in this instance, molar extinction coefficients associated with each species.

For all but the simplest cases, general analytic expressions for the C_{ij} and τ_j are not obtainable, and numerical methods must be employed. The process of decomposing a complex system into much simpler systems, capable of explicit solution, thus assumes a role of great importance as an analytical tool. In particular, the three-state model used to describe the RNase A kinetics is amenable to an exact analytical treatment, and such a treatment will therefore be described in more detail.

Three-State Model. In the exact treatment of the linear three-state model, equations for all three species (eq 1) take the form

$$c_i(t) = C_{i0} + C_{i1} \exp(-t/\tau_1) + C_{i2} \exp(-t/\tau_2) \tag{3}$$

where c_1, c_2, and c_3 refer to the concentrations of U_1, U_2, and N in our case. The nine C_{ij} are coefficients associated with each kinetic phase and depend on both initial and final conditions (see Appendix). The τ_j depend only on the final conditions.

The expressions for τ_1 and τ_2 are well known (cf. Eigen and DeMaeyer, 1963) and are given by

$$\tau_{1(2)}^{-1} = \{\textstyle\sum k - (+)\ [(\sum k)^2 - 4\prod k]^{1/2}\}/2 \tag{4}$$

where

$$\textstyle\sum k = k_{12} + k_{21} + k_{23} + k_{32} = \tau_1^{-1} + \tau_2^{-1} \tag{5}$$

and

$$\textstyle\prod k = k_{12}(k_{23} + k_{32}) + k_{21}k_{32} = (\tau_1\tau_2)^{-1} \tag{6}$$

Under limiting conditions where $\tau_1/\tau_2 \gg 1$ and where k_{23}, $k_{32} \gg k_{12}$, k_{21}, τ_1 and τ_2 are given by the limiting forms

$$\tau_{1,\mathrm{lim}}^{-1} = k_{12} + k_{21}/(1 + K_{23}) \tag{7}$$

and

$$\tau_{2,\mathrm{lim}}^{-1} = k_{23} + k_{32} \tag{8}$$

In experiments where absorbance or binding capacity is measured in RNase A folding experiments, the expression for the relative amplitude associated with the fast process in refolding (R) or unfolding (U) is given by

$$\alpha_2^{R,U} = 1/[1 + |C_{31}^{R,U}|/|C_{32}^{R,U}|] \tag{9}$$

where C_{31} and C_{32} refer to the concentration (c_3) of species N and are simply the coefficients of the exponential terms

for τ_1 and τ_2, respectively. The specific amplitude analysis of RNase A kinetics is simplified by the fact that the molar absorbances of U_1 and U_2 are approximately equal, and only N binds 2′-CMP. Consequently the kinetics can be described in terms of the time dependence of N, when folding is measured either by 2′-CMP binding or by solvent exposure of buried tyrosine groups.

For refolding jumps beginning in conditions where f_N (fraction native, N) is approximately zero, eq 9 simplifies to[4,5] (see Appendix)

$$\alpha_2^R = 1/[1 + |\tau_1 k_{12} - U_2^0|/|\tau_2 k_{12} - U_2^0|] \quad (10)$$

where U_2^0 represents the initial concentration of U_2. In previous work (Garel and Baldwin, 1975a), it has been shown that the apparent concentration of U_2 relative to U_1 is independent of pH between pH 2 and pH 3. Therefore for refolding jumps considered in this paper, U_2^0/U_1^0 and hence K_{12} is constant. For unfolding, starting from f_N approximately equal to unity

$$\alpha_2^U = 1/[1 + |1 - \tau_1(k_{12} + k_{21})|/|1 - \tau_2(k_{12} + k_{21})|] \quad (11)$$

Under the limiting conditions specified above for eq 7 and 8, eq 10 reduces to

$$\alpha_{2,\mathrm{lim}}^R = U_2^0[1 + K_{21}/(1 + K_{23})] \quad (12)$$

Experimental Section

Materials. (a) RNase A: (i) Worthington No. RAF 53N523 phosphate-free lyophilized powder, stored at 0 °C; (ii) Worthington No. N3J405 solution passed over Sephadex G-25 (Pharmacia) to remove phenol and phosphate; stored in 0.1 N $NaClO_4$ at neutral pH and 0 °C. Concentration measured by absorbance at 278 nm, neutral pH, using molar absorbance of 9.8×10^3 (Sela and Anfinsen, 1957). (b) 2′-CMP: P.L. Biochemicals No. 273-10; concentration measured using a molar absorbance of 7.6×10^3 at 260 nm, neutral pH (Beaven et al., 1955). (c) Other reagents: $NaClO_4$, Fisher Scientific; cacodylic acid, Fisher Scientific; glycine, J. T. Baker. These reagents were used without further purification.

Solutions. All solutions were prepared with quartz-distilled water. Solutions were prepared according to protocol, adjusted to final pH, and finally filtered through 0.43 μm pore size Millipore filters. All solutions employed in measurements at or above ambient temperature were degassed for ½ to 1 h immediately prior to use. Solutions containing protein were heated to 60 °C for 10 min at neutral pH prior to pH adjustment and filtration in order to remove dimers (Crestfield et al., 1963). Final buffer conditions were as follows: (a) pH 3.0, 0.1 N $NaClO_4$, 0.1 M glycine; (b) pH 5.8, 0.1 N $NaClO_4$, 0.05 M sodium cacodylate. Initial protein solutions were unbuffered. All solutions were handled in glass syringes in order to avoid reactions between protein and plastic syringes (Garel and Baldwin, 1973).

pH Measurements. All reported pH values were measured at 22 °C on a Radiometer type PHM 26 pH meter using a Radiometer type GK 2301 C combined electrode. Beckman standard buffers were used for reference.

Equilibrium Absorbance Measurements. Equilibrium transition curves were performed on a Cary 15 recording spectrophotometer with a water-jacketed quartz cuvette. The cuvette temperature was driven linearly at 0.67 °C/min and was monitored continuously with an Atkins Model 3H5T-C45 meter and sensing probe. The possibility of incomplete equilibration of either temperature or protein was ruled out by running cooling curves immediately following each heating curve. Absence of any optical hysteresis ensured complete reversibility as well as complete equilibration. Other equilibrium absorbance measurements were made with unjacketed cells mounted in temperature-regulated brass blocks. A wavelength of 286.5 nm was employed for protein measurements.

Stopped-Flow Measurements. The apparatus employed for all stopped-flow measurements consisted of a Gibson-Durrum stopped-flow instrument with a number of specific modifications. (1) Light was provided by a 12V tungsten filament lamp which was supplied by a Hewlett-Packard 6267 B power supply set at 12 V, 8.4 A. (2) Wavelength selection (287 nm) was carried out with a Bausch and Lomb high intensity grating monochromator, 2700 grooves/mm, with a 1.34-mm entrance slit and a 0.75-mm exit slit. A 286-nm interference filter was placed between the monochromator and the observation chamber in order to remove scattered light passing through the monochromator, which becomes appreciable at wavelengths less than 300 nm. With this configuration, Beer's law was obeyed to optical densities greater than 1.0. (3) Transmitted light was detected with an EMI 9558Q end-on photomultiplier powered by a Power Designs Model 2K20 high-voltage power supply. It was observed that, when voltages less than 300 V were used, the response time of the PM increased to the point where kinetic complexity appears as a consequence of slow PM response. For this reason, PM voltages were always greater than 400 V and were usually in the range 500–600 V. (4) PM output was recorded on a Tektronix 564 storage oscilloscope. The signal was passed through a rise-time filter set at τ_{RC} less than one-tenth that of the smallest experimental τ. A bias voltage was provided to the oscilloscope from a Kepco bias supply. (5) The thermal regulation of the Gibson-Durrum instrument does not provide adequate temperature equivalence between the driving syringes and the observation chamber at temperatures away from ambient temperature, even if high-velocity-flow water baths are employed. For this reason, a temperature-control jacket was constructed to enclose the observation chamber and was regulated with an auxiliary water bath.

(b) Kinetic artifacts which appeared with some regularity could be divided into two main categories: those due to inadequate thermal control and those due to pressure effects. Temperature-related artifacts can arise as a result of schlieren effects due to mixing of solutions of different temperatures (Gibson, 1964), even if these temperatures differ by as little as a few hundredths of a degree Celsius. These effects produce complex kinetic behavior in the millisecond to second time range, and can be observed by mixing water with water in a stopped-flow experiment. To eliminate this artifact, water/water controls were run for every temperature setting, and the baths were adjusted until schlieren effects disappeared. An additional temperature-related artifact involves slight changes in τ's caused by temperature changes on mixing. These effects can be eliminated by premixing protein and buffer solutions, placing the mixture in both driving syringes, and adjusting the temperature baths until no kinetics are seen upon stopped-flow mixing of the

[4] The symbols U_1, U_2, N, and I refer to individual species, whereas with superscripts these symbols refer to the corresponding concentrations. The superscript (0) refers to initial concentration, and the superscript (f) refers to final concentration.

[5] In this paper, all k_{ij} refer to final conditions.

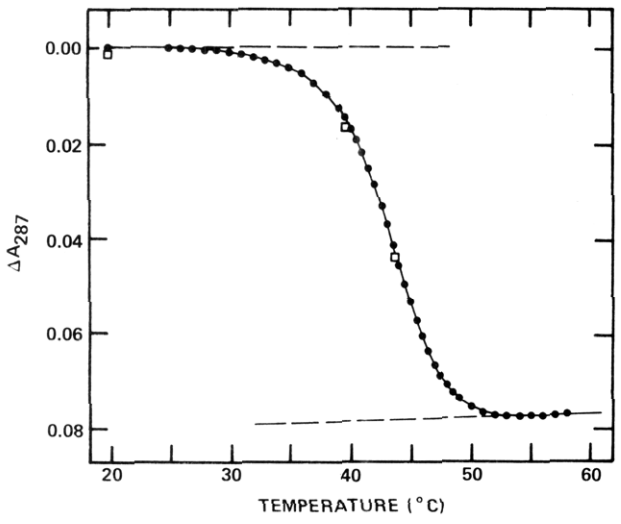


FIGURE 1: The transition curve for thermal unfolding of RNase A at pH 3.0, measured by following changes in absorbance of buried tyrosine groups at λ 287 nm. ΔA_{287} refers to the decrease in absorbance of the sample relative to an identical RNase A solution held constant at 20.0 °C. Closed circles (●) represent unfolding accompanying a continuous temperature rise of 0.67 °C/min. Open squares (□) represent refolding at equilibrated temperatures. Concentration of RNase A: 0.65 mg/ml (see Experimental Section for buffers).

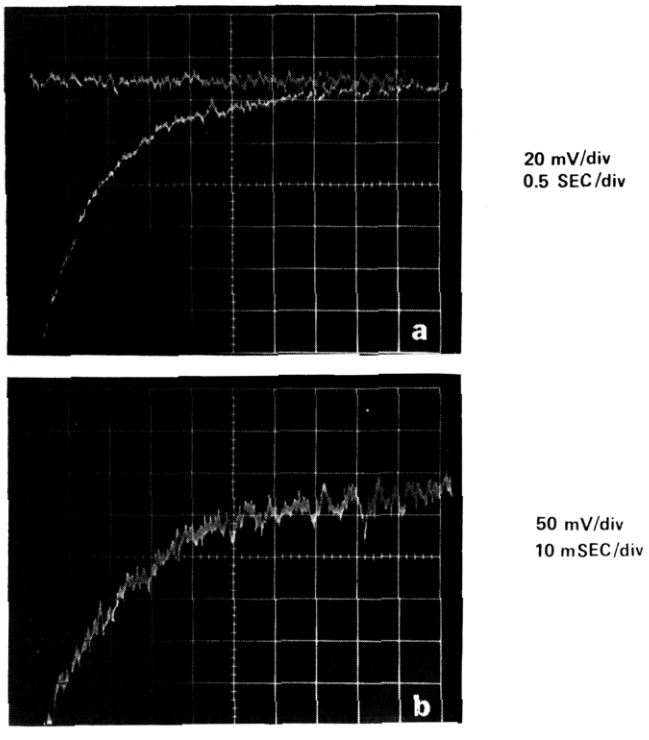


FIGURE 2: Oscilloscope recordings of unfolding kinetics for RNase A above T_m (pH 7.0 → 3.0, 53.6 °C). Final concentration of RNase A: 0.8 mg/ml (see Experimental Section for buffers). Kinetic analysis is based on successive recording of several mixing experiments performed under identical conditions.

identical protein solutions. This procedure was performed following the H_2O/H_2O control at each temperature setting. For all experiments, the contents of the driving syringes were allowed to equilibrate for at least 15 min prior to each set of mixing experiments. Pressure-related artifacts were found to be due to cavitation (inadequate degassing at high temperatures), slight optical misalignment, and movement of the observation chamber following the 80-ms pressure cutoff in the Durrum pressure actuator. The first of these artifacts can be removed by degassing all solutions for at least ½ h immediately before use. The last two artifacts can be eliminated by slightly defocusing the light entering the observation chamber.

Down Temperature Jumps. Solutions containing protein at pH 3 were equilibrated at 54 °C (or at 70 °C for pH 5.8, 2′-CMP solution) for 10 min, followed by rapid injection into a pre-cooled jacketed cuvette. Approximately 60 s were required for temperature equilibration and complete disappearance of schlieren effects. Final solution temperature was monitored with an in-dwelling temperature probe as described above. Absorbance changes at 286.5 nm were recorded as a function of time in a synchronous-time mode chart drive on a Cary 15 recording spectrophotometer.

Data Analysis. Oscilloscope traces were photographed with a Polaroid camera using Type 107 black and white film. Curves were retrieved from the Polaroid oscillographs by overlaying a transparent millimeter grid in order to digitize the trace. Plots of log $\Delta I(t)$ were constructed for data analysis from the digitized scope traces, which record $\Delta I(t)$. For all of the measurements performed in the present study, the transmittance changes are much smaller than the absolute transmittance, and the resultant errors in τ estimation associated with plotting log $\Delta I(t)$ instead of log $\Delta A(t)$ were always less than 4%. Analysis of these log plots was always carried out by "peeling back" the various kinetic phases in order of decreasing τ. The tail of the largest time component was typically extrapolated back to zero time to provide both an amplitude and a τ value for the terminal phase. This component was then subtracted from the log plot and the process was repeated until all amplitude had been accounted for. The use of several time domains of data collection greatly improved the precision of the data analysis and, when τ's were not widely separated, the kinetic experiment was repeated several times to further increase the precision of the data analysis.

Computations. All computer computations were performed on the IBM 360/67 campus facility using WATFIV or FORTRAN H compilers, and using extended precision mode arithmetic.

Results

pH-Jump Determination of Refolding and Unfolding Kinetics throughout the Thermal Transition Zone at Final pH 3.0. All pH-jump experiments consisted of rapid mixing of protein and glycine buffer solutions at constant temperature to produce a final pH of 3.0 and a final protein concentration of 0.5–1.0 mg/ml, for both refolding and unfolding experiments. The initial pH in unfolding experiments is 7.0 and in refolding experiments 2.0. The transition curve for pH 3.0 is shown in Figure 1 and is situated roughly halfway between the transition regions for pH 2.0 ($T_m \simeq 34$ °C) and pH 7.0 ($T_m \simeq 62$ °C) (B. Nall, unpublished data), providing optimal signal changes for both refolding and unfolding experiments. Optical changes were monitored at 287 nm. Semilog plots of the resultant oscillographs reveal two distinct kinetic phases below T_m (= 43 °C) and three distinct phases above T_m. Data from a typical pH jump, in this case unfolding above T_m, are depicted in Figure 2. Semilog plots of these data demonstrate triphasic kinetics (Figure 3) and from these plots are obtained the time constants (τ's) and relative amplitudes (α's) of the various components. The constants for a series of experiments, both refolding and unfolding, are plotted as a function of temper-

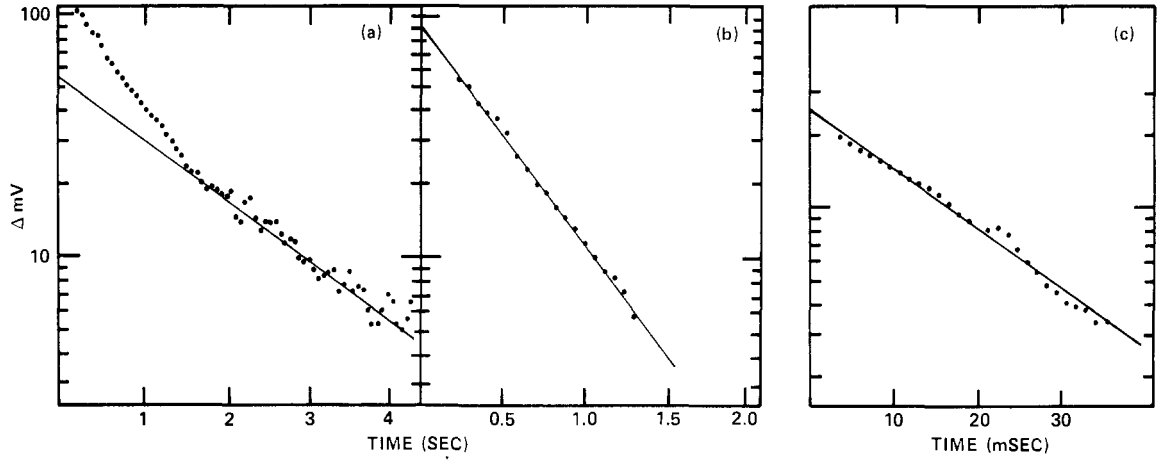


FIGURE 3: Semilog plots of oscillographs shown in Figure 2a,b. (a) Plot of Figure 2a; τ_1 = 1.7 s, obtained by extrapolation of terminal phase to zero time. (b) Replot of the fast component of Figure 3a; τ_2 = 0.48 s. (c) Plot of Figure 2b; τ_3 = 17.7 ms.

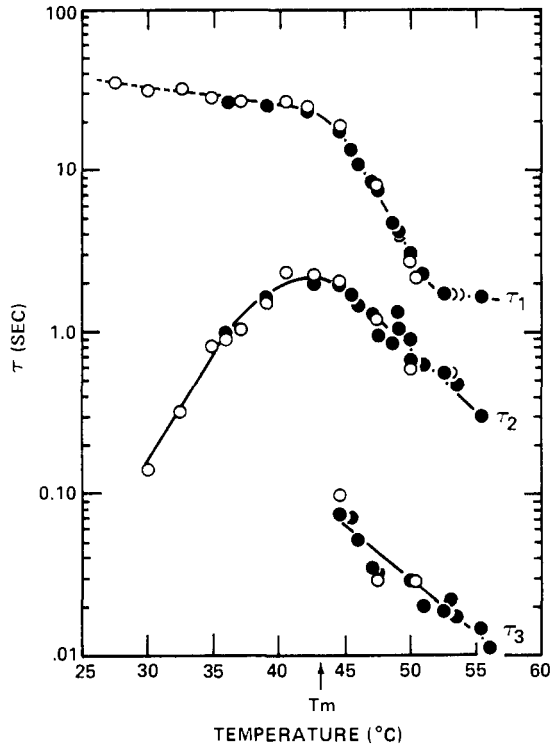


FIGURE 4: Temperature dependence of the kinetic time constants (τ_1, τ_2, and τ_3) for both unfolding (●) and refolding (○) of RNase A in and above the final transition zone (pH 3.0), as determined by pH-jump experiments. For unfolding, pH 7.0 → 3.0 and for refolding, pH 2.0 → 3.0. RNase A concentration: 0.5-1.0 mg/ml. Optical changes followed at λ 287 nm. T_m = transition midpoint.

ature in Figure 4. Amplitude data are similarly plotted in Figure 5a–c. It should be noted that the τ_i values for refolding and unfolding to the same final conditions agree quite closely, thus providing a basic kinetic check on the reversibility of the system.

Measurement of the Slow Phase of Refolding below the Transition Zone at pH 3 by Down Temperature Jump. It is not possible to measure refolding by the pH-jump technique at some low temperatures because RNase A is then native at pH 2.0. The experimental procedure for down temperature-jump measurements is described in the Experimental Section. Protein concentration varies from 1 to 2 mg/ml in glycine buffer. At temperatures below the final transition zone, the slow phase becomes complex, as noted previously in pH-jump studies of refolding at pH 7.0 (Tsong et al., 1972a). Down temperature-jump measurements have also been performed at pH 5.8, employing the inhibitor 2′-CMP binding at 250 nm where there is little absorbance change in RNase A upon refolding (Garel and Baldwin, 1973). Complex kinetics are observed at low temperatures in the slow phase at pH 5.8 as well as at pH 3.0. The fact that complex slow kinetics are observed by 2′-CMP binding suggests that refolding to form the native enzyme occurs in both the early and late parts of the slow kinetic phase.

Prediction of the Equilibrium Transition Curve from Purely Kinetic Parameters. One additional check of the linear three-state model lies in its ability to reproduce successfully the equilibrium transition curve, based solely on kinetic data. The fraction native enzyme (f_N) is defined as the relative (normalized) concentration of species N in the equilibrium distribution:

$$f_N = 1/(1 + K_{32} + K_{32}K_{21}) \quad (13)$$

Since τ_1, τ_2, k_{12}, and k_{21} are known everywhere in the transition zone (see Discussion section), all that remains in expressing $f_N(T)$ is to solve for $K_{23}(T)$. By employing eq 4, along with the identity

$$k_{23} + k_{32} = \tau_1^{-1} + \tau_2^{-1} - k_{12}(1 + K_{21}) \quad (5')$$

one can solve for $K_{23}(T)$ in terms of τ_1 and τ_2. After some algebra, one obtains

$$K_{23} = K_{21}/\beta - 1 \quad (14)$$

where

$$\beta = 1/k_{12}[\tau_1 + \tau_2 - \tau_1\tau_2 k_{12}(1 + K_{21})] - 1 \quad (15)$$

Substituting eq 14 into the expression for f_N (eq 13) gives

$$f_N = (1 - \beta/K_{21})/(1 + \beta) \quad (16)$$

Moreover, expressions 5′ and 14 can be used to obtain k_{23} and k_{32} directly. A comparison of f_N obtained from eq 16 with an experimental transition curve, after correcting for the presence of α_3 (see Appendix, part B), is shown in Figure 6. The agreement is within the combined experimental error of the two experimental methods.

Discussion

(*a*) *Predicted Behavior of* α_2^R *and* α_2^U. Appendix, part

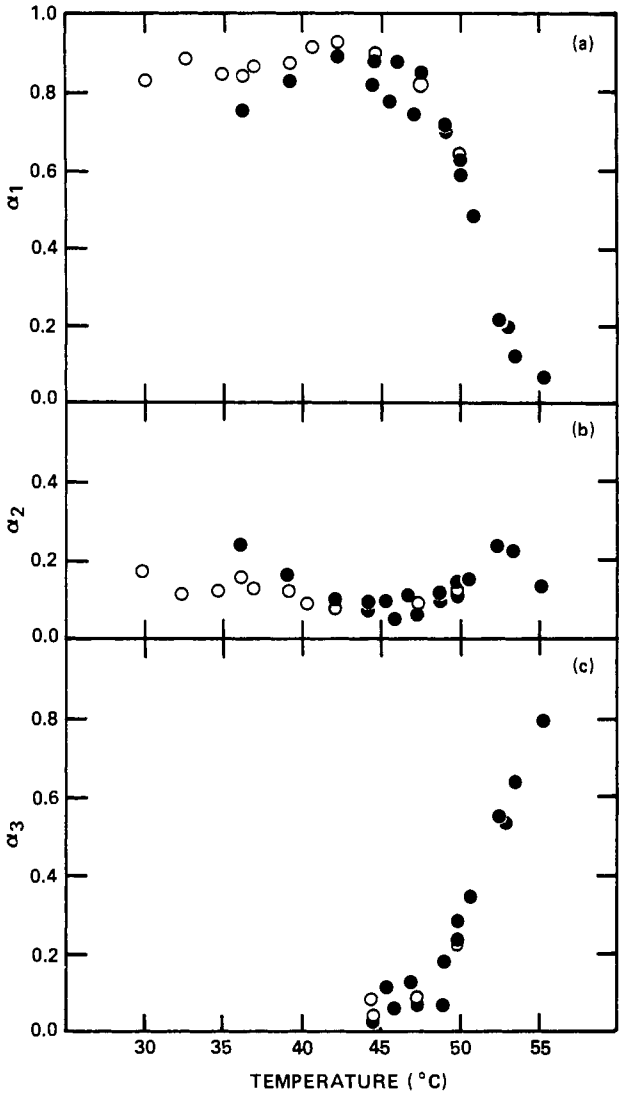


FIGURE 5: Temperature dependence of the amplitudes associated with each time constant in Figure 4. Amplitude subscripts refer to correspondingly numbered τ's.

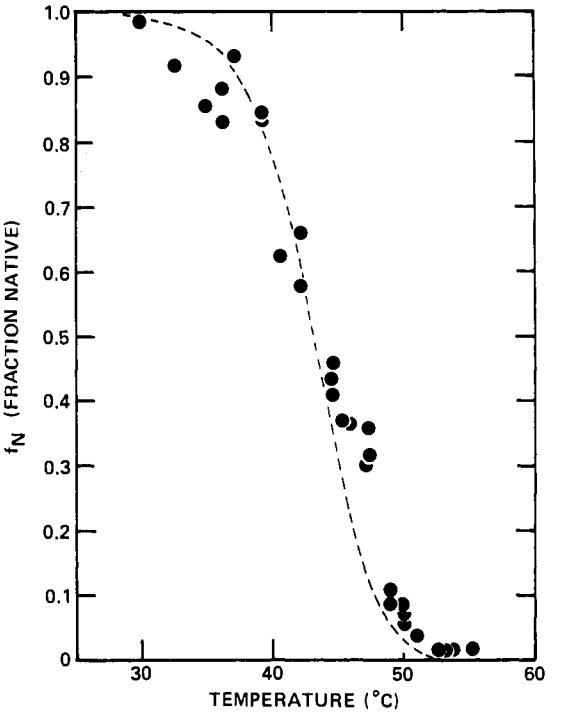


FIGURE 6: Comparison of the measured equilibrium transition curve with the corresponding curve predicted from a purely kinetic analysis. The dashed line represents the experimental transition curve at pH 3.0 (Figure 1) expressed as the normalized fraction of native RNase A (f_N). The closed circles represent the transition curve predicted from kinetic data, corrected above T_m for the presence of a third kinetic component (see text and Appendix, part B).

A, demonstrates analytically the equality $\alpha_2^R = \alpha_2^U$ for the present case and Figure 7a confirms this equality experimentally; thus, a discussion of the predicted behavior of α_2^R in terms of eq 10 is equivalent to that of α_2^U in terms of eq 11. These two equations contain four parameters: (a) τ_1 and τ_2, the measured time constants of the slow and fast kinetic phases; (b) U_2^0, which is obtained from α_2^R in a single refolding measurement (see eq 12); and (c) k_{12}, which equals $(1/\tau_1)$ below the transition zone (see eq 7) but which must be obtained by extrapolation of ln k_{12} vs. $(1/T)$ inside or above the transition zone. Since $K_{12} = k_{12}/k_{21}$ is found to be independent of temperature and can be computed from $K_{12} = U_2^0/(1 - U_2^0)$, k_{21} in eq 11 can be obtained from U_2^0 and k_{12}.

Two properties of τ_1 and τ_2 taken in conjunction with eq 10 and 11 provide a basis for understanding the temperature dependence of α_2 inside the transition zone. First, $(\tau_1 k_{12})$ changes from a value of 1 below the transition zone (eq 7) to a value close to U_2^0 above the transition zone. Equation 7 gives $(\tau_1 k_{12}) = U_2^0$ when $K_{23} = 0$ above the transition zone, but eq 7 applies only approximately in this case because the condition $\tau_1 \gg \tau_2$ is not strictly satisfied. Substituting these values for $(\tau_1 k_{12})$ into eq 10 gives a transition from $\alpha_2 = U_2^0 = 0.2$ below the unfolding zone to $\alpha_2 = 1$ above the unfolding zone. Figure 7b shows the expected behavior of α_2 inside the transition zone when the limiting condition $\tau_1 \gg \tau_2$ is always satisfied: $\alpha_{2,lim}$ increases monotonically and follows a sigmoid curve. The second relevant property of τ_1 and τ_2 is that τ_2 approaches τ_1 near T_m (Figure 4) so that the limiting expressions for τ_1 and τ_2 (eq 7 and 8) do not apply here. The drawing together of τ_2 and τ_1 has an effect on α_2 which is shown diagrammatically at T_m in the lower part of Figure 8. The predicted value of α_2 decreases toward zero as τ_2 approaches τ_1, while the limiting value of α_2 for $\tau_1 \gg \tau_2$ is 0.33 at T_m. The effects of τ_2 approaching τ_1 on the entire curve of α_2 vs. temperature are shown diagrammatically in Figure 7b, where the behavior of $\alpha_{2,lim}$ (for the hypothetical case of $\tau_1 \gg \tau_2$) is contrasted with the observed behavior of α_2.

(*b*) *Comparison of Theory with Experiment.* There are three experimental tests of the analysis given here: (1) the equality of α_2^U and α_2^R, derived in Appendix, part A; (2) the prediction of α_2 as a function of temperature (eq 10 and 11 and Appendix, part B); and (3) the prediction of the equilibrium transition curve from kinetic data (eq 16 and Appendix, part B). At temperatures above T_m, where α_3 contributes significantly, the simple analysis given in eq 1–16 must be modified to take account of a third kinetic phase. This is done in Appendix, part B: the quantity α_2, which is predictable for the three-species mechanism $U_1 \rightleftharpoons U_2 \rightleftharpoons N$, is changed to α_2(pred) for the four-species mechanism $U_1 \rightleftharpoons U_2 \rightleftharpoons I \rightleftharpoons N$ (see Appendix, part B). The experimental quantity corresponding to α_2(pred) is labeled α_2(cor) (Figure 7a). It should be noted that one of the experimental quantities needed in predicting amplitudes, namely k_{12}, must be obtained by extrapolation at tempera-

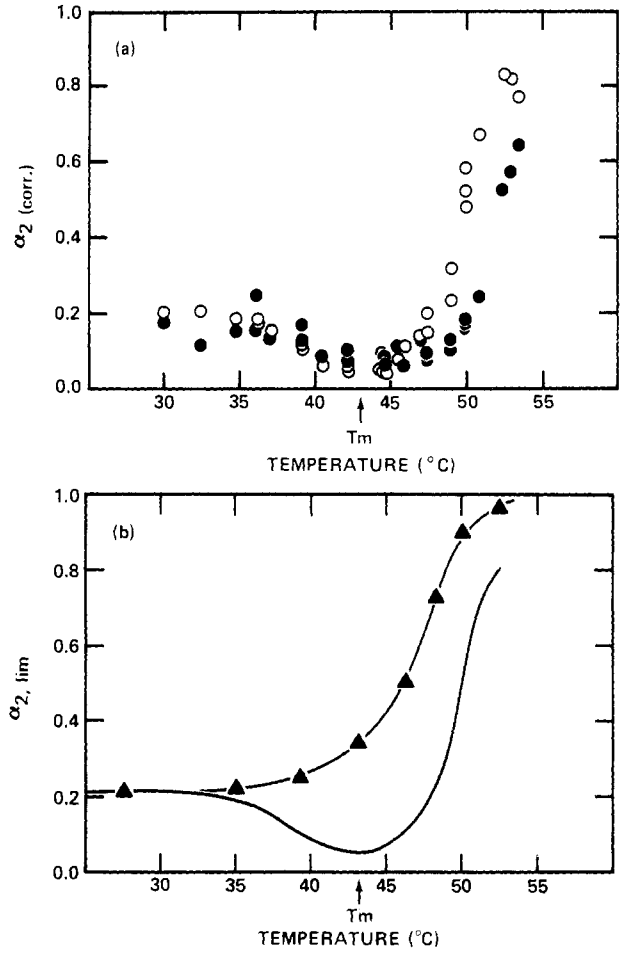


FIGURE 7: (a) Comparison of α_2 predicted from τ data (O) with corresponding experimental values of α_2 (●), corrected to account for the presence of α_3 above T_m (α_3(cor)), where α_2(cor) = α_2(expt)/(1 − α_3(expt)) (see Appendix, part B). (b) Comparison between the experimental temperature dependence of α_2 (solid line, representing an average of unfolding and refolding data) with the temperature dependence expected for the limiting case, $\tau_1/\tau_2 \rightarrow \infty$ ($\alpha_{2,lim}$, eq 12; as ▲).

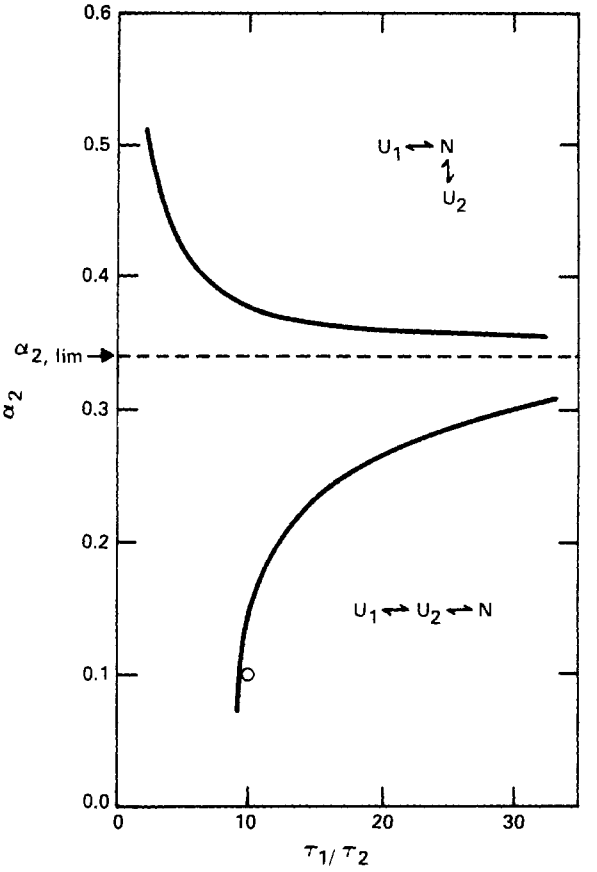


FIGURE 8: Comparison of α_2 predicted by the linear three-state model (model a) with the corresponding value predicted by the branched model (model b), plotted as a function of τ_1/τ_2. In both cases $f_N = 0.5$ in the final conditions, corresponding to a refolding or unfolding jump ending at T_m. For τ_1/τ_2 sufficiently large (>20), α_2 approaches $\alpha_{2,lim}$ for both models and the ability to distinguish between the two models is lost. The open circle (O) represents the measured value of α_2 at T_m. In the current analysis, the faster of the two kinetic phases present at T_m is always referred to as "phase 2" If the $U_2 \rightleftharpoons N$ interconversion were to become much slower than the $U_1 \rightleftharpoons U_2$ process, τ_1/τ_2 would again increase with α_2(linear) → 0, corresponding to conversion between optically identical species, and α_2(branched) → 0.89. These effects will be discussed in more detail elsewhere (Hagerman, in preparation).

tures in and above the transition zone. In principle, it should be possible to determine accurately the temperature dependence of k_{12} from refolding experiments below the transition zone (see eq 7). In practice, this is made difficult by the fact that the slow phase becomes kinetically complex in refolding at low temperatures, and the origin of this complexity is unknown. We have used an activation enthalpy of 13 kcal in extrapolating k_{12} on the basis of the observed slow kinetics, but this may be subject to future revision.

The results of comparing theory with experiment in these three tests are given in Figures 5, 6, and 7a, respectively. Figure 5 shows that $\alpha_i^R = \alpha_i^U$ for all three kinetic phases (i = 1, 2, 3) although data for α_3 are derived almost entirely from unfolding experiments. Figure 7a shows experimental and predicted values of α_2(cor) calculated according to eq A21 of Appendix. The predicted values are based on experimental values of τ_1 and τ_2, which account for their scatter. The experimental values are calculated by an expression (eq A21) which makes severe demands on experimental accuracy when α_3 is large. Thus, the agreement between theory and experiment is satisfactory. Finally, the equilibrium transition curve is predicted satisfactorily from τ_1, τ_2 and U_2^0 by eq 16 at temperatures where α_3 is negligible and by eq A38 when α_3 is significant (Figure 6). The only new experimental quantity needed is α_3.

In addition to these quantitative predictions, the analysis explains the following phenomena: (1) α_1 approaches zero at the top of the transition zone (Figure 5a) as predicted from eq 10 and 7; (2) τ_1 depends on temperature inside and just above the transition zone as predicted by eq 7 (see below); (3) the decrease toward zero of α_2 at temperatures near T_m results from τ_2 and τ_1 drawing together (see Figure 8).

Finally, the analysis serves to rule out the alternative three-species model $U_1 \rightleftharpoons N \rightleftharpoons U_2$, which predicts incorrectly that α_2 at T_m should increase as τ_2 approaches τ_1 (Figure 8).

(*c*) *Interpretation of the Temperature Dependence of* τ_1 *and* τ_2. The variation of τ_1 with temperature is predicted by eq 7 for the limiting case when $\tau_1 \gg \tau_2$ and the three-species mechanism is applicable. According to eq 7, τ_1 depends on the slow $U_1 \rightleftharpoons U_2$ reaction but is coupled inside the transition zone to the fast $U_2 \rightleftharpoons N$ reaction in the following way. Below the transition, τ_1 depends only on the single rate constant k_{12} since $K_{23} = [N^f/U_2^f] \gg 1$; above the transition zone, $K_{23} \rightarrow 0$ and $\tau_1 \rightarrow 1/(k_{12} + k_{21})$. Figure 4 shows that τ_1 does decrease sharply inside the transi-

tion zone, as predicted, and this rapid drop with temperature does appear to level off above the transition zone; however, τ_1 cannot be measured accurately when $\alpha_1 \rightarrow 0$ (Figure 5a). The value of τ_1 above the transition zone is predictable from U_2^0 and from k_{12} as a function of temperature. Unfortunately the latter is not known accurately (see above) and the activation enthalpy of 13 kcal for k_{12} has been chosen in part to ensure agreement of the predicted and observed values of τ_1 above the transition zone.

The variation of τ_2 with temperature is very interesting: the decrease in rate ($= 1/\tau_2$) with temperature below T_m means that the fast-refolding reaction $U_2 \rightleftharpoons N$ is actually a complex set of reactions. The behavior of the rate, which passes through a minimum at T_m, bears a striking similarity to the behavior of the α helix (cf. Schwarz, 1965) and to the polynucleotide double helix (cf. Ross and Sturtevant, 1960) and, therefore, is probably a nucleation-dependent reaction. At least two other small proteins (metmyoglobin, Shen and Hermans, 1972; ferricytochrome *c*, Ikai et al., 1973) show a fast refolding reaction whose rate also drops to a minimum at the midpoint of the transition. The fact that τ_2 approaches τ_1 near T_m is useful in the present analysis, because it allows us to decide between the alternative three-species mechanisms: $U_1 \rightleftharpoons N \rightleftharpoons U_2$ and $U_1 \rightleftharpoons U_2 \rightleftharpoons N$ (Figure 8).

(*d*) *Interpretation of the Temperature Dependence of* α_3. The intermediate I, which is responsible for τ_3, must be produced directly from N because τ_3 is the fastest process when the pH-jump unfolding starts from a pH at which $f_N = 1$. It is important to find out if the properties of I more closely resemble those of N or of U_2 and U_1. As regards its extinction coefficient at 287 nm, which measures the exposure to solvent of buried tyrosine groups, I must be nearly equivalent to U_1 and U_2 because the optical change associated with τ_3 approaches the total change for the entire unfolding process (Figure 5c). Also, the temperature dependence of α_3 is predicted satisfactorily (Figure 9) by using the mechanism $U_1 \rightleftharpoons U_2 \rightleftharpoons I \rightleftharpoons N$ and by assuming that the equilibrium ratio U_1-U_2-I is 0.78:0.20:0.02, independent of temperature. We cannot rule out the alternative mechanism $U_1 \rightleftharpoons U_2 \rightleftharpoons N \rightleftharpoons I$ because τ_2 and τ_3 are always well separated, and the test which differentiates mechanisms (a) and (b) (Figure 8) cannot be used here.

(*e*) *Possible Structural Interpretations of the* $U_1 \rightleftharpoons U_2 \rightleftharpoons I \rightleftharpoons N$ *Mechanism.* The present work defines two intermediates in the folding transition of RNase A, one of which (U_2) certainly, and the other of which (I) probably, has the properties of an unfolded species even to the extent of being populated at equilibrium above the transition zone for unfolding at low pH. This is a surprising result, which was not anticipated either in our early studies of the RNase A transition (Tsong et al., 1971, 1972), or in studies of the guanidine-induced unfolding of several small proteins by Tanford and co-workers (Ikai et al., 1973; Tanford et al., 1973). It remains to be seen whether U_2 and I persist in concentrated guanidine solutions, and whether similar stable intermediates are present after the guanidine-induced unfolding of lysozyme and ferricytochrome *c*. If stable intermediates like U_2 exist in concentrated guanidine solutions, then application of the kinetic analysis of Ikai and Tanford (1971, 1973) to lysozyme and ferricytochrome *c* is based on the incorrect assumption that there is only one species of the unfolded protein.

There are two consequences of the $U_1 \rightleftharpoons U_2 \rightleftharpoons I \rightleftharpoons N$ mechanism that demand a molecular explanation. First, the unfolding reaction, $N \rightleftharpoons I$, which can occur in a time as short as 10 ms (Figure 4), evidently exposes the interior of the protein to solvent. What is the molecular nature of this reaction, why is it so fast, and why are no partly native intermediates observed? Second, the $U_1 \rightleftharpoons U_2$ reaction and probably also the $U_2 \rightleftharpoons I$ reaction are comparatively slow reactions that take place after unfolding is nearly or entirely complete. What are these reactions, and why are they so slow? On the basis of studies of synthetic linear polymers, one would expect that conformational interconversions of an unfolded polypeptide chain should be fast reactions.

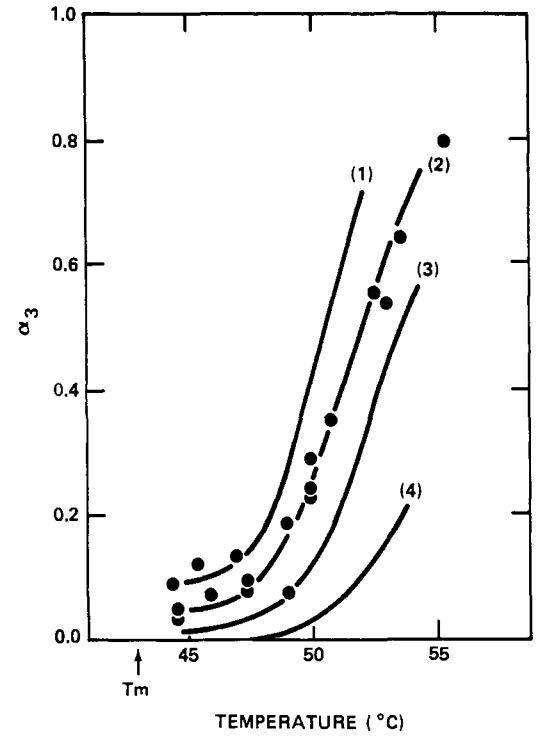


FIGURE 9: Determination of the equilibrium ratio I^0/U_2^0 from the temperature dependence of α_3. Closed circles (●) represent the experimental values of α_3 as a function of temperature. Solid lines represent computed values of α_3 for various constant values of I^0/U_2^0. I^0/U_2^0 = 0.2 (curve 1); 0.1 (curve 2); 0.04 (curve 3); 0.01 (curve 4). It can be seen from the figure that the ratio I^0/U_2^0 varies by less than a factor of 2 over a temperature range where the U_2^0/N^0 ratio varies by 1000.

Regarding the nature of U_2 and I, the first question is: does the formation of U_2 from U_1 (and of I from U_2) represent the formation of a nucleus of structure (e.g., held together by H bonds) such that successive steps in folding can occur rapidly, or does it represent a change in topology (i.e., configuration or conformation) of the unfolded polypeptide chain? The evidence is fairly strong that the latter is correct. It is difficult to imagine how an unstable nucleus of H-bonded structure could remain intact at temperatures above the unfolding transition zone. In fact, the $U_1 \rightleftharpoons U_2$ equilibrium at pH 2 appears to be independent of temperature (Garel and Baldwin, 1975a; cf. also Brandts et al., 1975). It does depend on pH, but this dependence is opposite to the behavior expected if U_2 contains a nucleus of structure that is stabilized at neutral pH. The only physical property that has been found thus far to differ between U_1 and U_2 is the average pK of the three freely ionizing nitrotyrosyl groups on the surface of nitrotyrosyl RNase A. The difference in pK between U_1 and U_2 is small: about 0.05 pH unit (Garel and Baldwin, 1975b).

Brandts et al. (1975) have proposed a plausible explana-

tion for the $U_1 \rightleftharpoons U_2$ reaction of RNase A. They show data for proline-containing dipeptides in which the rate of the cis–trans isomerization of the prolyl residue is slow (τ varies from 20 to 100 s at 25 °C) and the activation enthalpy is large (19 kcal). The ratio of the cis–trans forms is about 20:80 in some model compounds. They suggest that in U_2 each prolyl residue has the same cis or trans conformation as in N, while in U_1 one or more of the four prolyl residues of RNase A has an incorrect conformation and cannot fold up until the cis $\rightleftharpoons$ trans isomerization is completed. This hypothesis can explain the rate and the apparent activation enthalpy (which is not yet known accurately) of the $U_1 \rightleftharpoons U_2$ interconversion, but it cannot explain the pH dependence (Garel and Baldwin, 1975a) without some extension of the hypothesis involving the formation of residual structure in unfolded RNase A at neutral pH. The rate of the $U_2 \rightleftharpoons I$ interconversion appears to be much too fast to be the cis–trans isomerization of a prolyl residue, according to the data of Brandts et al., (1975), and the activation enthalpy of this reaction is unknown. An alternative explanation for the $U_1 \rightleftharpoons U_2$ and $U_2 \rightleftharpoons I$ reactions is that certain topologically difficult changes in overall chain configuration are involved, perhaps involving the -SS- bonded loops.

(*f*) *Comparison with Previous Kinetic Studies.* Scott and Scheraga (1963) published the first data on the kinetics of RNase A unfolding; they made stopped-flow, pH-jump measurements of unfolding as a function of temperature in the pH range 0.9–3.3. They did not report data taken in time ranges faster than seconds, and their results cannot be compared with the present measurements of τ_2 and τ_3. They found two steps in the time range where τ_1 is now observed. Although some evidence for complexity in the slow phase has been observed at low temperatures of unfolding (Tsong et al., 1972b) as well as of refolding (Tsong et al., 1972a), the pattern of behavior observed by Scott and Scheraga has not been reproduced either by these authors or by Pohl (1969), or by Summers and McPhie (1972), or in our present work. The reason is unknown. Pohl (1969) made measurements of thermal unfolding and refolding at pH 2 by a "slow temperature-jump" technique. He did not record data in time ranges faster than seconds and, instead of reporting τ_1, he gives derived values for the rates of unfolding and refolding that are obtained by use of the two-state approximation. Summers and McPhie measured α_1 as a function of temperature in stopped-flow measurements of unfolding at pH 1.13. They were chiefly concerned with pointing out the similarity in behavior between RNase A and metmyoglobin, both of which show a rapid decrease in α_1 with increasing temperature (cf. Figure 5a). Our present results correspond closely to those of Tsong et al. (1972a), except that they missed the triphasic kinetics at the high-temperature end of the transition zone and consequently misassigned τ_1 and τ_2 in and above this temperature range. The triphasic kinetics are easy to miss unless one is looking for them, because α_1 is approaching zero in this temperature range (Figure 5a).

Three models are reviewed below which, in addition to the present model, have been used as starting points for the interpretation of the kinetics of protein folding transitions. These are: (1) the two-state model in which all intermediates are negligible, which has been the standard model used in early studies of protein folding transitions (Eisenberg and Schwert, 1951; see also the early equilibrium studies of Northrop, 1930); (2) a three-state model in which an abortive or dead-end intermediate is formed rapidly after the initiation of refolding (Ikai and Tanford, 1971, 1973); and (3) a multi-state model in which folding begins by nucleation followed by a series of equal propagation steps (Tsong et al., 1972b). The two-state model: the argument is sometimes made that application of the two-state approximation to the slow phase of refolding or unfolding should provide a good first-approximation basis for interpreting the kinetics even when the kinetics demonstrate that intermediates are present, because the slow phase shows the major amplitude change, at least below T_m. The present work indicates that such an interpretation may be quite misleading. For example, on the basis of such a two-state analysis, one would view RNase A refolding as a rather slow process in the time range of 10–100 s. However, our present results show that the slow reaction in refolding represents an interconversion between two unfolded species, U_1 and U_2, both of which are present before refolding is initiated, whereas the actual folding process ($U_2 \rightleftharpoons N$) is a fast reaction (Figure 4). The three-state model of Ikai and Tanford (1971, 1973) has been applied to guanidine-induced transitions (Ikai et al., 1973; Tanford et al., 1973). This model cannot explain the present results for RNase A because it predicts that, starting with unfolded enzyme, the product of the fast refolding reaction should be an abortive intermediate, whereas it actually is the native enzyme (Garel and Baldwin, 1973).

The nucleation-dependent, sequential model of Tsong et al. (1972b) at first appeared to explain several properties of the unfolding kinetics of RNase A (Tsong et al., 1972a). However, it did not explain the observation of a fast refolding reaction, starting with unfolded enzyme. This led Garel and Baldwin (1973) to characterize the product of the fast refolding reaction and to the discovery that it is native enzyme, which in turn led to the present $U_1 \rightleftharpoons U_2 \rightleftharpoons N$ mechanism.

(*g*) *Comparison with Calorimetric Results.* Highly precise measurements of protein-unfolding transitions have been made with the aim of learning whether intermediates are detectable at equilibrium: see the review by Privalov (1974). The thermal folding transition of RNase A has been studied carefully by Tsong et al., (1971) and by Tiktopulo and Privalov (1974). The latter authors divide the unfolding curve measured calorimetrically into two parts: a noncooperative, premelting zone and a cooperative, melting zone. No unfolding kinetics have been found in the premelting zone by the fast temperature-jump technique (Tsong et al., 1971); the absorbance of the buried tyrosine groups was monitored at 286 nm. Calorimetric studies indicate that the melting process itself is highly cooperative: $\Delta H_{cal}/\Delta H_{vH}$ is quoted as 1.05 ± 0.03 (Tiktopulo and Privalov, 1974) and this ratio should be greater than one if there are detectable intermediates. However, to be detected, any intermediates must differ in partial molar enthalpy from the two end states, native and unfolded protein. The $U_1 \rightleftharpoons U_2$ reaction of RNase A should not be measurable by this technique, because the equilibrium is independent of temperature (Garel and Baldwin, 1975a).

Acknowledgments

We are grateful for the advice and discussions of Drs. E. L. Elson and J.-R. Garel and of B. T. Nall and J. A. Ridge.

Appendix

Part A: Three-State Amplitude Analysis

Amplitude Expressions for Refolding and Unfolding Ki-

netics. For a restricted set of kinetic models, exact analytic expressions for individual concentrations ($c_i(t)$) and time constants (τ_j) can be given in terms of initial and final elementary rate constants (it is equivalent to express initial conditions in terms of concentrations c_i^0). Such expressions are presented for the most general (cyclic) three-species system by Szabó (1969; see also Ikai, 1971). The nine coefficients (C_{ij}) of eq 1 can be obtained directly for the linear three-state model by expanding the expressions given by Szabó (with $k_{31}, k_{13} = 0$), and the resulting expressions are listed below ($\tau_i = \lambda_i^{-1}$):

$$C_{10} = (c_1^0 + c_2^0 + c_3^0)k_{21}k_{32}/\lambda_1\lambda_2 \quad \text{(A1)}$$

$$C_{20} = (c_1^0 + c_2^0 + c_3^0)k_{12}k_{32}/\lambda_1\lambda_2 \quad \text{(A2)}$$

$$C_{30} = (c_1^0 + c_2^0 + c_3^0)k_{12}k_{23}/\lambda_1\lambda_2 \quad \text{(A3)}$$

$$C_{12} = [1/\lambda_2(\lambda_2 - \lambda_1)][c_1^0k_{12}(\lambda_2 - k_{23} - k_{32}) - c_2^0k_{21}(\lambda_2 - k_{32}) + c_3^0k_{32}k_{21}] \quad \text{(A4)}$$

$$C_{22} = [1/\lambda_2(\lambda_2 - \lambda_1)]\{-c_1^0k_{12}(\lambda_2 - k_{32}) + c_2^0[k_{21}(\lambda_2 - k_{32}) + k_{23}(\lambda_2 - k_{12})] - c_3^0k_{32}(\lambda_2 - k_{12})\} \quad \text{(A5)}$$

$$C_{32} = [1/\lambda_2(\lambda_2 - \lambda_1)][c_1^0k_{12}k_{23} - c_2^0k_{23}(\lambda_2 - k_{12}) + c_3^0k_{32}(\lambda_2 - k_{12} - k_{21})] \quad \text{(A6)}$$

$$C_{11} = [1/\lambda_1(\lambda_2 - \lambda_1)][-c_1^0k_{12}(\lambda_1 - k_{23} - k_{32}) + c_2^0k_{21}(\lambda_1 - k_{32}) - c_3^0k_{32}k_{21}] \quad \text{(A7)}$$

$$C_{21} = [1/\lambda_1(\lambda_2 - \lambda_1)]\{c_1^0k_{12}(\lambda_1 - k_{32}) - c_2^0[k_{21}(\lambda_1 - k_{32}) + k_{23}(\lambda_1 - k_{12})] + c_3^0k_{32}(\lambda_1 - k_{12})\} \quad \text{(A8)}$$

$$C_{31} = [1/\lambda_1(\lambda_2 - \lambda_1)][-c_1^0k_{12}k_{23} + c_2^0k_{23}(\lambda_1 - k_{12}) - c_3^0k_{32}(\lambda_1 - k_{12} - k_{21})] \quad \text{(A9)}$$

As discussed in the theoretical section, these C_{ij} can be employed to predict the relative amplitude associated with each τ_i, solely on the basis of the k_{ij} and c_i^0. Considering the amplitude associated with the faster of the two phases (α_2), an expression can be given directly in terms of the C_{ij} as

$$\alpha_2^{U,R} = 1/(1 + |C_{31}^{U,R}|/|C_{32}^{U,R}|) \quad \text{(A10)}$$

where in the following analysis, U superscript specifies c_1^0, $c_2^0 = 0$, and R superscript specifies $c_3^0 = 0$. For unfolding

$$C_{31}^U/C_{32}^U = (\lambda_2/\lambda_1)[-(\lambda_1 - k_{12} - k_{21})/(\lambda_2 - k_{12} - k_{21})] \quad \text{(A11)}$$

and

$$\alpha_2^U = 1/[1 + |1 - \tau_1(k_{12} + k_{21})|/|1 - \tau_2(k_{12} + k_{21})|] \quad \text{(A12)}$$

Similarly for refolding

$$C_{31}^R/C_{32}^R = (\lambda_2/\lambda_1)\{[-c_1^0k_{12} + c_2^0(\lambda_1 - k_{12})]/[c_1^0k_{12} - c_2^0(\lambda_2 - k_{12})]\} \quad \text{(A13)}$$

and

$$\alpha_2^R = 1/(1 + |\tau_1k_{12} - c_2^0|/|\tau_2k_{12} - c_2^0|) \quad \text{(A14)}$$

where $c_1^0 + c_2^0 = 1$. It should be noted that eq A12 and A14 represent the magnitudes associated with each kinetic component, the relative signs of these amplitudes being given by the sign of the ratio C_{31}/C_{32}. The significance of $C_{31}/C_{32} < 0$ is considered in detail elsewhere (Hagerman, in preparation).

Proof of the Equality $\alpha_2^R = \alpha_2^U$ for RNase A Transition Kinetics at Final pH 3. The condition of equality is satisfied under conditions where $K_{12} \not\propto f(\text{pH})$ in the pH range of the refolding experiments, as will be shown below. It has been demonstrated previously that this condition is satisfied in the pH range 2–3. In verifying the equality $\alpha_2^R = \alpha_2^U$, it is sufficient to show that

$$|C_{31}^R|/|C_{32}^R| = |C_{31}^U|/|C_{32}^U| \quad \text{(A15)}$$

We use the relation $c_2^0 = k_{12}/(k_{12} + k_{21})$, which is valid when K_{12} has the same value in the initial conditions as in the final conditions of a refolding experiment. From eq A13

$$\begin{aligned}(\lambda_1/\lambda_2)|C_{31}^R|/|C_{32}^R| &= |k_{12} - c_2^0\lambda_1|/|k_{12} - c_2^0\lambda_2| \\ &= |k_{12} - \lambda_1k_{12}/(k_{12} + k_{21})|/|k_{12} - \lambda_2k_{12}/(k_{12} + k_{21})| \\ &= |\lambda_1 - k_{12} - k_{21}|/|\lambda_2 - k_{12} - k_{21}| \\ &= (\lambda_1/\lambda_2)|C_{31}^U|/|C_{32}^U|\end{aligned} \quad \text{(A16)}$$

Limiting Form of α_2 as $\tau_1/\tau_2 \to \infty$. As $\tau_1/\tau_2 \to \infty$

$$\alpha_2 \to \alpha_{2,\text{lim}} = 1/[1 + |1 - \tau_1(k_{12} + k_{21})|] \quad \text{(A17)}$$

since $\tau_2(k_{12} + k_{21}) \to 0$. Furthermore, since $\tau_1^{-1} \leq (k_{12} + k_{21})$

$$\begin{aligned}\alpha_{2,\text{lim}} &= 1/\tau_1(k_{12} + k_{21}) \\ &= [k_{12}/(k_{12} + k_{21})][1 + K_{21}/(1 + K_{23})] \\ &= c_2^0[1 + K_{21}/(1 + K_{23})]\end{aligned} \quad \text{(A18)}$$

Part B: Corrections to the Three-State Model to Account for the Appearance of a Third Kinetic Phase

Comparison of Experimental and Predicted Amplitudes. In comparing experimental amplitudes with those derived from theory (using τ_i data) under conditions in which the kinetic system displays three-state behavior ($\alpha_3 \simeq 0$), the predicted amplitudes, here referred to as $\alpha_i(\text{pred})$, correspond directly to their experimental counterparts:

$$\alpha_1(\text{pred}) = \alpha_1(\text{expt}) \quad \text{(A19)}$$

and

$$\alpha_2(\text{pred}) = \alpha_2(\text{expt}) \quad \text{(A20)}$$

Above T_m, where $\alpha_3(\text{expt}) > 0$, the three-state model must be regarded as an approximation to a more general model having at least four species. In comparing predicted and experimental amplitudes in this region, the correspondence

$$\alpha_2(\text{pred}) = \alpha_2(\text{expt})/(1 - \alpha_3(\text{expt})) \quad \text{(A21)}$$

was used for the present system ($\tau_3 \ll \tau_1, \tau_2$); its correctness will be demonstrated below. For the situation where τ_3 is not widely separated from τ_1 and τ_2, such a correspondence is no longer valid and the four-state system must be solved exactly.

Proof of the Correspondence $\alpha_2(pred) = \alpha_2(expt)/(1 - \alpha_3(expt))$ for the Case Where $\tau_3 \ll \tau_1, \tau_2$. For widely separated τ's, the amplitudes predicted from the exact four-state kinetic treatment should be equivalent to corresponding amplitudes predicted by stepwise equilibration. Furthermore, for a system that is totally separable ($\tau_1 \gg \tau_2 \gg \tau_3$), the experimental amplitudes should be given exactly by the corresponding stepwise equilibration (SE) expressions; thus

$$\alpha_{i,\text{lim}}^{4\text{-state}}(\text{pred}) = \alpha_i^{4\text{-state}}(\text{SE}) = \alpha_i(\text{expt}) \quad \text{(A22)}$$

For unfolding from completely native protein ($f_N^0 = 1$)

$$\alpha_1^{4\text{-state}}(\text{SE}) = K_{43}K_{32}K_{21}/(1 - f_N)(1 + K_{43} + K_{43}K_{32}) \times (1 + K_{43} + K_{43}K_{32} + K_{43}K_{32}K_{21}) \quad \text{(A23)}$$

$$\alpha_2^{4\text{-state}}(\text{SE}) = K_{43}K_{32}/(1 - f_N)(1 + K_{43})(1 + K_{43} + K_{43}K_{32}) \quad \text{(A24)}$$

and

$$\alpha_3^{4\text{-state}}(\text{SE}) = K_{43}/(1 - f_N)(1 + K_{43}) \quad \text{(A25)}$$

The limiting form of the three-species expression for $\alpha_2^{3\text{-state}}$(pred) (eq A18) is

$$\alpha_{2,\text{lim}}^{3\text{-state}}(\text{pred}) = 1/\tau_1(k_{12} + k_{21}) \quad \text{(A26)}$$

For a linear four-state model having I between U_2 and N

$$\tau_{1,\text{lim}}^{-1} = k_{12} + k_{21}/(1 + K_{23} + K_{23}K_{34}) \quad \text{(A27)}$$

and

$$\alpha_{2,\text{lim}}^{3\text{-state}}(\text{pred}) = (K_{21}/\gamma + 1)/(1 + K_{21}) \quad \text{(A28)}$$

where

$$\gamma = 1 + K_{23} + K_{23}K_{34} \quad \text{(A29)}$$

Therefore

$$\alpha_{2,\text{lim}}^{3\text{-state}}(\text{pred})(1 - \alpha_3(\text{SE})) = K_{43}K_{32}/(1 - f_N)(1 + K_{43})(1 + K_{43} + K_{43}K_{32}) = \alpha_{2,\text{lim}}^{4\text{-state}}(\text{pred}) = \alpha_2(\text{expt}) \quad \text{(A30)}$$

As τ_1 and τ_2 draw together, $\alpha_1^{3\text{-state}}(\text{pred})/\alpha_2^{3\text{-state}}(\text{pred})$ will change according to the equations given above for the three-state model. Nevertheless, the correction factor $(1 - \alpha_3(\text{expt}))$ still applies since, for $\tau_3 \ll \tau_1, \tau_2$, the relative amplitude $(\alpha_3(\text{expt}))$ only depends on the equilibrium ratio K_{34}. Physically, the correction factor refers to the fraction of the reaction that has already been completed by the time species U_1 and U_2 become populated. A linear four-state model with I between U_2 and N has been employed in the foregoing discussion, although a similar analysis with I as an off-pathway species yields the same result.

Comparison of f_N Obtained Directly from Equilibrium Data with f_N Predicted from Kinetic Data. The observation of an additional kinetic species (I) in unfolding experiments above T_m requires the correction of f_N predicted from τ_1 and τ_2, to account for the additional kinetic complexity. For the direct pathway, four-state model, the equilibrium fraction native is given by

$$f_N = 1/(1 + K_{43} + K_{43}K_{32} + K_{43}K_{32}K_{21}) \quad \text{(A31)}$$

In general, in predicting f_N from kinetic data, the four-state system must be solved exactly for all six k_{ij}, although for the RNase system certain approximations may be utilized since $k_{34}, k_{43} \gg k_{12}, k_{21}, k_{23}, k_{32}$; namely

$$\tau_{1(2)}^{-1} = \{\textstyle\sum k' - (+)\ [(\sum k')^2 - 4\prod k']^{1/2}\}/2 \quad \text{(A32)}$$

and

$$\tau_3^{-1} = k_{34} + k_{43} \quad \text{(A33)}$$

where

$$\textstyle\sum k' = k_{12} + k_{21} + k_{23} + k_{32}' \quad \text{(A34)}$$

$$\textstyle\prod k' = k_{12}(k_{23} + k_{32}') + k_{21}k_{32}' \quad \text{(A35)}$$

and

$$k_{32}' = k_{32}/(1 + K_{34}) \quad \text{(A36)}$$

All four rate constants (k_{12}, k_{21}, k_{23}, and k_{32}') can be obtained directly from τ_1^{-1} and τ_2^{-1} since K_{12} is known from $\alpha_{2,\text{lim}}$ (eq A18) and $\tau_1^{-1} \rightarrow k_{12}$ below the final transition zone (eq 7). Thus the magnitude of the error can be evaluated when the three-state approximation is used to express f_N in terms of kinetic data:

$$f_N^{3\text{-state}}(\text{kinetic}) = 1/(1 + K_{32}' + K_{32}'K_{21}) = (1 + K_{43})/(1 + K_{43} + K_{43}K_{32} + K_{43}K_{32}K_{21}) = (1 + K_{43})f_N(\text{equil}) \quad \text{(A37)}$$

Eq A25 can then be used to provide an explicit correction to $f_N^{3\text{-state}}$(kinetic) in terms of the experimental quantity α_3(expt):

$$f_N^{\text{cor}}(\text{kinetic}) = (1 - \alpha_3(\text{expt}))f_N^{3\text{-state}}(\text{kinetic})/(1 - \alpha_3(\text{expt})f_N^{3\text{-state}}(\text{kinetic})) \quad \text{(A38)}$$

An identical expression is obtained for the case where I is an off-pathway intermediate. The main point to be made here is that, under conditions where additional kinetic intermediates became evident (i.e., finite kinetic amplitudes) within the transition zone, consideration must be made of these additional species in predicting equilibrium properties of the system. The fact that a particular species is not significantly populated at equilibrium is not necessarily sufficient to exclude its importance in comparing kinetically determined equilibrium constants with those obtained by direct measurement. The foregoing analysis can be extended to more complicated systems where it holds that only components having finite amplitudes within the transition region need be considered in comparing kinetic and equilibrium results, even if additional kinetic components are observed above and below the transition zone.

References

Beaven, G. H., Holiday, E. R., and Johnson, E. H. (1955), in The Nucleic Acids, Vol. 1, Chargaff, E., and Davidson, J. N., Ed., New York, N.Y., Academic Press, p 513.

Brandts, J. F., Halvorson, H. R., and Brennan, M. (1975), *Biochemistry 14,* 4953.

Crestfield, A. M., Stein, W. H., and Moore, S. (1963), *J. Biol. Chem. 238,* 618.

Eigen, M., and DeMaeyer, L. (1963), in Technique of Organic Chemistry; Vol. 8, Part 2, Friess, S. L., Lewis, E. S., and Weissberger, A., Ed., New York, N.Y., Wiley-Interscience, p 895.

Garel, J.-R., and Baldwin, R. L. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 3347.

Garel, J.-R., and Baldwin, R. L. (1975a), *J. Mol. Biol. 94,* 611.

Garel, J.-R., and Baldwin, R. L. (1975b), *J. Mol. Biol. 94,* 621.

Gibson, Q. H. (1964), in Rapid Mixing and Sampling Techniques in Biochemistry, Chance, B., Gibson, Q. H., Eisenhardt, R. H., and Lonberg-Holm, K. K., Ed., New York, N.Y., Academic Press, p 115.

Ikai, A. (1971), Ph.D. Dissertation, Duke University.

Ikai, A., Fish, W. F., and Tanford, C. (1973), *J. Mol. Biol. 73,* 167.

Ikai, A., and Tanford, C. (1971), *Nature (London) 230,* 100.

Ikai, A., and Tanford, C. (1973), *J. Mol. Biol. 73,* 145.

Pohl, F. M. (1969), *FEBS Lett. 3,* 60.

Privalov, P. L. (1974), *FEBS Lett. 40,* S140.

Ross, P. D., and Sturtevant, J. M. (1960), *Proc. Natl. Acad. Sci. U.S.A. 46,* 1360.

Schwarz, G. (1965), *J. Mol. Biol. 11,* 64.

Scott, R. A., and Scheraga, H. A. (1963), *J. Am. Chem.*

Soc. 85, 3866.
Sela, M., and Anfinsen, C. B. (1957), *Biochim. Biophys. Acta 24,* 229.
Shen, L. L., and Hermans, J., Jr. (1972), *Biochemistry 11,* 1836.
Summers, M. R., and McPhie, P. (1972), *Biochem. Biophys. Res. Commun. 47,* 831.
Szabó, Z. G. (1969), in Comprehensive Chemical Kinetics, Bamford, C. H., and Tipper, C. F. H., Ed., Amsterdam, Elsevier, p 1.
Tanford, C., Aune, K. C., and Ikai, A. (1973), *J. Mol. Biol. 73,* 185.
Tiktopulo, E. I., and Privalov, P. L. (1974), *Biophys. Chem. 1,* 349.
Tsong, T. Y., Baldwin, R. L., and Elson, E. L. (1971), *Proc. Natl. Acad. Sci. U.S.A. 68,* 2712.
Tsong, T. Y., Baldwin, R. L., and Elson, E. L. (1972a), *Proc. Natl. Acad. Sci. U.S.A. 69,* 1809.
Tsong, T. Y., Baldwin, R. L., and McPhie, P. (1972b), *J. Mol. Biol. 63,* 453.

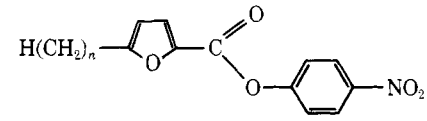
Rate Enhancement Specificity with α-Chymotrypsin: Temperature Dependence of Deacylation†

Joseph E. Baggott and Michael H. Klapper*

ABSTRACT: The relative rate of the hydrolysis of 2-(5-*n*-alkyl)furoyl-α-chymotrypsin reaches a maximum with the propyl derivative. The Arrhenius plots for the hydrolyses of the 2-furoyl-, 2-(5-ethyl)furoyl-, and 2-(5-*n*-propyl)furoyl-α-chymotrypsins display a discontinuity, while the plots obtained with the remaining furoyl derivatives 5-methyl, 5-*n*-butyl, and 5-*n*-amyl are linear. We conclude that the deacylation of the furoyl derivatives of α-chymotrypsin involves a minimum of two elementary reaction steps. Depending upon the reaction conditions, rate enhancement specificity appears to be either entropy or enthalpy controlled.

The protein α-chymotrypsin is useful for testing theories of rate enhancement, because it is one of the best understood enzymes. The crystal structure of this protein and of various derivatives has been determined (Mavridis et al., 1974; Blow and Steitz, 1970). The reaction pathway has been sufficiently described (Bender and Kezdy, 1965) for a reasonable chance of success in attempting the assignment of activation parameters to elementary reaction steps. The study of homologous substrate series allows the isolation of various factors controlling reaction rates, and a large number of series have been examined (e.g., Dupaix et al., 1973; Fife and Milstien, 1967; Marshall and Akgün, 1971). Therefore, a reasonable explanation of rate enhancement specificity (the variation of enzymatic turnover with different substrates) should be possible with this enzyme. But it is precisely here that a serious discrepancy is found in the literature. The reaction of esters and amides with chymotrypsin occurs through esterification of serine 195 on the enzyme, followed by hydrolysis of the acyl enzyme intermediate liberating free enzyme and acid. The turnover of ester substrates is generally limited by the latter, deacylation step. As a result, the deacylation rate constant can be determined with ester substrates either from steady-state kinetics, or from direct observation of the acyl enzyme decay. Bender and co-workers (1964), Kaplan and Laidler (1966), and Fife and Milstien (1967; with the straight-chain fatty acid series) have reported that rate enhancement specificity is entropy controlled—i.e., faster turnover rates between substrates are a reflection of more positive activation entropies—and that $\Delta H^{\ddagger}$ for different substrates remains relatively constant. Fife and Milstien (1967) simultaneously studied branched chain fatty acid derivatives, with which they observed compensatory changes between activation entropies and enthalpies, but enthalpy control of specificity; i.e., negative changes in both activation enthalpy and entropy accompany faster rates. Martinek et al. (1972) and Marshall and Chen (1973) also studied the straight-chain fatty acids and reported compensatory changes with activation enthalpy control of specificity. Finally, Cane and Wetlaufer (1966) studying the same straight-chain series reported compensatory changes with entropy control; i.e., positive changes in both activation enthalpy and entropy accompany faster rates.

Any successful predictive theory of rate enhancement specificity should include, or be based on, the calculation of activation enthalpies and entropies, and, thus, the resolution of this apparent discrepancy is important before a meaningful theory of chymotrypsin rate enhancement can be formulated. Since the different conclusions cited above may have arisen from the experimental difficulties inherent in accurate measurement of enzyme kinetics, we have reinvestigated this problem paying particular attention to the minimization of experimental errors. We chose to study the homologous series of *p*-nitrophenyl esters of 2-(5-*n*-alkyl)furoic acid



because of the properties of *p*-nitrophenyl 2-furoate hydrolysis described by Inward and Jencks (1965). The acylation of chymotrypsin by these esters is easily monitored from the

† From the Department of Chemistry, the Division of Biochemistry, The Ohio State University, Columbus, Ohio 43210. *Received August 7, 1975.* This research was supported in part by the National Science Foundation Grant GB-32777. M.H.K. is the recipient of a Research Career Development Award from the National Institutes of Health.