Commitment to Folded and Aggregated States Occurs Late in Interleukin-1β Folding[†]

John M. Finke, Larry A. Gross, Hoang M. Ho, David Sept, Bruno H. Zimm, and Patricia A. Jennings*

Department of Chemistry and Biochemistry, University of California at San Diego, La Jolla, California 92093-0359

Received June 7, 2000; Revised Manuscript Received September 5, 2000

ABSTRACT: A point mutation, lysine 97 to isoleucine, in the all-β cytokine interleukin-1β (IL-1β) exhibits an increased propensity to form inclusion bodies in vivo and aggregates in vitro. In an effort to better understand the aggregation reaction and determine when intervention may allow rescue of protein from aggregation during renaturation, we developed a novel application of mass spectrometry using isotopic labeling to determine the step(s) at which K97I commits to either the native or aggregated state. Interestingly, despite the early formation of a folding intermediate ensemble at an observed rate λ_2 of 4.0 s^{-1}, K97I commits to folding at a significantly slower rate λ_{CF} of 0.021 s^{-1}. This rate of commitment to folding is in excellent agreement with the observed rate of K97I native state formation ($\lambda_1 = 0.018$ s^{-1}). K97I also commits slowly to aggregation at an observed rate λ_{CA} of 0.023 s^{-1}. Earlier folding species and aggregates present prior to these commitment steps are likely to be in a reversible equilibrium between monomeric folding intermediates and higher-order oligomers. Kinetic and equilibrium experimental measurements of folding and aggregation processes are consistent with a nucleation-dependent model of aggregation.

Although major advances in recombinant DNA technology during the past two decades have enabled the expression of large quantities of protein, the desired protein is often expressed in insoluble biologically inactive aggregates known as "inclusion bodies" (*1*−*4*). Interestingly, the soluble native states of these proteins are often quite stable and do not demonstrate any propensity to aggregate once folded. One likely hypothesis for this observation is that early non-native states in the protein folding process compete between first-order intrachain and higher-order interchain binding events (*5*, *6*). The end result of these processes is a fraction of nascent chain folded properly in the soluble native state and a fraction misfolded in an insoluble protein aggregate.

Studying the events leading to protein aggregation during renaturation has posed a significant biophysical challenge (*7*). Specifically, aggregation from transiently populated kinetic states during protein folding (*8*, *9*), as opposed to aggregation under equilibrated conditions (*10*, *11*), is difficult to study because the kinetic lifetimes of most folding intermediates are short (*5*). Successful experiments used to address the mechanism of aggregation during refolding enable either the physical trapping of folding and aggregating intermediates (*12*) or the properties of these species (*13*−*16*). Indeed, the need for methods to address kinetic commitment, i.e., the "point of no return", to both the native and aggregated states is paramount (*14*). Comparing real-time folding/aggregation kinetics to commitment kinetics can determine which folding species can commit to folding/aggregation.

Commitment experiments typically follow a two-step procedure similar to double-jump experiments (*14*−*16*). The first step typically renatures the protein from the unfolded state to the desired folding/aggregating conditions. The second step suddenly changes experimental conditions such that committed species can be trapped and measured. Although the experimental details of this second step vary between different methods, the basic idea is that each data point in a commitment experiment is determined by the time delay between the first and second steps. Thus, commitment kinetics are always measured by varying the delay between steps 1 and 2 over a series of measurements, ultimately resulting in a kinetic curve as enough time points are collected.

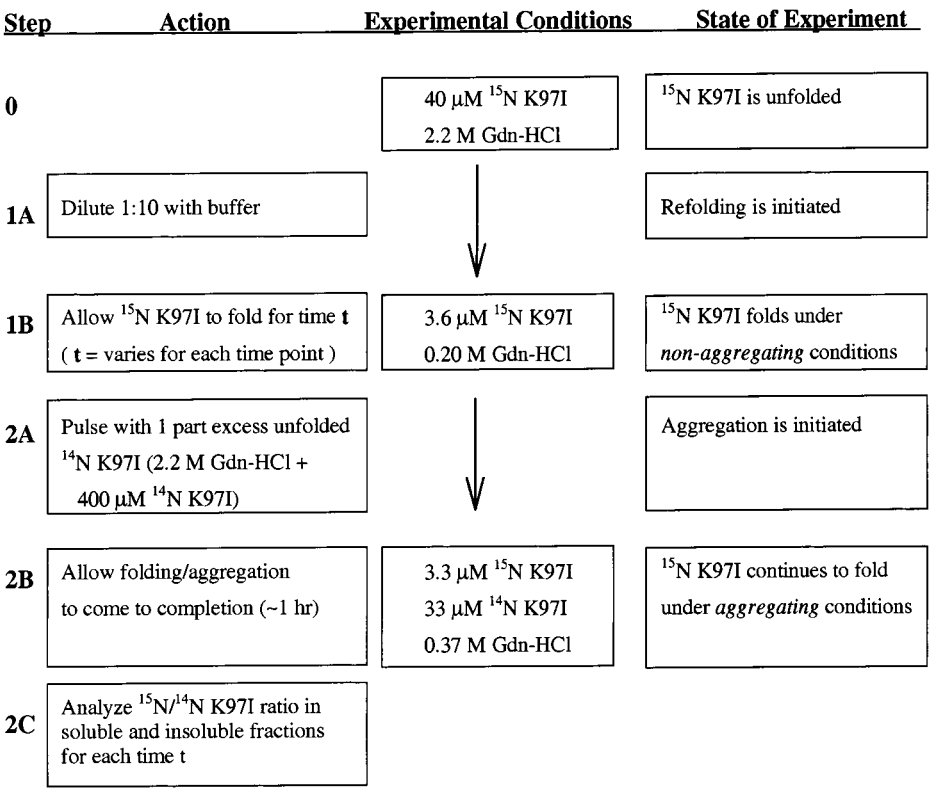
The goal of the commitment-to-folding experiment is to allow the protein to fold under nonaggregating conditions (step 1) and, at various time points, quickly alter the experiment conditions such that the monomeric protein will aggregate (step 2). Although large temperature jumps (*16*) and denaturant shifts (*13*, *15*) will modulate protein aggregation at step 2, these techniques do not allow comparable conditions between commitment to folding and the complementary experiment (commitment to aggregation). Furthermore, if step 2 consists of a large shift in temperature or denaturant, concern exists about whether the experimental conditions prior to step 2 still apply in the interpretation of results measured after step 2. Ideally, one would prefer to modulate aggregation at step 2 through large changes in protein concentration, with minimal change in denaturant,

[†] This work supported by U.S. Public Health Service Grant CA09523 (J.M.F.), a Hellman Faculty Fellowship (P.A.J.), a Sloan Fellowship (P.A.J.), NIH Grant GM54038 (P.A.J.), and a NSERC Fellowship (D.S.).

* To whom correspondence should be addressed: University of California at San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0359. Phone: (858) 534-6417. Fax: (858) 534-6714. E-mail: pajennin@ucsd.edu.

10.1021/bi001307d CCC: $19.00
Published on Web 11/23/2000

Scheme 1



temperature, and folding rate constants. Goldberg et al. succeeded with the commitment-to-folding experiment by renaturing low concentrations of turkey lysozyme in step 1 and renaturing high concentrations of hen lysozyme in step 2 (*14*). This commitment experiment proved to be successful due to intrinsic lysozyme activity and the selective inhibition of hen lysozyme using a specific antibody. However, this method may not be universally applicable as not all proteins have enzymatic activity or the availability of a selectively inhibiting antibody.

One strategy for overcoming these problems is to use a variation of pulse−chase experiments, where the protein is labeled with a tracer molecule. For instance, a protein can be radiolabeled with ^{14}C, ^{35}S, ^{32}P, or ^{3}H to allow quantitative assessment of the protein's incorporation into aggregates. However, purifying sufficient quantities of radiolabeled material as well as working with radioactive materials presents logistical complications. The method developed for the present commitment study uses ^{15}N isotopic labeling with mass spectrometry to accomplish the commitment-to-folding experiment by rapidly increasing the effective protein concentration in step 2 of the assay. The advantages of the ^{15}N labeling technique are that both commitment to folding and aggregation can be studied under similar solvent conditions and that it has wide applicability toward the study of protein aggregation. We now report the rates of commitment to folding and commitment to aggregation for the interleukin 1-β (IL-1β)[1] mutant K97I, a variant similar to wild-type IL-1β in all respects except for an increased susceptibility to aggregation (*17*, *18*). By comparing K97I rates of commitment to K97I rates of folding and aggregation, we have determined that only the final native state and final aggregated state(s) can commit. The results are consistent with the previous hypothesis of nucleation-dependent aggregation during renaturation, which can explain both the kinetic and equilibrium measurements.

MATERIALS AND METHODS

Expression and Purification of IL-1β Mutant K97I. The K97I protein was cloned, expressed, and purified as described by Finke et al. (*18*). ^{15}N-labeled K97I was expressed in cells grown in minimal medium containing ^{15}N-labeled ammonium chloride. Purified K97I was dialyzed extensively into buffer containing 10 mM MES, 2 mM EDTA, 90 mM NaCl, and 1 mM βme at pH 6.5 for all the experiments that are discussed here.

Commitment-to-Folding Experiments. A folding commitment experiment (Scheme 1) was performed to assess the time-dependent production of committed K97I protein. Prior to the experiment (step 0), ^{15}N-labeled K97I was unfolded with a 40 μM ^{15}N-labeled K97I/2.2 M Gdn-HCl mixture. Folding (step 1A) was initiated by diluting 1 part unfolded [^{15}N]K97I with 10 parts buffer to reach concentrations of 3.6 μM [^{15}N]K97I and 0.20 M Gdn-HCl. For each sample, the [^{15}N]K97I was allowed to refold for various times, *t* (step 1B), after the initiation of refolding in step 1A and before the [^{14}N]K97I pulse in step 2A. In step 2A, 1 part 400 μM unfolded [^{14}N]K97I/2.2 M Gdn-HCl mixture was added to reach final concentrations of 33 μM [^{14}N]K97I and 0.37 M Gdn-HCl. Immediately after step 2A, the isotopic mixture of renaturing K97I is allowed to reach an irreversible equilibrium consisting of soluble native and insoluble ag-

[1] Abbreviations: AU, absorbance units; EDTA, ethylenediaminetetraacetic acid; βme, 2-mercaptoethanol; MES, 2-(4-morpholino)-ethanesulfonic acid; Gdn-HCl, guanidine hydrochloride; ESI-MS, electrospray ionization mass spectrometry; HPLC, high-pressure liquid chromatography; IL-1β, interleukin-1β.

gregated states (step 2B). During step 2B, [^{15}N]K97I committed-to-folding monomers remain in solution and do not incorporate into aggregating [^{14}N]K97I. Also during step 2B, [^{15}N]K97I monomers which are not committed to folding can associate into [^{14}N]K97I aggregates. Subsequent procedures, described below (see ESI-MS Analysis), are then necessary to determine the ratio of ^{15}N to ^{14}N for each sample (step 2C).

It should be noted that the denaturant conditions necessarily differ between step 1 (0.20 M Gdn-HCl) and step 2 (0.37 M Gdn-HCl) due to the introduction of 1 part 2.2 M Gdn-HCl denatured [^{14}N]K97I at step 2A, an unavoidable consequence of the commitment-to-folding experiment. In the case of K97I, this denaturant shift corresponds to only a minor (10−20%) decrease in K97I folding rates and stability (*18*), effectively resulting in no change in experimental conditions prior to and after step 2A.

The soluble and insoluble K97I were separated by centrifugation at 7000*g*. The soluble portion was checked for the absence of insoluble particles by the absence of turbidity at 500 nm (<0.003 absorbance unit). Insoluble K97I was resolubilized and unfolded in 10 parts 6 M Gdn-HCl and refolded in 100 parts buffer, conditions where K97I refolds without aggregating (3.6 μM K97I and 0.55 M Gdn-HCl). Both the soluble and insoluble K97I samples were then dialyzed with 5 × 2 L exchanges using distilled H_2O. All samples were then concentrated to 200 μL using Centricon 10 concentrators (Amicon).

ESI-MS Analysis. The relative percentages of [^{15}N]- and [^{14}N]K97I within both the soluble and precipitated fractions at each time point, *t*, were determined by mass spectrometry. It is important to note that quantitation with mass spectrometry can be problematic for two reasons: (1) attempts to quantify protein in a multicomponent mixture may be in error as the signal of interest may be suppressed by contaminants and (2) the instrumental response is affected by the isotopic purity of a protein. To address these issues, we used purified K97I protein (>99%), demonstrated successful quantitation of IL-1β (*19*), and used [^{14}N]K97I as a standard to quantitate the relative amount of [^{15}N]K97I to [^{14}N]K97I (see below).

The mass spectra were obtained on a Hewlett-Packard (HP) 5989B mass spectrometer using a model 59987A atmospheric pressure ionization interface built by Analytica of Branford. The instrument is equipped with the extended mass range option. Electrospray ionization was performed under the conditions normally employed for protein spectra, i.e., in positive ion mode with a cylindrical potential of −5 kV, and the spray needle at ground potential. The sample was injected into a flowing solution containing 50% methanol (v/v) reverse-osmosis water with 1% acetic acid (Fisher HPLC grade), supplied by a HPLC isocratic pump through a Valco six-port valve at a rate of 10 μL/min.

The instrument was tuned using the standard protein solutions of cytochrome *c* and myoglobin, also supplied by HP. The known mass-to-charge peaks of these samples were used for mass calibration between 500 and 1600 *m*/*z* units. A solution of purified IL-1β was also used as a calibration reference. For each time, *t*, sample measurement, an average of 10−20 spectra were collected over an *x*-axis domain from 1500 to 1980 *m*/*z* units. The scanning step in each case was 0.1 *m*/*z* unit, and the scan rate was approximately 0.3 ms per step. One spectrum for each time, *t*, sample was measured to include the +9, +10, and +11 peaks. These mass-to-charge peaks were deconvoluted by the instrument software to produce the mass spectrum of the two isotopically characterized states of K97I. K97I gives a maximum signal on a multiply charged peak of +10 at *m*/*z* 1736 for ^{14}N-labeled K97I and *m*/*z* 1756 for ^{15}N-labeled K97I, measurements consistent with previous studies (*19*).

The commitment-to-folding method relies upon standardizing the relative amount of [^{15}N]K97I with the amount of [^{14}N]K97I in each sample. The raw amounts of [^{15}N]- and [^{14}N]K97I were determined in each sample using the area under the +10 peak centered at *m*/*z* 1736 for [^{14}N]K97I and *m*/*z* 1756 for [^{15}N]K97I. The relative amount of each [^{15}N]K97I peak was quantitated by dividing the [^{15}N]K97I peak area by the area under the [^{14}N]K97I peak present in the same mass spectrum. The sensitivity of the isotopically impure [^{14}N]K97I peak was shown to be similar to the isotopically pure [^{15}N]K97I, as determined by a peak height ratio of 0.95. The standardization of [^{15}N]K97I with [^{14}N]K97I was justified experimentally by the measurement of similar turbidity values after step 2B in Scheme 1 for all time, *t*, samples. This indicates a similar total aggregation reaction in all samples and minimal contributions from aggregating [^{15}N]K97I. As an additional control, the measurement of mass spectra from both soluble and insoluble samples in triplicate was used to confirm the consistency and precision of the results (see the Results).

The degree of commitment to folding was assessed initially by comparing the relative spectra of [^{15}N]K97I of each sample after standardization with the area under the [^{14}N]K97I peak (*m*/*z* 1736) using the equation

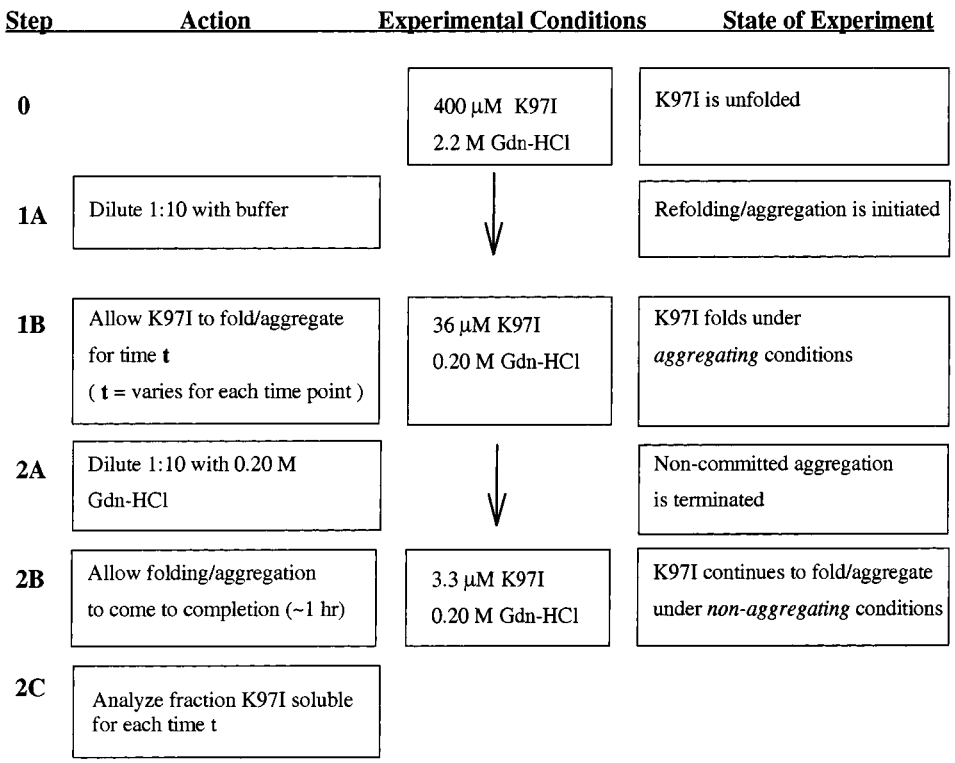
$$\mathrm{R15} = \frac{\mathrm{S15}}{A_{\mathrm{S14}}} \tag{1}$$

where R15 is the relative standardized mass spectra of the [^{15}N]K97I *m*/*z* +10 peak, S15 is the raw mass spectra of the [^{15}N]K97I *m*/*z* +10 peak, and A_{S14} is the area under the raw mass spectra of the [^{14}N]K97I *m*/*z* +10 peak. Thus, the relative standardized [^{15}N]K97I spectra are the raw [^{15}N]K97I spectra divided at each *m*/*z* increment by the single value corresponding to the total area under the [^{14}N]K97I. The purpose for the calculation of the standardized [^{15}N]K97I spectra R15 is to visually compare the time dependence of the corrected mass spectra in the commitment experiment. To calculate the molar fraction of [^{15}N]K97I to [^{14}N]K97I in the sample, the area under R15, A_{R15}, is measured:

$$A_{\mathrm{R15}} = \frac{[[^{15}\mathrm{N}]\mathrm{K97I}]}{[[^{14}\mathrm{N}]\mathrm{K97I}]} \tag{2}$$

Refolding Assessed in Real Time with Tryptophan Fluorescence. Protein refolding was monitored prior to 20 s with an Applied Photophysics SX.17MV (Applied Photophysics, London) stopped-flow unit with a path length of 0.1 cm. Protein refolding was monitored after 20 s using a Fluoromax-2 spectrofluorimeter equipped with a Neslab RTE-111 temperature controller. Refolding experiments were initiated by a 1:10 dilution of unfolded K97I at 2.2 M Gdn-HCl into final Gdn-HCl concentrations of 0.20 M and a final K97I concentration of 6 μM. Stopped-flow fluorescence was

Scheme 2



measured with excitation at 293 nm and emission collected through a >320 nm cutoff filter. Manual-mixing fluorescence was measured by the time-dependent change in fluorescence emission at 343 nm (slit of 1 mm) while exciting at 293 nm (slit of 1 mm). The stopped-flow kinetic trace is the average of 10−20 data acquisitions. The manual-mixing kinetic trace is the average of three data acquisitions.

Commitment-to-Aggregation Experiments. A commitment-to-aggregation experiment (Scheme 2) was performed to assess the time-dependent production of committed (i.e., refolding resistant) K97I aggregates. Prior to the experiment (step 0), K97I was unfolded at 400 μM/2.2 M Gdn-HCl. Folding (step 1A) was initiated by diluting 1 part unfolded K97I with 10 parts buffer to reach concentrations of 36 μM K97I and 0.20 M Gdn-HCl. For each sample, the K97I was allowed to refold for various times, *t* (step 1B), from the initiation of refolding and before dilution in step 2A. In step 2A, the solution was diluted 1:10 to reach final concentrations of 3.3 μM K97I and 0.20 M Gdn-HCl. Immediately after step 2A, the renaturing K97I is allowed to reach folding/aggregation irreversible states (step 2B). Committed aggregates continue to form insoluble aggregates, while noncommitted aggregates dissociate and refold properly as soluble monomers. Noncommitted K97I monomers will associate into K97I aggregates. Subsequent procedures, described below, are then necessary to determine the fraction of soluble K97I (step 2C). In step 2C, the soluble and insoluble K97I were separated by centrifugation at 7000*g*. The soluble portion was checked for the absence of insoluble particles by the lack of turbidity at 500 nm (<0.003 absorbance unit). The soluble protein concentration was measured by the absorbance at 280 nm (AU_{280}) and calculated using the extinction coefficient ϵ_{280} of 11.26 mM^{-1} cm^{-1} with the Beer−Lambert relationship.

Aggregation Assessed in Real Time with Light Scattering. Protein aggregation during refolding was monitored prior to 20 s with an Applied Photophysics SX.17MV stopped-flow unit with a path length of 0.1 cm. Slower aggregation reactions were monitored using manual-mixing techniques monitored with a Fluoromax-2 spectrofluorimeter equipped with a Neslab RTE-111 temperature controller. Aggregation was initiated by a 1:10 dilution of unfolded K97I at 2.2 M Gdn-HCl into a final Gdn-HCl concentration of 0.20 M and a final K97I concentration of 33 μM. Stopped-flow light scattering was measured at a wavelength of 500 nm and collected at an angle of 90° through a >320 nm cutoff filter. Manual-mixing light scattering was measured by the time-dependent change in light scattering at 500 nm (slit of 1 mm) while exciting at 500 nm (slit of 1 mm). Each stopped-flow kinetic trace is the average of 10−20 data acquisitions. Each manual-mixing kinetic trace is the average of five data acquisitions.

Soluble Fraction. Refolding experiments were initiated by a 1:10 dilution of K97I from 2.2 M Gdn-HCl to the indicated final K97I concentrations of protein and 0.20 M Gdn-HCl. Samples were centrifuged at 7000*g*, and the completeness of the aggregate separation was assessed by the absence of turbidity at 500 nm in the soluble fraction. The soluble protein concentration was measured by the absorbance at 280 nm (AU_{280}) and calculated using the extinction coefficient ϵ_{280} of 11.26 mM^{-1} cm^{-1} with the Beer−Lambert relationship.

Data Analysis. Time-dependent changes in commitment-to-folding tracks with the fraction of [^{15}N]K97I/[^{14}N]K97I as measured by variable A_{R15}. Time-dependent changes in commitment to aggregation track with the 280 nm absorbance (AU_{280}) of soluble K97I remaining in solution. The fraction of K97I commitment in either a commitment-to-

folding or commitment-to-aggregation experiment can be quantitated through the following equation

$$C(t) = \frac{V(t) - V_0}{V_\infty - V_0} \quad (3)$$

where $C(t)$ is the fractional degree of commitment at time t, $V(t)$ is the variable measuring commitment at time t, V_0 is the variable measuring commitment at time zero, and V_∞ is the variable measuring commitment at time ∞ (i.e., native K97I protein). The variable $C(t)$ reflects the fraction of commitment in both folding and aggregation commitment studies. To distinguish between different commitment experiments, this value is denoted $C_I(t)$ for the insoluble fraction of the commitment-to-folding experiment, $C_S(t)$ for the soluble fraction of the commitment-to-folding experiment, and $C_{AG}(t)$ for commitment to aggregation. The variable $V(t)$ is the parameter measured in the commitment study. In the case of commitment to folding, $V(t)$ is the parameter A_{R15}. In the case of commitment to aggregation, $V(t)$ is the parameter of K97I AU_{280}, measured after centrifugation. In the commitment-to-folding experiment, a decrease in A_{R15} in the insoluble fraction should result in an increase in A_{R15} in the soluble fraction. Consequently, a convincing control for the commitment-to-folding experiment is agreement between the variables $C_I(t)$ and $C_S(t)$.

$$C_1(t) \cong C_s(t) \quad (4)$$

Real-time kinetics measured with stopped-flow and manual-mixing techniques, the commitment to folding in the insoluble fractions $C_I(t)$, the commitment to folding in soluble fractions $C_S(t)$, and the commitment to aggregation $C_{AG}(t)$ were fit to eq 5 using in-house software.

$$A(t) = \sum^{i} A_i e^{-\lambda_1 t} + A_\infty \quad (5)$$

where $A(t)$ is the measured signal [i.e., fluorescence, $C_I(t)$, $C_S(t)$, or $C_{AG}(t)$] at time t, i is the number of observed aggregation processes measured by first-order fits, A_i is the amplitude of kinetic process i, A_∞ is the equilibrium value of $A(t)$ at time ∞, and λ_i is the observed rate constant for process i. The values of i, A_i, A_∞, and λ_i were determined using the fit quality represented in the χ_r^2 values (*20*), the random dispersion of residuals, and the logical consistency of the generated fitting parameters. The λ_i value is an observed rate constant and does not directly measure the microscopic rate constant for a kinetic process.

Stopped-flow and manual-mixing aggregation kinetics cannot be fit rigorously without information regarding the quantity and character of oligomeric states involved in the aggregation process. Nonetheless, approximate rates can be derived with a fit of the time-dependent light scattering signal to the sum of first-order processes using eq 5. Due to the higher-order nature of aggregation, fits of light scattering data with observed rates λ_i should not be interpreted as evidence of either a microscopic rate constant or even a first-order kinetic process. To avoid confusion between observed first-order processes in folding and approximated first-order aggregation processes measured with light scattering, the parameter λ_{AGG} will be used to characterize aggregation processes. These approximated first-order aggregation rates, λ_{AGG}, can be used as qualitative parameters to determine the effect of protein concentration to these steps through eq 6 (*21*):

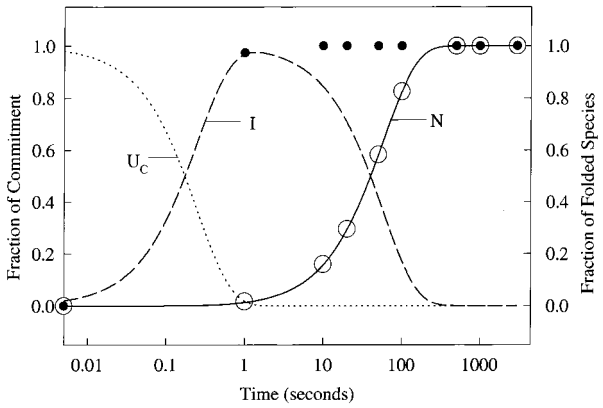


FIGURE 1: K97I folding species formed during renaturation with two commitment scenarios. Unfolded K97I ensemble U_C (•••) folds to the intermediate ensemble I (− − −) with an observed rate constant λ_2 of 4.0 s^{-1}. Native K97I (—) populates at the observed rate λ_1 of 0.018 s^{-1}. If the intermediate ensemble (●) commits to folding, complete commitment will occur before kinetic data points can be collected. If the native state (○) commits to folding, commitment will occur on a time scale where kinetic data points will measure the commitment process.

$$\log(\lambda_{AGG-i}^{[K97I]}) = R \log[K97I] + K \quad (6)$$

where $\lambda_{AGG}^{[K97I]}$ is the rate of an aggregation process at a given protein concentration [K97I], R is the apparent rate order of the process described by λ_{AGG}, and K is $\log(\lambda_{AGG})$ in the limit [K97I] $\rightarrow$ 0.

RESULTS

Commitment Experiments. We want to address the question of when a protein is irreversibly committed to either the native or misfolded aggregated state. Our experiment uses the competition between folded and aggregated states during the renaturation of K97I to measure commitment. In this way, the species populated during folding, which have been well characterized by stopped-flow circular dichroism (J. M. Finke, unpublished data), quench-flow pulse labeling (*19*, *22*), and double-jump experiments (*18*), can be correlated with the time-dependent increase in commitment. A plot of the time evolution of species populated during folding of K97I at 0.20 M Gdn-HCl and 25 °C is given in Figure 1. Refolding is best characterized as the population of an initial unstructured ensemble U_C, which folds within 1 s to a partially structured ensemble of intermediates I and finally folds slowly to the native state N. The unfolded form is designated U_C (indicating partial collapse) instead of U, because, although lacking in significant secondary structure, it not likely to be completely devoid of it (*23*, *24*).

Figure 1 also indicates two possible scenarios for commitment during K97I folding at 0.20 M Gdn-HCl and 25 °C. In each scenario, the population of the committed species would be measured at various times t when step 2A in Scheme 1 is executed. In scenario 1, intermediate ensemble I commits to folding. In this case, commitment will occur fast such that manual mixing data points will not be acquired quickly enough to measure the kinetics. In our case, the

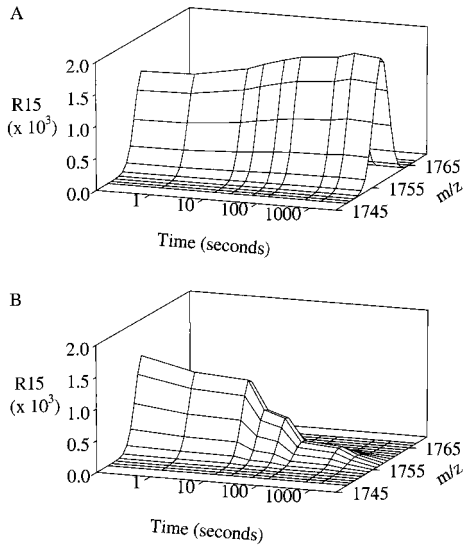


FIGURE 2: Time dependence of the relative standardized mass spectra (R15) of the +10 *m*/*z* peak of ^{15}N-labeled K97I in the commitment-to-folding experiment. (A) Time dependence of the R15 in the soluble sample fraction. The soluble R15 peak increases in area from its initial value of 70% maximum area at time zero to 100% maximum area at 3000 s. (B) Time dependence of R15 in the insoluble sample fraction. The insoluble R15 peak increases in area from its maximum area at time zero to zero area at 3000 s.

protein would therefore appear to commit from zero commitment to complete commitment within 1 s. In scenario 2, only the native state can commit to folding. In this case, manual-mixing techniques will adequately measure commitment kinetics as the native state slowly populates.

Commitment to Folding. In these studies, we take advantage of the fact that ^{14}N- and ^{15}N-labeled K97I differ in molecular mass by 200 Da and can be easily distinguished by mass spectrometric techniques. In Figure 2, the [^{15}N]K97I mass spectra are shown after the standardization described by eq 1 (Materials and Methods). These standardized mass spectra are graphed as slices along the time axis and follow the time-dependent process by which K97I commits to the native folded state. In Figure 2A, the magnitude of the R15 peak in the soluble portion is shown to increase by 30% between time zero and 3000 s. To confirm this time-dependent increase in the level of [^{15}N]K97I in the soluble samples, the time-dependent decrease in the level of [^{15}N]K97I from the insoluble aggregates was also measured. In Figure 2B, the R15 peak in the insoluble fraction is shown to decrease from its maximal value at time zero to zero at time *t* (3000 s). Together, these results indicate that K97I ultimately reaches a state which will not aggregate and is committed to folding to the soluble native state.

Note that, even under conditions where the greatest amount of aggregation occurs (time zero), a significant portion of [^{15}N]K97I folds to the soluble state. This is due to the fact that experimental limitations do not allow 100% K97I aggregation at time zero. This 30% difference in the soluble peak height between the spectrum at 0 s and the spectrum at 3000 s is consistent with the extent of K97I aggregation measured at 33 μM in previous studies (*18*). The similarity in the R15 peak size measured in the soluble (Figure 2A) and insoluble (Figure 2B) fractions at time zero is consistent with equal soluble/insoluble partitioning of [^{15}N]K97I relative to [^{14}N]K97I expected at time zero.

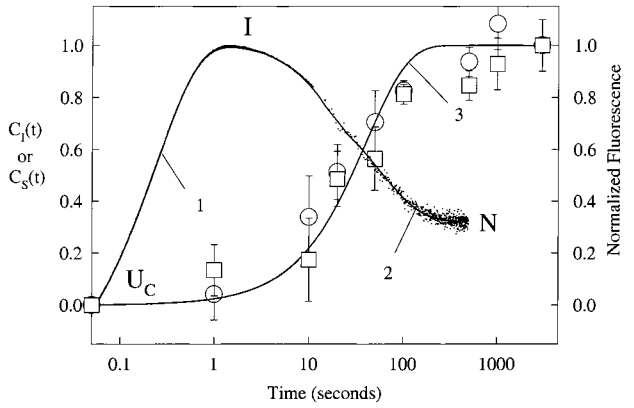


FIGURE 3: Real-time K97I folding compared to K97I commitment to folding. Stopped-flow/tryptophan fluorescence (trace 1) characterizes the transition from an unfolded ensemble U_C and a partially folded intermediate ensemble I (λ_2 = 4.0 s^{-1}). Manual-mixing/tryptophan fluorescence (trace 2) characterizes the final folding step from intermediate ensemble I to the native state N (λ_1 = 0.018 s^{-1}). The K97I fraction committed to folding (trace 3) follows a single-exponential process (λ_{CF} = 0.021 s^{-1}) when fit to the commitment data from the soluble $C_S(t)$ (○) and insoluble $C_I(t)$ (□) samples at various time points. Error bars for $C_S(t)$ and $C_I(t)$ indicate the standard deviation of three separate sample measurements. The left-hand *y*-axis [$C_S(t)$ or $C_I(t)$] refers to commitment-to-folding data (trace 3), and the right-hand *y*-axis (normalized fluorescence) refers to real-time fluorescence data (traces 1 and 2). The observed rate of the commitment to folding (λ_{CF} = 0.021 s^{-1}) correlates well with the I to N folding rate (λ_1 = 0.018 s^{-1}).

Figure 3 shows the real-time folding process as well as the global fit of the commitment-to-folding kinetics. Traces 1 and 2 indicate respective stopped-flow and manual-mixing measurements of real-time tryptophan fluorescence kinetics for the K97I folding at 0.20 M Gdn-HCl. These kinetic data fit best to two exponentials, with observed rate constants λ_2 of 4.0 s^{-1}, characterizing the folding of U_C to I, and λ_1 of 0.018 s^{-1}, characterizing the folding of I to N. These measurements are consistent with the simple folding mechanism proposed for K97I folding (Figure 1).

Trace 3 indicates the fit of the commitment-to-folding kinetic data as measured by variables $C_S(t)$ for the soluble fractions and $C_I(t)$ for the insoluble fractions. The *y*-axis label, "$C_S(t)$ or $C_I(t)$", is a measure of the fraction of [^{15}N]K97I committed at time *t*. The variables $C_S(t)$ (circles) and $C_I(t)$ (squares) are in good agreement, as expected from eq 4. Error bars of both $C_S(t)$ and $C_I(t)$ indicate the standard deviation from three separate sample measurements at each time (*t*). A fit of the combined soluble and insoluble kinetic data indicates that the commitment to folding is well characterized by a single-exponential process with an observed rate constant λ_{CF} of 0.021 s^{-1}. This result is consistent with native state commitment to folding as described in Figure 1.

Commitment to Aggregation. In conjunction with commitment-to-folding studies, commitment-to-aggregation studies are also important in determining the irreversible steps in the aggregation pathway. Like commitment to folding, the commitment to aggregation may follow a fast or slow kinetic rate. However, there is no reason, a priori, for the rates of commitment to folding and commitment to aggregation to be on similar time scales. Figure 4 shows the commitment-to-aggregation kinetics overlaid on the real-time aggregation process observed by light scattering. Traces 1

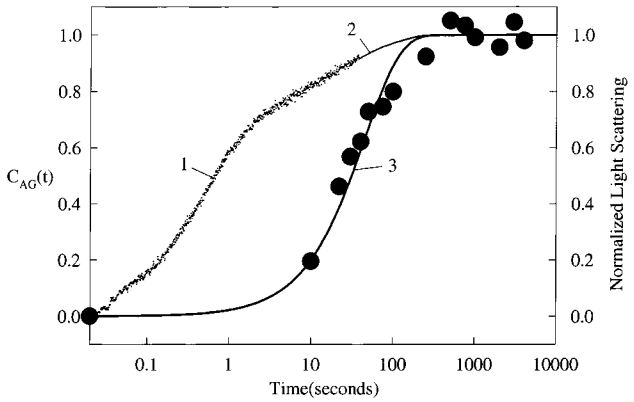


FIGURE 4: Real-time K97I aggregation compared to K97I commitment-to-aggregation. Stopped-flow/light scattering (trace 1) measures early aggregation step(s) between 0.01 and 10 s. Manual-mixing/light scattering (trace 2) measures a slower aggregation process between 10 and 500 s, after which the aggregation is complete. The fraction of aggregated K97I committed to aggregation (trace 3) follows a single-exponential process ($\lambda_{CA} = 0.023$ s^{-1}) when fit to the commitment data $C_{AG}(t)$ (●) derived from the remaining soluble K97I concentration. The left-hand *y*-axis [$C_{AG}(t)$] refers to commitment-to-aggregation data (trace 3), and the right-hand *y*-axis (normalized fluorescence) refers to real-time light scattering data (traces 1 and 2). The observed rate of the commitment to aggregation ($\lambda_{CA} = 0.023$ s^{-1}) correlates well with the observed rate of the slow aggregation process ($\lambda_{AGG}^{slow} = 0.018$ s^{-1}).

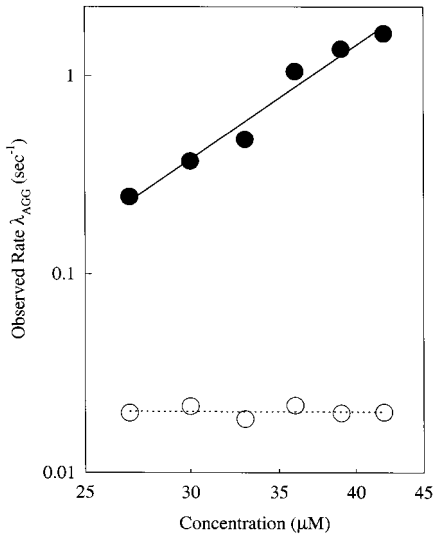


FIGURE 5: Effect of concentration on observed aggregation rates for the fast aggregation process λ_{AGG}^{fast} (●) and the slow aggregation process λ_{AGG}^{slow} (○). The linear fit of log(λ_{AGG}^{fast}) to log(μM K97I) yields a slope of 4.5 (—), indicating a concentration dependence of 4−5. The linear fit of log(λ_{AGG}^{slow}) to log(μM K97I) yields a slope of 0.1 (- - -), indicating a first-order concentration dependence.

(stopped-flow) and 2 (manual-mixing) describe the typical aggregation reaction during refolding at 42 μM K97I in real time using light scattering at 500 nm. Rigorous fits of the light scattering data are not warranted without more information regarding the size and populations of the aggregated states. However, two distinct aggregation processes can be observed with light scattering, with approximate rates λ_{AGG}^{fast} of $\sim$1 s^{-1} and λ_{AGG}^{slow} of $\sim$0.02 s^{-1}. After the relaxation of the slower step, λ_{AGG}^{slow}, the aggregation reaction equilibrates at 32% K97I aggregation. Trace 3 indicates the fit of the commitment-to-aggregation data and is well characterized by a single-exponential process with an observed rate constant λ_{CA} of 0.023 s^{-1}.

Aggregation Kinetics. Estimated rate(s), λ_{AGG}^{fast} and λ_{AGG}^{slow}, calculated using eq 5 (see Data Analysis in Materials and Methods), observed during the aggregation process are displayed with respect to K97I concentration in Figure 5. The observed rate for the fast process λ_{AGG}^{fast} is highly concentration-dependent, whereas the slow process λ_{AGG}^{slow} appears to be concentration-independent (first-order). Fits of the observed rates to concentration using eq 6 (see Data Analysis in Materials and Methods) fit the apparent rate order *R* of 4.5 for λ_{AGG}^{fast} and *R* of 0.1 for λ_{AGG}^{slow}.

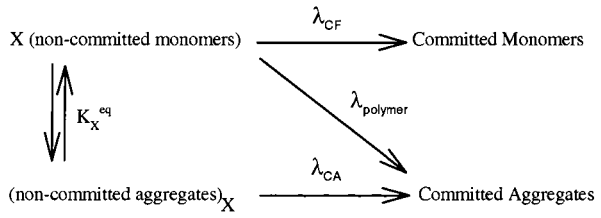
DISCUSSION

Although not defined outright as such, a variety of commitment studies have been performed previously. Brems performed a two-step experiment which, in the first step, diluted denatured bovine growth hormone to protein and Gdn-HCl concentrations supporting folding (*15*). To test commitment to folding after various delay times, the sample was transferred to protein and Gdn-HCl concentrations in which aggregation was favorable. Using similar mixing techniques, Cleland and Wang probed whether species formed after 15 min of bovine carbonic anhydrase refolding were resistant to aggregation (*13*). Danner and Seckler also performed both folding and aggregation commitment studies using temperature jumps during refolding of the temperature-sensitive P22 tailspike protein (*16*). Goldberg et al. performed the first definitive commitment experiment, measuring commitment kinetics in both the folding and aggregation pathways of denatured, reduced turkey lysozyme (*14*). Determining the commitment in either pathway was established by altering the effective protein concentrations and minimally perturbing the solvent conditions, thereby providing directly comparable commitment rates between folding and aggregation. To study folding commitment steps in K97I aggregation, this study uses an isotopic labeling−mass spectrometry application, which can be easily used to study other aggregating proteins. The method also allows for the study of both commitment to folding and aggregation under the same denaturant, pH, and temperature conditions, conditions crucial to a direct comparison of folding and aggregation commitment rates.

Interleukin-1β Commits to Folding Late in Renaturation. In Figure 3, the commitment to folding, indicated by $C_S(t)$ and $C_I(t)$ (trace 3), is essentially simultaneous with the slow folding step λ_1 observed in K97I folding (trace 2). The observed folding rate λ_1 has been shown to correspond to the formation of native K97I from a partially folded intermediate ensemble I (*18*, *19*). No significant change in the values of $C_S(t)$ or $C_I(t)$ is observed during the formation of the intermediate ensemble I which accumulates at a rate λ_2 of 4.0 s^{-1}. Thus, intermediate ensemble I is not a state which commits K97I to successful folding. The observed rate constant fitted to $C_S(t)$ and $C_I(t)$ ($\lambda_{CF} = 0.021$ s^{-1}) is highly similar to that of λ_1 ($\lambda_1 = 0.018$ s^{-1}). Therefore, the state which commits to folding is the native state and only the native state.

Interleukin-1β Commits to Aggregation Late in Renaturation. The commitment to aggregation (Figure 4) fits to a first-order kinetic model and is coincident with the slow aggregation step λ_{AGG}^{slow} observed in K97I light scattering

Scheme 3



kinetics. The first-order time constant ($\lambda_{CA} = 0.023\ s^{-1}$) fitted to the increase in K97I commitment to aggregation is similar to λ_{AGG}^{slow} (0.022 s^{-1}). The observed rate λ_{AGG}^{slow} corresponds to the formation of the final equilibrated aggregation ensemble as no further increase in signal magnitude is observed. As indicated in Figure 4, committed aggregates are only 20% populated when the sample is diluted after aggregation for 10 s (see Materials and Methods). Thus, the early process measured by λ_{AGG}^{fast} does not result in a commitment, and the slower process λ_{AGG}^{slow} correlates with commitment to aggregation.

The Commitment to Aggregation Is First-Order. Observed aggregation rates λ_{AGG}^{fast} and λ_{AGG}^{slow} should not be used in a rigorous kinetic analysis due to the mechanistic complexity of the aggregation process. However, they can be used as an indicator of rate changes with respect to protein concentration. Figure 5 shows that the early step λ_{AGG}^{fast} is highly concentration-dependent ($R = 4.5$), whereas the final process λ_{AGG}^{slow} is concentration-independent ($R = 0.1$, i.e., first-order). The high rate order ($R = 4.5$) of λ_{AGG}^{fast} is suggestive of an aggregate consisting of at least four monomers.

Late Commitment Is Consistent with a Nucleation-Dependent Aggregation Mechanism. The nucleation-dependent mechanism, which has been proposed to describe amyloid fibril formation (*25*), was also proposed by Finke et al. to describe the aggregation observed in K97I refolding (*18*). This mechanism is defined as "a series of unfavorable protein−protein association equilibria leading to an unstable nucleus, followed by a series of favorable equilibria", culminating in mature aggregate formation (*25*). Nucleation-dependent polymerization is experimentally characterized by (1) mature aggregate formation only at monomer concentrations exceeding a critical concentration and (2) first-order kinetics observed in formation of the mature aggregate. In a previous study, it was shown that K97I produces mature aggregates only when refolded above the critical concentration of 20 μM (*18*), one of the hallmarks of nucleation-dependent polymerization. In addition to a critical concentration, the nucleation mechanism has been further supported by the existence of two distinct aggregation processes during refolding (Figure 4): noncommitted aggregates formed at a rate λ_{AGG}^{fast} and committed aggregates formed at a rate λ_{AGG}^{slow}. In a nucleation model, the rate λ_{AGG}^{fast} would characterize the formation of small aggregates involved in nucleus formation and the rate λ_{AGG}^{slow} would correspond to formation of the mature aggregate (*25*). Small aggregates assessed by the rate λ_{AGG}^{fast} could be dissociated and properly refolded if high-level dilution was initiated early enough in the refolding process. Postnucleus aggregates, assessed by the rate λ_{AGG}^{slow}, resist dissociation and continue to grow at an observed first-order rate, the second observation expected in a nucleation−polymerization model.

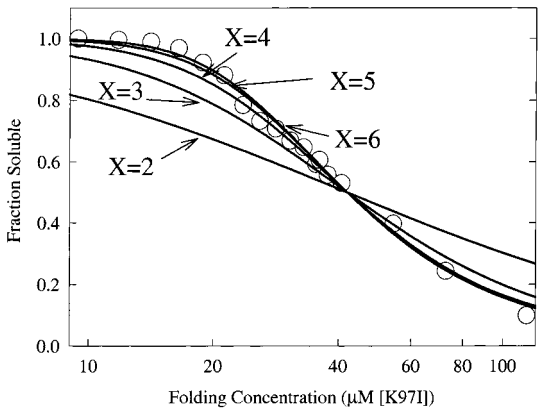


FIGURE 6: Dependence of K97I concentration on the final equilibrium fraction of soluble native K97I refolded from 2.2 M Gdn-HCl to 0.20 M Gdn-HCl. The fraction of soluble K97I (○) remains at 1.0 until 10 μM and decreases to 0.09 at 103 μM. Simulations of the equilibrium experimental data use the simple nucleation model of Scheme 3. The best fit of the nucleation model to the experimental data is shown using a nucleus size of 2, 3, 4, 5, and 6 monomers.

A Nucleation-Dependent Model for K97I. Using a simplistic approach, a nucleation mechanism for K97I aggregation during refolding is modeled using four species: (1) uncommitted monomers, (2) uncommitted aggregates (i.e., nucleus), (3) committed monomers (i.e., native K97I), and (4) committed aggregates. To address whether this simple nucleation model can accurately describe the aggregation of K97I, these four species are arranged in Scheme 3, where X is the number of monomers forming the nucleus, K_X^{eq} is the equilibrium constant of nucleus formation, λ_{CF} is the rate of commitment to folding (0.018 s^{-1}), λ_{CA} is the rate of commitment to aggregation (0.021 s^{-1}), and $\lambda_{polymer}$ is the rate of monomer incorporation into the final aggregate. The equilibrium percentage of soluble native K97I recovered from refolding at increasing K97I concentrations can be quantitated using numerical methods (NDsolve from Mathematica). Comparing this model to equilibrium raw data accomplishes two basic goals: (1) determining whether a nucleation model can reasonably fit experimental data and (2) addressing the estimate of a four-monomer minimum nucleus size. As with all kinetic mechanisms, successful fits of experimental data to this model neither prove the model's validity nor preclude additional steps.

The fraction of soluble K97I recovered after refolding at increasing K97I concentrations was fit to experimental data using kinetic steps described in Scheme 3. The simulations started with a fraction of uncommitted monomers of 1 and all other species at 0. To fit the simulated mechanism to experimental data, integer values of X (i.e., 2 for dimer, 3 for trimer, etc.) were selected, $\lambda_{polymer}$ was fixed at 1.0 s^{-1}, and K_X^{eq} was optimized at each fixed value of X. Figure 6 shows the fits of X values of 2, 3, 4, 5, and 6. Clearly, X values of 2−3 undershoot the raw data, while X values of 4−5 appear to fit much more accurately. If the value $\lambda_{polymer}$ is varied, different nucleus sizes X can be fit to the raw data. However, small nucleus sizes ($X = 2-3$) could not be fit to the raw data at any value of $\lambda_{polymer}$. Thus, these simulations (1) demonstrate that a nucleation mechanism can describe kinetic and equilibrium experimental measurements of K97I aggregation and (2) provide evidence that nucleus size X is comprised of 4 or more monomers.

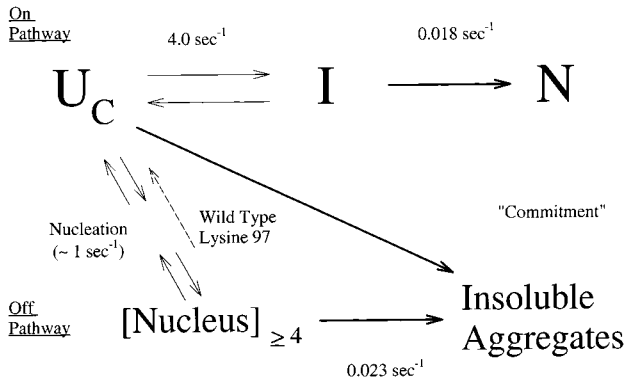


FIGURE 7: Proposed mechanism of K97I on-pathway folding and off-pathway aggregation. The on-pathway folding mechanism is initiated with an unstructured ensemble U_C. U_C equilibrates reversibly with a partially folded intermediate ensemble I at an observed rate λ_2 of 4.0 s^{-1}. Intermediate ensemble I commits to folding to the native state N at an observed rate λ_1 of 0.018 s^{-1}. Off-pathway aggregation is initiated from U_C and reversibly equilibrates with a nucleus ($\geq$4 monomers) at an observed rate λ_{AGG}^{fast} of $\sim$1 s^{-1}. The nucleus commits to aggregation as it adds monomers with an apparent first-order commitment rate λ_{AGG}^{slow} of 0.023 s^{-1}.

Conclusions. Building upon previous work on IL-1β folding (*17*−*19*, *22*, *26*) and the kinetics of commitment to folding/aggregation, we propose a comprehensive mechanism to describe K97I folding and aggregation (Figure 7). Early in the folding process, K97I forms a pre-equilibrium between the unstructured ensemble U_C, the monomeric intermediate ensemble I, and a nucleus aggregate ($N \geq 4$ monomers). The pre-equilibrium between U_C and I has been observed directly by pulse labeling/mass spectrometry with an approximate ratio of 4:1 (*19*). It has also been demonstrated that a drastic reduction in the K97I aggregation rate occurs as intermediate ensemble I is formed, indicating that aggregation initiates predominately from the unstructured ensemble U_C (*18*). Thus, the intermediate ensemble I does not aggregate but still does not commit K97I to folding. In short, fast folding (λ_2) and aggregation (λ_{AGG}^{fast}) steps are reversible and do not result in commitment to either folding or aggregation. In the slower phase of the refolding reaction, the committed native state is formed at a rate λ_1 and committed insoluble aggregates are formed at a rate λ_{AGG}^{slow}. In wild-type IL-1β, the pre-equilibrium is tipped against nucleus formation due to the presence of the charged lysine 97 (Figure 7). The nucleus will not be populated and will therefore not result in insoluble aggregates during the folding of wild-type IL-1β.

Biological Inferences. Although both folding and aggregation have been studied in many proteins, it has been difficult to ascribe specific properties which are common to all, other than their propensity to aggregate during renaturation. These proteins do not fall into a particular class of secondary structure since α-helical (*15*), β-sheet (*18*, *27*), and mixed α/β (*1*, *28*) proteins have been shown to aggregate during refolding. The folding pathways of these proteins are also varied, as simple two-state proteins (*28*), proteins with folding intermediates (*4*, *13*, *15*, *18*), and oligomeric proteins (*4*, *21*, *29*) all have been shown to aggregate. Although the folding species leading to aggregation is, in many cases, a species formed early in the dead time of folding measurements, it is unclear whether this species is always a largely unfolded state (*28*) or could possibly be an early intermediate species (*15*, *27*).

An important question in protein aggregation is when proteins commit to their soluble native state. It has been found in all proteins studied with commitment experiments, including K97I, that native state formation is necessary to commit the protein to the folded state (*13*−*16*). This may have important consequences on the biological importance of rate-limiting folding steps as certain natural proteins may face evolutionary pressure to avoid sequences which fold so slowly that aggregation predominates over productive folding. Perhaps a lower speed limit may exist on the folding of certain sequences which aggregate during refolding from non-native conformations. Conversely, if commitment occurs late in the folding, slow folding may allow for rescue of initially associated monomers during in vivo expression. Whether native state commitment is ubiquitous in protein folding remains an interesting question to be addressed with further studies.

ACKNOWLEDGMENT

We thank Dr. Joe Adams for the use of his stopped-flow fluorimeter and Dr. Osman Bilsel for providing his global analysis software.

REFERENCES

1. De Bernardez-Clark, E., and Georgiou, G. (1991) in *Protein Refolding* (Georgiou, G., and De Bernardez-Clark, E., Eds.) pp 1−20, American Chemical Society, Washington, DC.
2. Chan, W., Helms, L. R., Brooks, I., Lee, G., Ngola, S., McNulty, D., Maleeff, B., Hensley, P., and Wetzel, R. (1996) *Folding Des. 1*, 77−89.
3. Chrunyk, B. A., Evans, J., Lillquist, J., Young, P., and Wetzel, R. (1993) *J. Biol. Chem. 268*, 18053−18061.
4. Haase-Pettingell, C. A., and King, J. (1988) *J. Biol. Chem. 263*, 4977−4983.
5. Wetzel, R. (1996) *Cell 86*, 699−702.
6. London, J., Skrzynia, C., and Goldberg, M. E. (1974) *Eur. J. Biochem. 47*, 409−415.
7. Fink, A. L. (1998) *Folding Des. 3*, 9−23.
8. Mitraki, A., Haase-Pettingell, C., and King, J. (1991) in *Protein Refolding* (Georgiou, G., and De Bernardez-Clark, E., Eds.) pp 35−49, American Chemical Society, Washington, DC.
9. Georgiou, G., Valax, P., Ostermeier, M., and Horowitz, P. M. (1994) *Protein Sci. 3*, 1953−1960.
10. Kelly, J. W. (1997) *Structure 5*, 595−600.
11. Booth, D. R., Sunde, M., Bellotti, V., Robinson, C. V., Hutchinson, W. L., Fraser, P. E., Hawkins, P. N., Dobson, C. M., Radford, S. E., Blake, C. C., and Pepys, M. B. (1997) *Nature 385*, 787−793.
12. Goldenberg, D., and King, J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 3403−3407.
13. Cleland, J. L., and Wang, D. I. (1990) *Biochemistry 29*, 11072−11078.
14. Goldberg, M. E., Rudolph, R., and Jaenicke, R. (1991) *Biochemistry 30*, 2790−2797.
15. Brems, D. N. (1988) *Biochemistry 27*, 4541−4546.
16. Danner, M., and Seckler, R. (1993) *Protein Sci. 2*, 1869−1881.
17. Wetzel, R., and Chrunyk, B. A. (1994) *FEBS Lett. 350*, 245−248.
18. Finke, J. M., Roy, M., Zimm, B. H., and Jennings, P. A. (2000) *Biochemistry 39*, 575−583.
19. Heidary, D. K., Gross, L. A., Roy, M., and Jennings, P. A. (1997) *Nat. Struct. Biol. 4*, 1−10.
20. Bevington, P. R. (1969) *Data Reduction and Error Analysis for the Physical Sciences*, Vol. 1, McGraw-Hill, New York.

21. Zettlmeissl, G., Rudolph, R., and Jaenicke, R. (1979) *Biochemistry 18*, 5567−5571.
22. Varley, P., Gronenborn, A. M., Christensen, H., Wingfield, P. T., Pain, R. H., and Clore, G. M. (1993) *Science 260*, 1110−1113.
23. Sosnick, T. R., Shtilerman, M. D., Mayne, L., and Englander, S. W. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 8545−8550.
24. Jennings, P. A. (1998) *Nat. Struct. Biol. 5*, 846−848.
25. Harper, J. D., and Lansbury, P. T., Jr. (1997) *Annu. Rev. Biochem. 66*, 385−407.
26. Craig, S., Schmeissner, U., Wingfield, P., and Pain, R. H. (1987) *Biochemistry 26*, 3570−3576.
27. King, J., Haase-Pettingell, C., Robinson, A. S., Speed, M., and Mitraki, A. (1996) *Faseb J. 10*, 57−66.
28. Silow, M., Tan, Y. J., Fersht, A. R., and Oliveberg, M. (1999) *Biochemistry 38*, 13006−13012.
29. Plomer, J. J., and Gafni, A. (1993) *Biochim. Biophys. Acta 1163*, 89−96.

BI001307D