Aggregation Events Occur Prior to Stable Intermediate Formation during Refolding of Interleukin 1β†

John M. Finke, Melinda Roy, Bruno H. Zimm, and Patricia A. Jennings*

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

Received July 1, 1999; Revised Manuscript Received October 26, 1999

ABSTRACT: A point mutation, lysine 97 → isoleucine (K97I), in a surface loop in the β-sheet protein interleukin 1β (IL-1β), exhibits increased levels of inclusion body (IB) formation relative to the wild-type protein (WT) when expressed in *Escherichia coli*. Despite the common observation that less stable proteins are often found in IBs, K97I is more stable than WT. We examined the folding pathway of the mutant and wild-type proteins at pH 6.5 and 25 °C with manual-mixing and stopped-flow optical spectroscopy to determine whether changes in the properties of transiently populated species in vitro correlate with the observation of increased aggregation in vivo. The refolding reactions of the WT and K97I proteins are both described by three exponential processes. Two exponential processes characterize fast events (0.1−1.0 s) in folding while the third exponential process correlates with a slow (70 s) single pathway to and from the native state. The K97I replacement affects the earlier steps in the refolding pathway. Aggregation, absent in the WT refolding reaction, occurs in K97I above a critical protein concentration of 18 μM. This observation is consistent with an initial nucleation step mediating protein aggregation. Stopped-flow kinetic studies of the K97I aggregation process demonstrate that K97I aggregates most rapidly during the earliest refolding times, when unfolded protein conformers remain highly populated and the concentration of folding intermediates is low. Folding and aggregation studies together support a model in which the formation of stable folding intermediates afford protection against further K97I aggregation.

Understanding the factors responsible for the aggregation of expressed proteins is a problem of importance not only in biotechnology but also in the health-related industries. Deposition of insoluble protein aggregates results in the formation of inclusion bodies during the bacterial expression of recombinant proteins (*3*−*8*) as well as a variety of pathological conditions in mammals, including Alzheimer's and other amyloid-related disease states (*3, 9*−*12*). A common characteristic of protein deposition diseases, as well as inclusion body formation during bacterial expression, is the sensitivity of the extent of aggregation to point mutations in the protein of interest (*3, 4, 11, 13*). The common belief is that single amino acid replacements creating off-pathway association events are mediated either by intermediates or by native states, as a single mutation is thought to be unlikely to be able to alter significantly the unfolded biopolymer (*14*). However, it is known that mutations can and do affect the stability and transient structure of the unfolded state (*15*). Thus, the effects of point mutations on aggregation formed during protein folding may help identify the species and factors involved in protein deposition.

To address fully protein aggregation phenomena, detailed structural and biochemical information on the protein being studied must be available in addition to a consistent method of modulating aggregation. Consequently, recombinant human interleukin 1β (IL-1β)[1] is an excellent model system for characterizing the effects of point mutations on the competition between on-pathway folding events and off-pathway aggregation. The wild-type protein (WT) is found predominantly (>90%) in the soluble fraction of whole-cell lysates when expressed at 37 °C in *Escherichia coli*, and a number of point mutations have been identified that show dramatic effects on the levels of inclusion bodies formed (>90% insoluble) (*4, 16, 17*). Of particular interest is the mutation in the surface loop 86−99 that replaces the positively charged lysine side chain at position 97 with the hydrophobic side chain of isoleucine (*18*; see Figure 1 and Materials and Methods). This mutation K97I results in a native state that is thermodynamically more stable than the wild-type protein yet is found predominantly in inclusion bodies when expressed at 37 °C in *E. coli*. No alteration in the kinetics of the slow step in folding, nor the ribosomal synthesis rates, nor protein−chaperone interactions have been found with this mutation (*17*). FTIR analysis indicates that

† This work supported by Hellman Faculty Fellowship (P.A.J.), Sloan Fellowship (P.A.J.), NIH Grant GM54038 (P.A.J.), and Public Health Service Grant CA09523 (J.M.F. and M.R.).

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

[1] Abbreviations: FTIR, Fourier transform infrared spectroscopy; IPTG, isopropyl β-D-thiogalactopyranoside; PMSF, phenylmethanesulfonyl fluoride; EDTA, ethylenediaminetetraacetic acid; DTT, dithiothreitol; βme, 2-mercaptoethanol; Tris-HCl, tris(hydroxymethyl)-aminomethane hydrochloride; MES, 2-(4-morpholino)ethanesulfonic acid; GdnHCl, guanidine hydrochloride; Gdn-SCN, guanidine thiocyanate; FPLC, fast protein liquid chromatography; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; ESI-MS, electrospray ionization mass spectrometry; IL-1β, interleukin-1β.

10.1021/bi991518m CCC: $19.00
Published on Web 12/28/1999

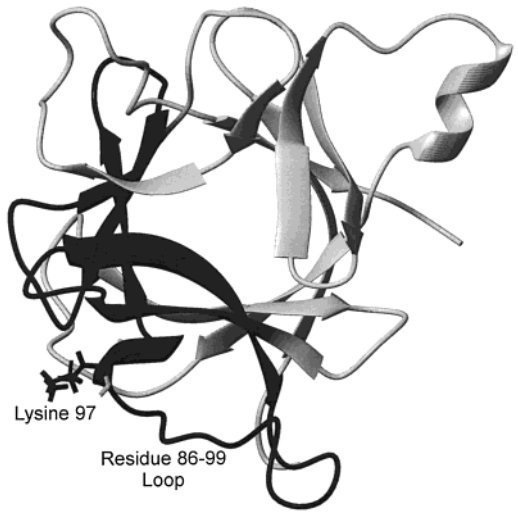


FIGURE 1: MOLMOL representation (*44*) of the solution structure of human interleukin 1β (*45*) down the β-barrel with secondary structure elements as defined by Kabsch and Sander (*46*). Residues demonstrating early amide-proton protection are shown as dark strands. Residues demonstrating late amide-proton protection are shown as light strands. The residue 86−99 loop microdomain is indicated and the extended side chain of lysine 97 is shown.

K97I inclusion bodies in vivo and aggregates in vitro contain a significant and similar degree of secondary structure, yet both are different from the native state (*19*).

We now report the results of manual-mixing and stopped-flow optical experiments on the folding reaction of WT and K97I at pH 6.5 and 25 °C. Using kinetic analyses of both folding and aggregation, we have determined that K97I self-associates when unfolded conformers predominate and that formation of stable folding intermediates protects against further association. We have proposed structural details of the initial aggregation stage and how aggregation is subsequently prevented during the folding process, a prerequisite to any further studies of the final insoluble aggregates.

MATERIALS AND METHODS

Expression and Purification of IL-1β and K97I. The gene encoding either wild-type IL-1β (WT) or the K97I sequence variant were subcloned into a pET 24-d vector (Novagen) and transformed into *E. coli* BL21(DE3) cells. Cells were grown in LB to an OD_{600} of 0.8 and protein expression was induced with IPTG at a final concentration of 1 mM. Rifampicin was added 45 min after induction to a final concentration of 0.1 mg/mL. Three hours after induction, cells were harvested by centrifugation at 10000*g* for 30 min.

Purification of IL-1β was based on the procedure of Meyers et al. (*20*) but with a number of modifications. Cells were resuspended in 10 mM KPO_4, 0.2 mM EDTA, 5 mM DTT, and 1 mM PMSF at pH 8.0 and then lysed by sonication at 4 °C, followed by centrifugation at 4300*g* for 30 min. Soluble IL-1β in the supernatant was made 80% saturated in ammonium sulfate and then precipitated by centrifugation as above. The pellet was then dissolved in buffer A (25 mM NH_4OAc, 2 mM EDTA, and 1 mM βme at pH 4.5) and dialyzed overnight in buffer A at 4 °C.

The dialysate was centrifuged as above, filtered, and applied to a Resource S cation-exchange column (Pharmacia) equilibrated in buffer A. IL-1β was eluted in a 40 column volume linear gradient of 25−240 mM NH_4OAc, pH 4.5, at a flow rate of 3 mL/min. Purity was judged to be greater than 95% as assessed by SDS−PAGE. Protein concentrations were calculated by using an experimentally determined $\epsilon_{280} = 11.26\ mM^{-1}\ cm^{-1}$ (*21*). Mass spectral analysis and DNA sequencing identified the mutant IL-1β as a K97I variant, and not K97V as it had been previously reported (*4*). Purified protein was dialyzed extensively into 10 mM MES buffer containing 2 mM EDTA, 90 mM NaCl, and 1 mM βme at pH 6.5 for all experiments discussed.

Equilibrium Fluorescence. Equilibrium unfolding titrations were measured by use of intrinsic tryptophan fluorescence emission intensity (*22, 23*). Protein samples were diluted to varying final denaturant concentrations and equilibrated overnight. Protein concentrations ranged from 6 to 18 μM in separate experiments. Fluorescence spectra were acquired on a Fluoromax-2 spectrofluorometer (SPEX, Edison, NJ). Fluorescence emission was measured as total intensity over the 300−450 nm emission range after excitation at 293 nm.

Manual-Mixing Fluorescence. Unfolding experiments were initiated by addition of native IL-1β stock into a cuvette to final GdnHCl concentrations ranging from 2 to 5 M and final IL-1β concentrations ranging from 6 to 24 μM. Refolding experiments were initiated by dilution of unfolded IL-1β at 2.2 M GdnHCl to final GdnHCl concentrations ranging from 0.2 to 1.4 M and final IL-1β concentrations ranging from 6 to 36 μM. The kinetics of folding reactions with relaxation times greater than 10 s were measured by the time-dependent change in fluorescence emission at 343 nm (slit 1 mm) while excitation was at 293 nm (slit 1 mm) with a Fluoromax-2 spectrofluorometer equipped with a Neslab RTE−111 temperature controller.

Stopped-Flow Fluorescence. Folding reactions with relaxation times less than 10 s were monitored with an Applied Photophysics SX.17MV (Applied Photophysics, London) stopped-flow unit with a path length of 0.1 cm. Refolding experiments were initiated by a 1:10 dilution of unfolded IL-1β at 2.2 M GdnHCl, except where indicated, into varying final IL-1β concentrations and final GdnHCl concentrations. Unfolding experiments used 1:1 ratio mixing sizes such that 1 part native IL-1β dilutes into 1 part GdnHCl buffer. Excitation was at 293 nm and emission was collected through a >320 nm cutoff filter. Each kinetic trace is the average of 10−20 data acquisitions.

Interrupted Refolding Experiments. The reverse double-jump experiment (*24*)

$$\text{unfolded} \xrightarrow[(t_i)]{\text{step 1}} \text{folded} \xrightarrow{\text{step 2}} \text{unfolded}$$

was performed to measure the time-dependent production of native wild-type and K97I protein. Folding (step 1) was initiated by diluting 4 parts unfolded protein in 3.0 M GdnHCl with 9 parts buffer to reach a final buffer concentration of 0.8 M GdnHCl. After various times (t_i), refolding was interrupted by transferring the solution into a final GdnHCl concentration of 4.5 M (step 2). Partially folded intermediates present immediately prior to step 2 are extremely unstable (0.8 M GdnHCl refolding conditions), unfold quickly during step 2, and are therefore not measured. However, under these conditions, the population of native IL-1β unfolds at the expected rate ($\tau_{step2} \sim 400$ s). Consequently, native IL-1β at each time point t_i can be reliably

quantitated by the amplitude of the resulting unfolding kinetic exponential τ_{step2}. All amplitudes were normalized to the amplitude recovered from step 2 on 100% native protein. No amplitude was observed when unfolded IL-1β (3.0 M GdnHCl) was diluted into the final conditions (4.5 M GdnHCl), confirming that measured amplitudes could result only from native IL-1β formed from folding after step 1 and not from extraneous processes. Spectroscopic conditions were as described under Manual-Mixing Fluorescence.

Stopped-Flow Light Scattering. Protein aggregation during refolding was monitored with an Applied Photophysics SX.17MV (Applied Photophysics, London) stopped-flow unit with a path length of 0.1 cm. Aggregation was initiated by a 1:10 dilution of IL-1β from 2.2 M GdnHCl (except where noted) to the indicated final IL-1β concentrations of protein and final GdnHCl concentrations. Aggregation was measured by light scattering at a wavelength of 500 nm and collected at a 90° angle through a >320 nm cutoff filter. Each kinetic trace is the average of 10−20 data acquisitions.

Fraction Aggregated. Refolding experiments were initiated by a 1:10 dilution of IL-1β from 2.2 M GdnHCl, except where noted, to the indicated final IL-1β concentrations of protein and 0.4 M GdnHCl. Samples were centrifuged at 7000*g* and the completeness of the aggregate separation was assessed by the absence of turbidity at 500 nm. The dissolved protein concentration was measured by the absorbance at 280 nm and calculated by use of the extinction coefficient $\epsilon_{280} = 11.26\ \mathrm{mM^{-1}\ cm^{-1}}$ and the Beer−Lambert relationship (*21*).

Data Analysis. Equilibrium data were fit as described previously (*4*). Manual mixing kinetics, stopped-flow kinetics, and interrupted refolding data were fit globally (*25*) to eq 1 using the Marquardt algorithm (*26*) and in-house software:

$$A(t) = \sum_{i}^{i} A_i \exp(-t/\tau_i) + A(\infty) \qquad (1)$$

The number of kinetic processes i, relaxation time τ_i, signal amplitude A_i of each exponential kinetic process i, and the final signal value $A(\infty)$ at equilibrium were determined using the fit quality represented in the chi-squared (χ_r^2) values (*27*), the random dispersion of residuals, and the logical consistency of the generated fitting parameters. The relaxation time τ equals the inverse of the observed rate constant (i.e., $1/k_{obs}$) and does not directly measure the microscopic rate constant for a kinetic process (*28*).

RESULTS

Equilibrium Fluorescence. The equilibrium transition for K97I IL-1β is shifted to higher denaturant concentrations, consistent with a 0.9 kcal/mol increase in thermodynamic stability of the native protein with mutation, as observed previously (*16, 18, 22*).

Fluorescence Measured Kinetics. A plot of the change in fluorescence intensity above 320 nm as a function of time for a stopped-flow refolding jump of WT IL-1β from 2.2 to 0.2 M GdnHCl is given in Figure 2A. The first 20 s of the reaction are shown for clarity. The initial increase in fluorescence intensity is characterized by two exponential processes and is followed by a slow process in which the decay in fluorescence intensity approaches the expected

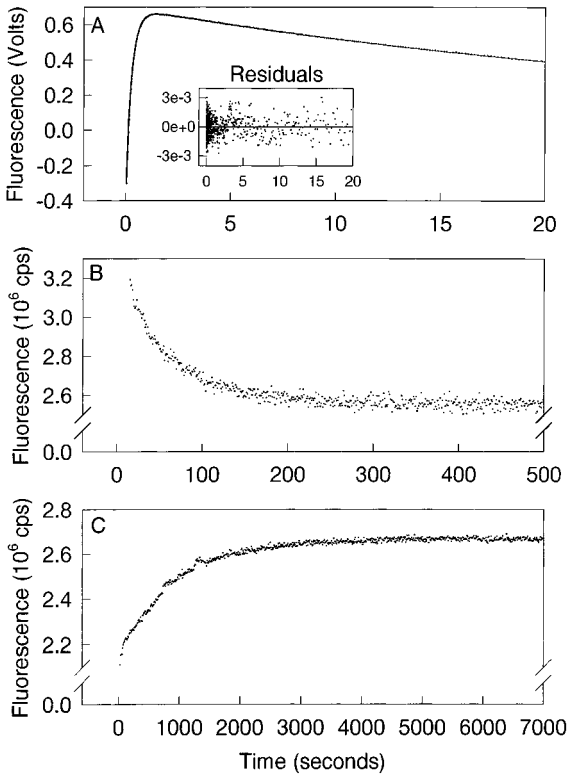


FIGURE 2: Real-time kinetic fluorescence measurements during the refolding of WT IL-1β. (A) Refolding from 2.2 to 0.2 M GdnHCl by stopped-flow mixing. Individual data points (·) are shown along with the 3-exponential fit (—). (Inset) Residuals of the kinetics fit to three exponential processes τ_1, τ_2, and τ_3. (B) Refolding from 2.2 to 0.2 M GdnHCl by manual mixing. Under these conditions, the displayed τ_1 process is characterized by a decrease in fluorescence with time. (C) Refolding from 2.2 to 1.0 M GdnHCl by manual mixing. Under these conditions, the displayed τ_1 process is characterized by an increase in fluorescence with time. Given the change in sign for the amplitude of τ_1 between panels B and C, interpretation of amplitude changes over GdnHCl concentrations for the three kinetic processes τ_1, τ_2, and τ_3 will be ambiguous.

equilibrium value. This nonmonotonic change in fluorescence intensity as a function of time measures the population of a highly fluorescent intermediate formed prior to the native protein and is consistent with the finding of a populated intermediate with pulse-labeling techniques (*1, 2*). The calculated fit of the data to a three-exponential process is shown for comparison (Figure 2A). The plot of the residual errors for the fit of the data to a three-exponential equation is given in the inset to Figure 2A. As demonstrated in Figure 2 panels B (0.2 M GdnHCl) and C (1.0 M GdnHCl), the slowest process has either a negative (Figure 2B) or positive (Figure 2C) fluorescence amplitude, depending upon the final denaturant concentration in the refolding jumps. This observation is consistent with the changes in equilibrium fluorescence measurements of the native state observed at 0.2 and 1.0 M GdnHCl, respectively. Global analysis of both WT and K97I *re*folding data consistently fit best to three exponential processes and are designated as τ_3 (fast), τ_2 (medium), and τ_1 (slow). Analysis of both WT and K97I *un*folding data under all conditions consistently fit best to one exponential process and is designated τ_1(slow). The two faster processes observed in the refolding reactions, τ_3 and τ_2, were not observed in unfolding by either manual-mixing or stopped-flow techniques (data not shown).

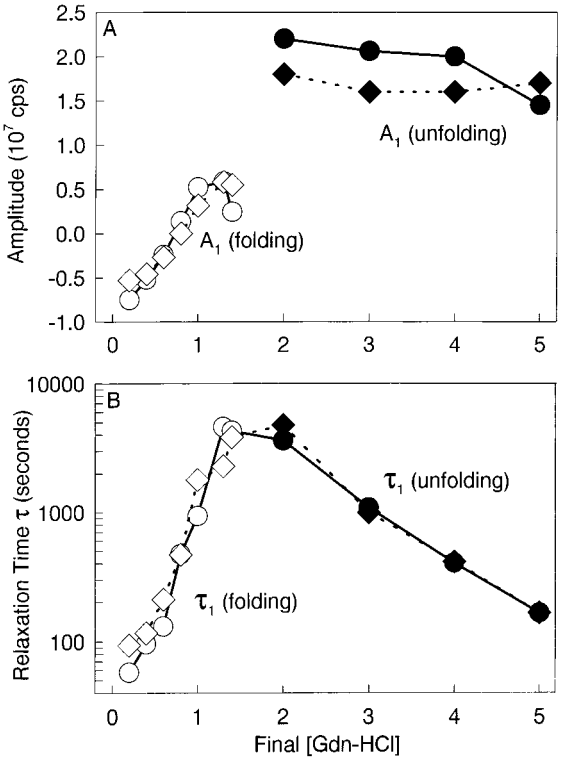


FIGURE 3: Manual mixing kinetic fluorescence measurement of folding and unfolding kinetic parameters for WT A_1/τ_1 refolding (○), WT A_1/τ_1 unfolding (●), K97I A_1/τ_1 refolding (◇), and K97I A_1/τ_1 unfolding (◆). (A) Amplitude A_1 of τ_1 folding/unfolding process with respect to final GdnHCl concentration. (B) Relaxation time of τ_1 folding/unfolding process with respect to final GdnHCl concentration. The values of τ_1 are essentially identical for WT and K97I, as observed previously (*17*).

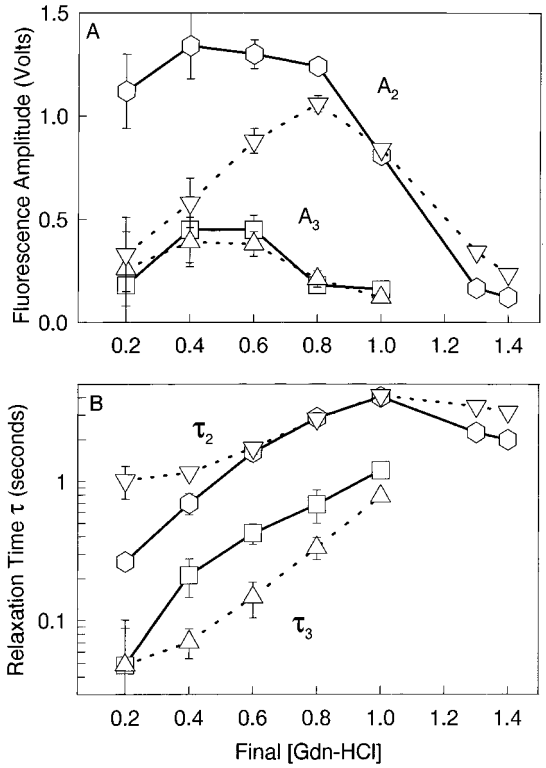


FIGURE 4: Stopped-flow kinetic fluorescence measurements for refolding kinetic parameters WT A_3/τ_3 (□), WT A_2/τ_2 (⬡), K97I A_3/τ_3 (△), and K97I A_2/τ_2 (▽). Error bars represent 95% confidence intervals. WT and K97I A_3/τ_3 are not shown above 1.0 M GdnHCl because the amplitude A_3 can no longer be resolved. (A) Amplitude (in PMT volts) of kinetic folding processes WT A_3, WT A_2, K97I A_3, and K97I A_2. (B) Relaxation time of WT τ_3, WT τ_2, K97I τ_3, and K97I τ_2 folding processes with respect to final GdnHCl concentration. The values of K97I τ_2 do not follow simple denaturant concentration dependence.

Fluorescence-detected folding kinetics for both WT and K97I were performed over a series of denaturant concentrations and 18 μM protein. The observed amplitude A_1 and slow relaxation time τ_1, determined in manual-mixing experiments as a function of final denaturant concentration, are given in Figure 3. The refolding relaxation time τ_1 varies from 30 s under strongly refolding conditions (0.2 M GdnHCl) to 1900 s in the transition region (1.0 M GdnHCl), where solution conditions support the formation of a highly fluorescent nativelike species (*18*). Unfolding relaxation times vary from 160 s under strongly unfolding conditions (5 M GdnHCl) to 3100 s in the transition region (2 M GdnHCl). Neither the observed relaxation times nor the fluorescence amplitudes for the slow τ_1 process differ appreciably between the WT and K97I mutant proteins, consistent with previously published observations (*16*).

In stopped-flow fluorescence experiments, an initial increase in fluorescence intensity observed in refolding experiments at 18 μM protein is described by two exponential process, τ_3 and τ_2. Interpretation of the GdnHCl dependence of the amplitudes of τ_3 and τ_2, A_3 and A_2, is unreliable because the amplitude A_1 changes from negative to positive at increasing GdnHCl concentrations (Figure 2A) and affects the fitted values of A_3, A_2, and A_1. Therefore, a qualitative analysis of fluorescence amplitudes follows. The observed amplitudes A_3 and A_2 for both the WT and K97I protein are shown as a function of final denaturant concentration in Figure 4A. The major difference in amplitudes of the fast processes between WT and K97I is in A_2. K97I A_2 is much lower than WT A_2 at 0.2 M GdnHCl but increases to the value of the WT A_2 at 0.8 M GdnHCl and behaves similarly to WT at higher GdnHCl concentrations. Although a quantitative interpretation of the amplitudes is precluded (see above), it is nonetheless clear that the major amplitude differences between K97I and WT occur in A_2.

In Figure 4B, the WT relaxation times τ_3 and τ_2 at 18 μM protein display the characteristic denaturant concentration dependence expected for a protein folding reaction (*29*). K97I τ_3 is slightly accelerated with respect to WT τ_3 and increases logarithmically from 50 ms at 0.2 M GdnHCl to 780 ms at 1.0 M GdnHCl. In contrast, K97I τ_2 displays kinetic behavior that does not follow simple denaturant concentration dependence. K97I τ_2 displays a relaxation time of approximately 1 s, which is independent of GdnHCl concentrations below 0.6 M, as opposed to WT τ_2, which decreases to 260 ms at 0.2 M GdnHCl. Repeated experiments (4×) with different protein preparations demonstrate that differences between WT and K97I are reproducible and statistically significant, as shown by 95% confidence intervals in Figure 4. Above 0.6 M GdnHCl, K97I τ_2 is denaturant-dependent and similar in rate to WT τ_2. For both WT and K97I above 1.0 M GdnHCl, the τ_3 process is no longer observable and τ_2 reaches a maximum near 4 s.

Interrupted Refolding Experiments. At pH 5.0, the fast processes in the refolding reaction of the WT protein are not a result of parallel paths to the native protein (*2*). To

test whether native WT and K97I also accumulate from a single folding route at pH 6.5, we performed interrupted refolding experiments (*24*). The time-dependent recoveries of the native-state amplitude during step 2 of the interrupted folding experiment (see Materials and Methods) for both WT and K97I are each accurately fit to a single-exponential process. The relaxation time for this process is 420 s for the WT and 370 s for K97I (data not shown), consistent with the slowest refolding process τ_1, as seen in Figure 3B. If a fast refolding process (τ_3 or τ_2) also led to native protein production, a significant amount of native-state amplitude would be observed during step 2 at the earliest time point (t_i = 10 s), which is not the case here.

Variation of Initial Denaturant Concentrations. To address whether residual structure in the unfolded state under high denaturant concentration is responsible for the observed differences in the τ_2 process for the WT and K97I proteins, a series of experiments were performed at 18 μM protein in which the *initial* GdnHCl concentrations in the refolding reactions were varied. The amplitudes and relaxation times for the refolding processes for both WT and K97I are independent of the initial unfolding conditions.

Aggregation. The final protein concentration in refolding was varied to determine whether protein association could account for the observed differences in the A_2 amplitude and the τ_2 relaxation time at 0.4 M final GdnHCl for K97I as compared to WT (Figure 4). We observed protein association during K97I refolding by two detection methods: (1) fluorescence at 293 nm excitation, 320−400 nm emission, and (2) light scattering at 500 nm.

When the refolding relaxation times τ_3 and τ_2 of WT and K97I are observed by stopped-flow fluorescence with 0.4 M final GdnHCl, only the τ_2 process of K97I is dependent on protein concentration (Figure 5). The τ_2 rate decreases 2-fold and the normalized amplitude A_2 (PMT volts/μM) decreases 40% as the protein concentration is increased from 3 to 42 μM.

Protein association during refolding in 0.4 M final GdnHCl was followed more directly by stopped-flow light scattering at 500 nm. A plot of the signal change at 500 nm as a function of refolding time for WT and K97I is given in Figure 6. As IL-1β contains no chromophores that either absorb or emit light at this wavelength, any observed signal change can be attributed to particulate formation as a result of protein aggregation. Over the range of protein concentrations measured, 3−48 μM, WT did not show any appreciable increase in light scattering. At 18 μM K97I, there is no measurable change in the optical signal with time. However, at higher protein concentrations, K97I displays measurable light scattering occurring within the first seconds of acquisition.

Light scattering behavior during refolding was also studied while first varying the initial GdnHCl concentration (Figure 7A) and then varying the final GdnHCl concentration (Figure 7B) with a constant protein concentration of 0.55 mg/mL K97I. Variations in the *initial* GdnHCl did not affect the light scattering behavior of K97I at either 0.4 or 0.6 M final GdnHCl (Figure 7A). This is consistent with the independence of the folding process to initial unfolding conditions as monitored by fluorescence (see above). However, increasing the *final* GdnHCl concentration in the refolding process demonstrated a reduction in the light scattering amplitude, which is due to a drop in aggregation (Figure 7B).

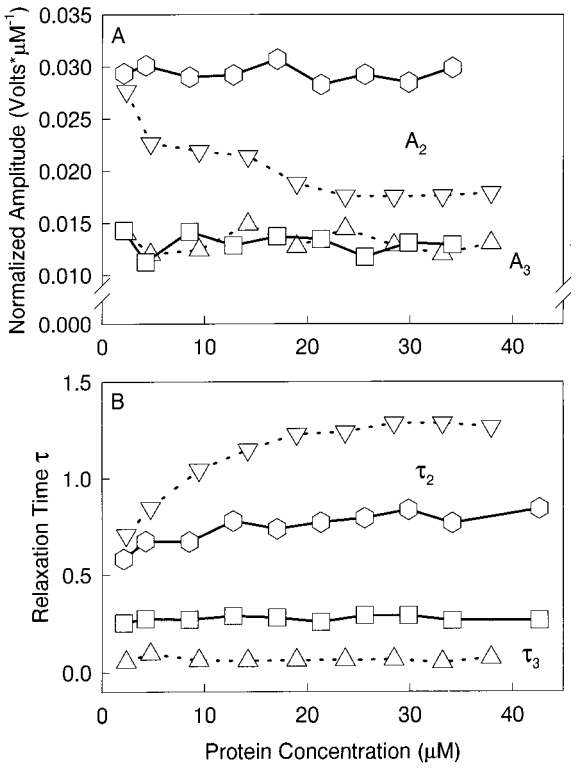


FIGURE 5: Effect of increasing protein concentration from 3 to 48 μM on measured folding parameters WT A_3/τ_3 (□), WT A_2/τ_2 (⬡), K97I A_3/τ_3 (△), and K97I A_2/τ_2 (▽). (A) Amplitudes A_3 and A_2 of respective kinetic processes τ_3 and τ_2 measured during the refolding of WT and K97I at 3−48 μM protein. (B) Refolding relaxation times τ_3 and τ_2 measured during the refolding of WT and K97I at 3−48 μM protein. K97I amplitude A_2 decreases 2-fold and K97I relaxation time τ_2 increases 2-fold as concentration is increased, indicating an association process during K97I refolding not present in WT refolding.

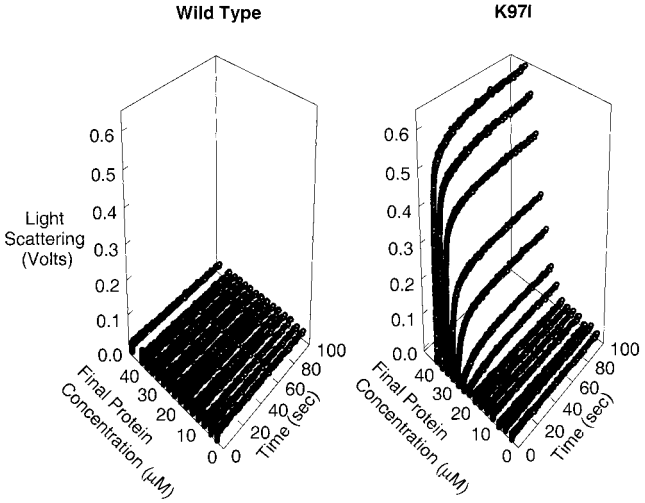


FIGURE 6: Stopped-flow light scattering behavior during refolding of WT and K97I at increasing protein concentrations from 3 to 48 μM. (Left panel) Time dependence of light scattering in WT refolding versus protein concentration. (Right panel) Time dependence of light scattering in K97I refolding versus protein concentration. Both panels A and B are represented on the same scale. At a final GdnHCl concentration of 0.4 M GdnHCl and protein concentrations >18 μM, K97I exhibits light scattering behavior consistent with aggregation.

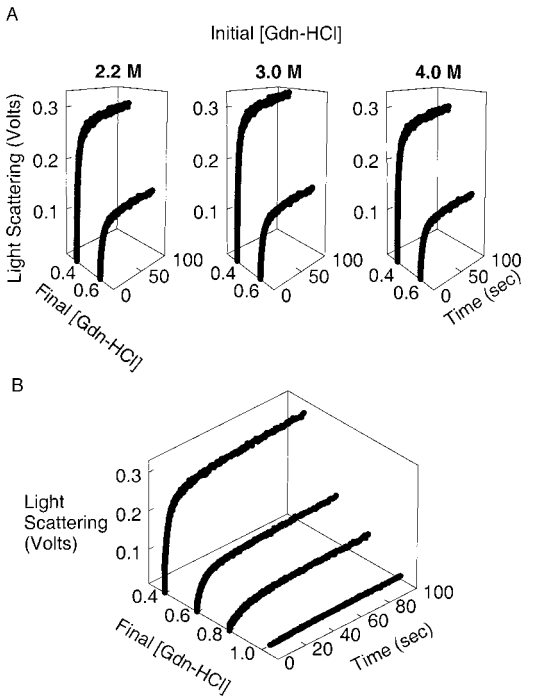


FIGURE 7: Effect of initial and final GdnHCl concentration during refolding of K97I on stopped-flow light scattering behavior. (A) Light scattering at initial GdnHCl concentrations 2.2, 3.0, and 4.0 M and final concentrations 0.4 and 0.6 M. (B) Light scattering at an initial GdnHCl concentration of 2.2 M and final concentrations 0.4, 0.6, 0.8, and 1.0 M. Initial GdnHCl concentration does not affect the extent of aggregation, while increasing the final GdnHCl decreases aggregation.

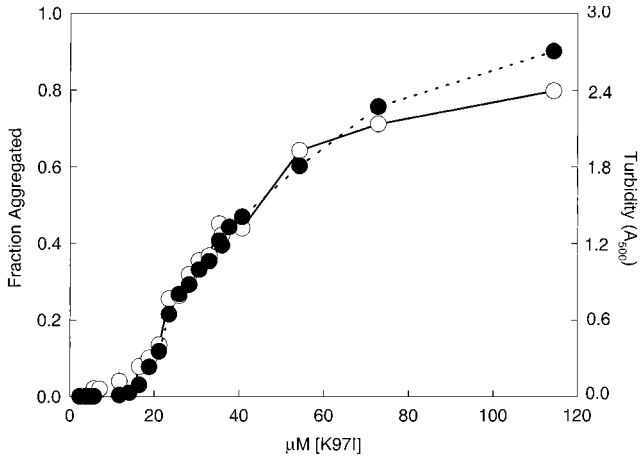


FIGURE 8: Fraction of K97I aggregated in post-refolding equilibrium relative to soluble K97I as a function of the total concentration of K97I. The fraction of aggregated K97I was determined quantitatively through UV absorbance at 280 nm of the dissolved protein (○) and qualitatively through solution turbidity at 500 nm (●). The relatively low fraction of K97I aggregate (<50%) during refolding experiments supports an aggregation model mediated by an early, transiently populated species.

Aggregation was quantified by measuring the fraction of K97I found in the dissolved and precipitated fractions after the completion of folding at 0.4 M final GdnHCl. The percentage of protein found in the precipitated fraction was assayed by the observed sample turbidity at 500 nm and by quantitating the amount of protein remaining in the soluble portion (Figure 8). There is good agreement between the two sets of data, which indicates a cooperative transition between the soluble and insoluble percentage as a function of total K97I concentration. The percent of K97I aggregation at any final protein concentration in Figure 8 has been shown to be independent of the initial concentration of unfolded protein prior to dilution and refolding (data not shown).

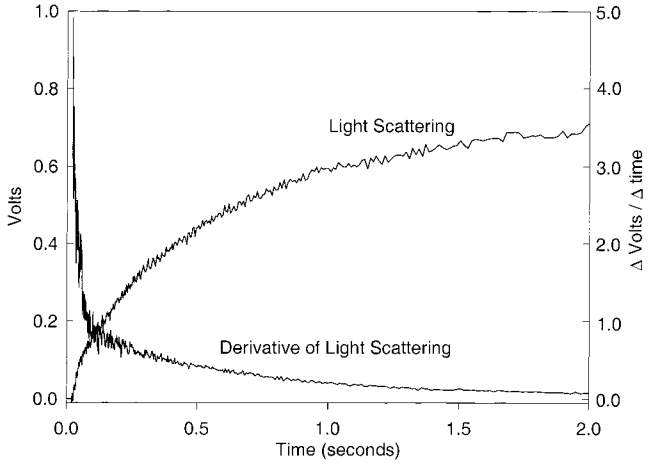


FIGURE 9: Analysis of stopped-flow light scattering data during refolding of K97I to 0.4 M GdnHCl and 42 μM K97I: Light scattering intensity in volts over 0−2 s after mixing and time derivative of light scattering in Δvolts/Δtime over 0−2 s after mixing where Δtime interval = ±0.005 s. The decrease of the time derivative, taken together with the observed folding rates, supports a model in which aggregation takes place before stable IL-1β structure formation.

DISCUSSION

Protein aggregation is a problem in the recovery of bioactive recombinant proteins and may also lead to pathological conditions in vivo and in vitro. Although many proteins aggregate under native conditions, it may be that pathological or recombinant protein aggregation reactions are mediated through association of a nonnative species (*7*, *11*, *12*, *19*, *30*−*38*). If nonnative states are responsible for aggregation, it is essential that they be characterized. In the current study, we performed kinetic experiments under conditions where folding and aggregation compete to gain a better understanding of these processes.

Aggregation Occurs Prior to Stable Intermediate Formation. At elevated protein concentrations the mutation of K97 to isoleucine affects early events in the folding reaction of IL-1β (Figures 4 and 5), particularly the τ_2 folding process and the extent of self-association. With regard to increasing protein concentration, the correlation between the lowered amplitude and rate of the τ_2 refolding process (Figure 5) and the increasing amplitude of light scattering (Figure 6) indicates that folding and aggregation are competing processes in K97I.

Figure 9 shows the intensity of light scattering and its time derivative plotted as a function of time. The time derivative of the observed light scattering signal is directly proportional to the concentration and the square of the molecular weight of the species involved in aggregation (*39*). The aggregation rate should be maximal under conditions where the aggregating species is most highly populated, as the aggregation reaction is at least second-order in the concentration dependence. The time derivative of the light scattering intensity is dependent on the aggregation rate, and consequently the concentration of the associating species. Therefore, if ag-

gregation occurs when the unfolded populations predominate, the time derivative is expected to start from its maximum value and subsequently fall as the species depopulates. If an intermediate (I) aggregates, the time derivative is expected to rise and fall in concert with the transient population of this species. If the native state (N) aggregates, the time derivative should initially be zero and increase with a rate reflecting the population of the native protein.

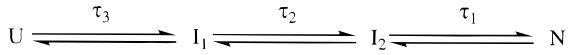
By use of the measured folding rates (Figure 4A) and the light scattering data (Figure 6) at 42 μM K97I and 0.4 M GdnHCl, the most probable aggregating species can be identified. The light scattering time derivative falls from an initial maximal value to nearly zero at 2 s after the initiation of folding (Figure 9). This same result is found at all protein and GdnHCl concentrations investigated. Consequently, the species whose predicted aggregation explains these observations is the unfolded ensemble U. The aggregation of U is also supported by the observation that refolding at 0.4 M GdnHCl and 42 μM K97I produces a relatively low fraction (<50%) of aggregated protein (Figure 8), since a less transient species would be expected to partition a larger fraction of protein to the aggregated state. Since the initial guanidine concentration does not affect the observed aggregation kinetics (Figure 7A) and the unfolded protein in high denaturant is not observed to associate, it is the properties of the unfolded ensemble under native conditions that lead to aggregation. While folding events in proteins can occur on a microsecond time scale resulting in the formation of an early species in the dead time of stopped-flow mixing, in the case of IL-1β, the entirety of the signal between unfolded and native folding is accounted for in rapid quench/pulse labeling experiments as well as stopped-flow/circular dichroism kinetic measurements (*2*). There is no evidence suggesting production of a stable structured intermediate in the dead time of mixing. However, it is possible that an unstructured folding intermediate or hydrophobic collapse of the unfolded protein (*40*) could occur in the dead time and be involved in aggregation.

K97I Aggregation Is Consistent with the Nucleation Model. Aggregation is absent during refolding to protein concentrations at and below 18 μM K97I in 0.4 M GdnHCl (Figures 6 and 8). However, aggregation is observed by light scattering at final protein concentrations of 21 μM and above in 0.4 M GdnHCl. Thus, there is a critical protein concentration above which protein aggregation occurs. Similar critical concentrations are observed in both refolding experiments (*32*, *35*, *38*, *41*) and in the static aggregation of the Aβ peptide (*42*). To explain this phenomenon, Harper and Lansbury (*42*) have proposed the necessity of nucleation, i.e., the thermodynamically unfavorable formation of a nucleus which, having attained a certain size, is stable and readily incorporates additional monomeric protein units into a growing aggregate.

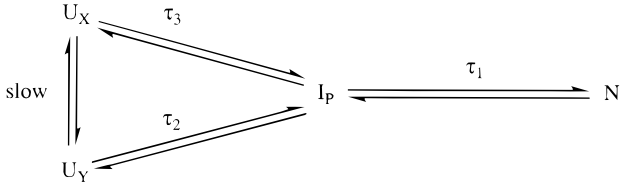
Nucleation-initiated aggregation exhibits no detectable aggregation at low protein concentrations followed by a dramatic increase in visible aggregation immediately above the *critical concentration*; that is, the concentration at which substantial numbers of nuclei have formed. Consequently, the observation of a critical concentration between 18 and 21 mM in K97I aggregation supports the nucleation-growth model for the K97I association process. Furthermore, the nucleation model necessarily requires that a number of multimers between the monomer and the stable aggregate are linked through reversible solution equilibria (*42*). These reversible off-pathway multimeric intermediates would be expected to alter folding rates. Consistent with this nucleation model, the K97I τ_2 folding rate is observed to slow at increasing K97I concentrations (Figure 5B).

Aggregation in Context of the Folding Model. At low protein concentrations, WT and K97I display similar multiphasic refolding kinetics over a range of denaturant concentrations at pH 6.5 and 25 °C. Therefore, the folding mechanism is not likely altered by the mutation. The two fast processes (τ_2 and τ_3) are associated with the population of highly fluorescent intermediate species (Figure 2A). Interrupted refolding experiments indicate that only the slow τ_1 process results in the formation of the native protein. Thus, there are no parallel pathways to N. These results are consistent with either the folding model given by Scheme 1 or Scheme 2. Under strongly folding conditions, the forward reactions predominate, while under strongly unfolding conditions, the reverse reactions dominate. The effect of denaturant on the forward and reverse rate constants can yield a complex effect on an observed τ as the denaturant concentration is changed (*28*, *29*). However, a rigorous analysis of the folding reaction under all denaturant concentrations, which would yield microscopic rate constants, is precluded in this case because of the opposing fluorescence amplitudes of the τ_3/τ_2 and τ_1 reactions (see Results, Figure 2A). Nevertheless, our three observable rate processes indicate a minimum of four species involved in folding. The observed τ values under strongly refolding conditions approximate the microscopic rates and can be associated with three specific transitions.

Scheme 1



Scheme 2



With the appropriate modification of kinetic rates, either Scheme 1 or Scheme 2 can explain the observed concentration dependence of folding rates (Figure 5B) and aggregation (Figures 6 and 8) in K97I. However, in either scheme, aggregation would initiate from the unfolded ensemble. In either Scheme 1 or 2, the aggregation rate decreases as intermediates accumulate on the time scale of the τ_3 process (Figure 9). A simple model of K97I folding and aggregation is shown in Figure 10A. The unfolded ensemble, with the hydrophobic isoleucine at position 97, competes between formation of folding intermediates at rates τ_3 and τ_2 and rapid association in a nucleus that, once formed, will irreversibly lead to larger aggregates. In Figure 10B, wild-type IL-1β containing a charged lysine at position 97 deters association of the unfolded form to a nucleus and thereby prevents aggregation.

Intermediate Formation and Commitment to Folding. Recent experimental evidence has supported the idea of

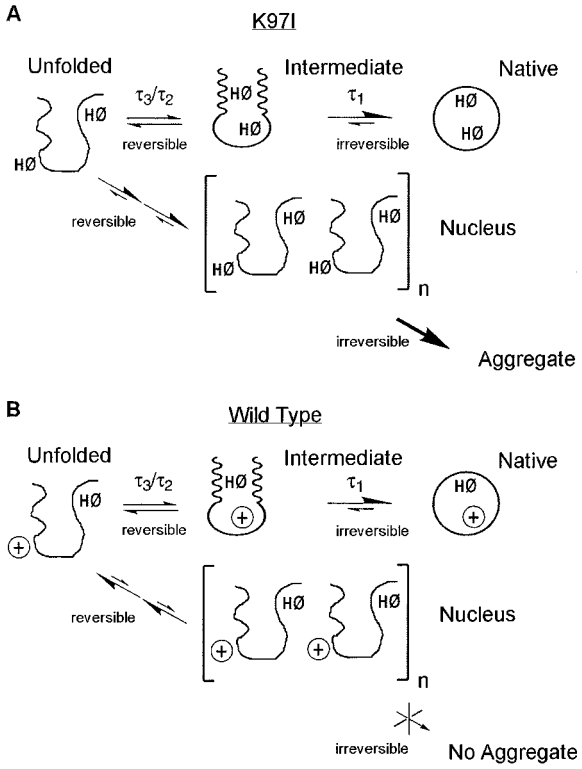


FIGURE 10: Aggregation model of K97I. (A) Unfolded K97I competes between folding and aggregation due to the presence of an isoleucine in the hydrophobic loop cluster, which facilitates self-association. Formation of the intermediate protects K97I from aggregation. (B) The lysine 97 in WT IL-1β prevents association of the unfolded species by contributing an unfavorable charged group to the hydrophobic cluster.

aggregation-sensitive folding intermediates (*13*, *14*, *30*, *36*). A simple interpretation of K97I data would have predicted aggregation through a folding intermediate, since aggregation appears to correlate with the reduction in the τ_2 folding rate but not the earlier τ_3 folding rate (Figure 5B). However, in this case the accumulation of K97I intermediates correlates with a complete reduction in the time derivative of the light scattering, supporting the hypothesis that intermediate formation protects the protein from further aggregation from the unfolded state.

Structural Interpretation. A picture of early IL-1β intermediate structure emerges from pulse-labeling hydrogen-exchange experiments analyzed with NMR (*1*) and mass spectrometric methods (*2*).[2] At 25 °C, amide hydrogens are distinctly protected within 1 s of folding in β-strands 6−10 (Figure 1, dark strands). Within this intermediate structure are a number of hydrophobic residues within the 86−99 loop, which form a structured microdomain (*18*) containing the K97I mutation site (Figure 1). In the subsequent formation of native structure, the remainder of the protein is sequestered (Figure 1, light strands).

Prior to forming intermediates, the unfolded ensemble may exist with residues 86−99 sampling a hydrophobic microdomain. The mutation K97I would increase the available hydrophobic surface area available for protein association in this loop (Figure 10A), whereas the wild-type lysine would serve as a positively charged deterrent (Figure 10B). Once a structured intermediate is formed, this hydrophobic microdomain may pack against strands 6−10 and no longer have a propensity to associate. FTIR studies indicate that the final K97I aggregates contain a high degree of nonnative β-sheet structure (*36*). The unfolded protein may acquire β-sheet structure through both the association reaction and the folding of β-strands involved in the aggregation nucleus. A similar aggregating structure has been supported in amyloid formation from the Alzheimer's peptide Aβ (*43*). Monomeric Aβ contains exposed nonnative hydrophobic structure even though Aβ amyloid fibrils contain a high degree of β-sheet structure.

IL-1β appears to belong to a class of proteins that form amorphous aggregates from the association of species in the unfolded ensemble. This behavior contrasts with the ordered fibrillar aggregates obtained from amyloidogenic proteins, which appear to populate aggregating species slowly under equilibrated conditions and subsequently to folding. Whether this difference exists due to the kinetic difficulty of organizing transient unfolded structures into amyloid precursors or from the inherent structures of the building blocks themselves poses an interesting challenge to future research.

[2] Studies of the dependence of τ on pH and temperature (J. M. Finke and P. A. Jennings, unpublished results), indicate that the intermediate in this study corresponds to the intermediate X_3 in Varley et al. (*1*); τ = 1 s at pH 5 and 4 °C) and the intermediate detected in Heidary et al. (*2*); τ = 126 ms at pH 5 and 25 °C).

ACKNOWLEDGMENT

We thank Ron Wetzel and SmithKline-Beecham Pharmaceuticals for the human IL-1β gene, J. Adams for the use of his stopped-flow fluorometer, P. Wright for the use of both his stopped-flow and static CD spectrometers, O. Bilsel for providing his global analysis software, L. Gross for mass spectrometric analysis, Hoang Ho for fraction aggregated studies, and B. Chrunyk and D. Schultz for helpful discussions and critical reviews of the manuscript.

REFERENCES

1. Varley, P., Gronenborn, A. M., Christensen, H., Wingfield, P. T., Pain, R. H., and Clore, G. M. (1993) *Science 260*, 1110−1113.
2. Heidary, D. K., Gross, L. A., Roy, M., and Jennings, P. A. (1997) *Nat. Struct. Biol. 4*, 1−10.
3. 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.
4. Chrunyk, B. A., Evans, J., Lillquist, J., Young, P., and Wetzel, R. (1993) *J. Biol. Chem. 268*, 18053−18061.
5. Carlson, A., and Reddy, J. (1991) in *Protein Refolding* (Georgiou, G., and De Bernardez-Clark, E., Eds.) pp 132−152, American Chemical Society, Washington, DC.
6. Klein, J., and Dhurjati, P. (1995) *Appl. Environ. Microbiol. 61*, 1220−5.
7. Haase-Pettingell, C. A., and King, J. (1988) *J. Biol. Chem. 263*, 4977−83.
8. Valax, P., and Georgiou, G. (1991) in *Protein Refolding* (Georgiou, G., and De Bernardez-Clark, E., Eds.) pp 96−109, American Chemical Society, Washington, DC.
9. Wood, S. J., Maleeff, B., Hart, T., and Wetzel, R. (1996) *J. Mol. Biol. 256*, 870−877.
10. Jarrett, J. T., Berger, E. P., and Lansbury, P. T., Jr. (1993) *Ann. N.Y. Acad. Sci. 695*, 144−8.
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*, 787−93.

12. Cohen, F. E., Pan, K. M., Huang, Z., Baldwin, M., Fletterick, R. J., and Prusiner, S. B. (1994) *Science 264*, 530−1.
13. 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.
14. Wetzel, R. (1996) *Cell 86*, 699−702.
15. Saab-Rincon, G., Gualfetti, P. J., and Matthews, C. R. (1996) *Biochemistry 35*, 1988−1994.
16. Chrunyk, B. A., Evans, J., and Wetzel, R. (1993) in *Protein Folding In Vivo and In Vitro* (Cleland, J. L., Ed.) pp 46−58, American Chemical Society, Washington, DC.
17. Wetzel, R., and Chrunyk, B. A. (1994) *FEBS Lett. 350*, 245−248.
18. Chrunyk, B. A., and Wetzel, R. (1993) *Protein Eng. 6*, 733−738.
19. Oberg, K., Chrunyk, B. A., Wetzel, R., and Fink, A. L. (1994) *Biochemistry 33*, 2628−34.
20. Meyers, C. A., Johanson, K. O., Miles, L. M., McDevitt, P. J., Simon, P. L., Webb, R. L., Chen, M.-J., Holskin, B. P., Lillquist, J. S., and Young, P. R. (1987) *J. Biol. Chem. 262*, 11176−11181.
21. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
22. Craig, S., Schmeissner, U., Wingfield, P., and Pain, R. H. (1987) *Biochemistry 26*, 3570−3576.
23. Finn, B. E., Chen, X., Jennings, P. A., Saalau-Bethel, S. M., and Matthews, C. R. (1992) in *Protein Engineering − A Practical Approach* (Rees, A. R., Wetzel, R., and Sternberg, J. E., Eds.) pp 167−189, IRL Press, Oxford, England.
24. Kiefhaber, T. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 9029−9033.
25. Knutson, J. R., Beecham, J. M., and Brand, L. (1983) *Chem. Phys. Lett. 102*, 501−507.
26. Marquardt, D. W. (1963) *J. Soc. Ind. Appl. Math. 11*, 431−441.
27. Bevington, P. R. (1969) *Data Reduction and Error Analysis for the Physical Sciences*, Vol. 1, McGraw-Hill, New York.
28. Bernasconi, C. F. (1976) *Relaxation Kinetics*, Academic Press, New York.
29. Schellman, J. A. (1987) *Annu. Rev. Biophys. Biophys. Chem. 16*, 115−37.
30. Brems, D. N. (1988) *Biochemistry 27*, 4541−6.
31. DeFelippis, M. R., Alter, L. A., Pekar, A. H., Havel, H. A., and Brems, D. N. (1993) *Biochemistry 32*, 1555−62.
32. Cleland, J. L., and Wang, D. I. (1990) *Biochemistry 29*, 11072−8.
33. Plomer, J. J., and Gafni, A. (1993) *Biochim. Biophys. Acta 1163*, 89−96.
34. Goldberg, M. E., Rudolph, R., and Jaenicke, R. (1991) *Biochemistry 30*, 2790−7.
35. Georgiou, G., Valax, P., Ostermeier, M., and Horowitz, P. M. (1994) *Protein Sci. 3*, 1953−60.
36. Fink, A. L. (1998) *Folding Des. 3*, 9−23.
37. Kelly, J. W. (1997) *Structure 5*, 595−600.
38. Zettlmeissl, G., Rudolph, R., and Jaenicke, R. (1979) *Biochemistry 18*, 5567−71.
39. Whytlaw-Gray, R., and Patterson, H. S. (1932) *Smoke: A Study of Aerial Disperse Systems*, E. Arnold, London.
40. Sosnick, T. R., Shtilerman, M. D., Mayne, L., and Englander, S. W. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 8545−50.
41. Herbst, R., Gast, K., and Seckler, R. (1998) *Biochemistry 37*, 6586−97.
42. Harper, J. D., and Lansbury, P. T., Jr. (1997) *Annu. Rev. Biochem. 66*, 385−407.
43. Zhang, S., Casey, N., and Lee, J. P. (1998) *Folding Des. 3*, 413−22.
44. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 51−5, 29−32.
45. Clore, G. M., Wingfield, P. T., and Gronenborn, A. M. (1991) *Biochemistry 30*, 2315−2323.
46. Kabsch, W., and Sander, C. (1983) *Biopolymers 22*, 2577−637.

BI991518M