Biophysical Chemistry 141 (2009) 214–221

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

Kinetic folding studies of the P22 tailspike beta-helix domain reveal multiple unfolded states

M.L. Spatara, C.J. Roberts, A.S. Robinson *

Department of Chemical Engineering, University of Delaware, Newark, Delaware 19716, United States

ARTICLE INFO

Article history:
Received 24 September 2008
Received in revised form 3 February 2009
Accepted 5 February 2009
Available online 12 February 2009

Keywords:
Beta-helix
Proline isomerization
Folding pathway
Viral adhesin
Aggregation

ABSTRACT

The beta-helix is an important protein fold in many pathogens, and is a challenging system for folding pathway prediction because it primarily is stabilized by non-local interactions along the primary sequence. A useful experimental model of this fold is a monomeric truncation of P22 tailspike protein, the beta-helix domain (bhx). This report describes a systematic *in vitro* study of the chemical denaturation and refolding of bhx. Results from equilibrium chemical denaturation experiments were consistent with a two-state folding mechanism, but showed only partial reversibility. Stopped-flow fluorescence studies showed a single unfolding step, but two refolding steps. The slow refolding step could be partly attributed to proline isomerization, based on an increased rate during refolding in the presence of PPIase and an increased relative amplitude of this step with increasing delay time in double-jump refolding experiments observed over delays of 5–100 s. However, double-jump refolding experiments with delay times longer than 100 s along with size exclusion chromatography and dynamic light scattering of refolding samples showed that the overall refolding yield decreased as bhx was unfolded for longer periods of time. Furthermore, the losses resulted from aggregate formation during refolding. This suggests that a change occurs over time in the unfolded or denatured state ensemble that increases the propensity for aggregation upon the shift to more native-favoring conditions. Alternatively aggregate nuclei may be able to form even under high denaturant conditions, and these subsequently grow and consume monomer when placed under native-favoring conditions.

1. Introduction

In protein folding, the formation of beta-sheet structure often involves interactions between residues that are far from each other in the primary sequence. For this reason, beta-sheet formation is often more difficult to predict than alpha-helix formation. However, the formation of beta-sheet structures during folding has gained recent attention, partially due to the prevalence of that structure in amyloid and other protein aggregates present in disease states [1]. One of the proposed model structures for amyloid is the parallel beta-helix [2,3], a coiled structure in which each rung of the helix contributes a beta-strand to each of three parallel beta-sheets that run along the helix surface. The beta-helix structure is being increasingly discovered as a native protein fold, particularly in proteins associated with infectious disease such as viral adhesin and tail proteins [4]. Based on sequence information, more than 97% of the passenger domains from autotransporters – secreted virulence proteins of gram-negative pathogenic bacteria – are predicted to adopt a beta-helix structure even though they show little sequence homology [5]. A detailed understanding of beta-helix formation is obviously important given this fold's utility as a model for protein aggregates in human disease, association with pathogenic systems, and ability to form robustly from diverse and non-repetitive protein sequences.

The current work focuses on the tailspike protein from the P22 bacteriophage. Fig. 1 shows the crystal structure of the tailspike trimer [6–8]. Tailspike has been used as a model system to study the folding of large, multimeric proteins, and its folding pathway has been extensively characterized both *in vitro* and *in vivo* [9–17]. Although tailspike is thermally stable and resistant to aggregation at temperatures up to 80 °C in its mature, trimeric form [18], it is highly susceptible to aggregation during folding, particularly *in vitro*, at elevated temperature, or in amino acid variants [12]. The aggregation is believed to occur through folding intermediates, which are destabilized under these conditions [12]. During refolding, tailspike adopts the majority of its secondary structure quickly, and in a monomeric state. It is believed that the successful formation of the beta-helix structure in the monomeric state is the critical step in determining the partitioning between a productive folding pathway or misfolding and aggregation [19]. Unfortunately, the complexity and irreversibility of the tailspike folding and assembly pathway, including competing aggregation events, make it difficult to study the beta-helix formation as an isolated event. For this reason, we are studying the *in vitro* folding of a truncated form of tailspike, comprising the beta-helix domain of the protein (bhx). The approximate location of the truncation points (residues 109 and 544) in the full-length structure

* Corresponding author. Tel.: +1 302 831 0557; fax: +1 301 831 6262.
E-mail address: asr@udel.edu (A.S. Robinson).

0301-4622/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2009.02.001

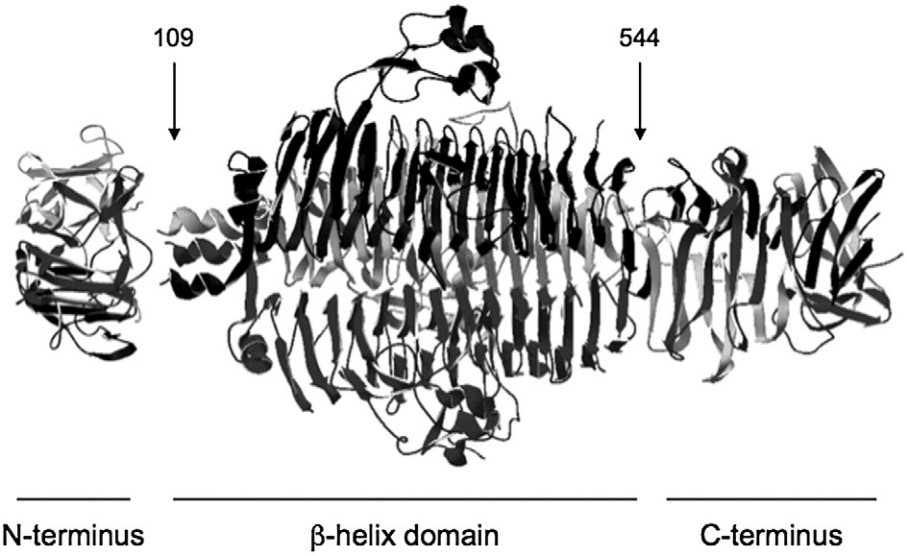


Fig. 1. P22 tailspike protein structure. Ribbon diagram of full-length, trimeric tailspike protein. Arrows indicate the approximate locations of residues 109 and 544, the truncation points for bhx.

is shown in Fig. 1. The bhx truncation has a molecular weight of 47.2 kDa, and exists primarily as a monomer at low concentrations [20]. Preliminary studies have demonstrated that bhx is spectroscopically similar to full length tailspike, retains partial enzymatic activity, has folding reversibility [20], and readily forms fibrous aggregates [21]. These observations indicate that the bhx truncation is structurally similar to the helix domain in the full-length protein and can act as a good model for a structured, monomeric folding intermediate.

2. Experimental

2.1. Materials

All chemicals used were obtained from major commercial suppliers. The cyclophilin enzyme was purchased from Sigma. Ultrapure urea was obtained from MP Biomedicals and used in all spectroscopic experiments.

2.2. Protein expression

Chemically competent *E. coli* BL21(DE3) cells (Novagen) were transformed with a pET11a plasmid containing the DNA encoding residues 109–544 of P22 tailspike (1tyr: pdb.org) under control of the T7 promoter [22]. Cells were selected on LB-Ampicillin (100 μg/mL) plates. For protein expression, an individual colony was selected and grown to an OD_{600} of approximately 0.6 at 20 °C before expression was induced by addition of 1 mM IPTG. Following induction for 18 h at 20 °C, cells were harvested by centrifugation at 4000 ×*g* for 20 min, resuspended in lysis buffer (50 mM Tris, pH 7.6, 25 mM NaCl, 2 mM EDTA, 20 mM $MgSO_4$, 20 μg/ml DNase, 100 μg/ml lysozyme, 0.1% Triton-X 100) and subjected to two freeze/thaw cycles (−80 °C/20 °C). Cellular debris was removed by centrifugation at 13,000 ×*g* for 30 min and the resulting supernatant was used for subsequent purification.

2.3. Protein purification

Purification of beta helix protein (bhx) was performed as previously described for the full length tailspike protein [23]. Purified bhx was applied to a Superdex 75 size exclusion column (GE Healthcare) and eluted isocratically in 100 mM phosphate buffer, pH 7, to remove any larger multimers formed during the purification prior to folding experiments. Protein concentration was determined by absorbance at 280 nm (1 OD_{280} = 1.33 mg/mL) [22]. Protein purity was >99% as determined by SDS–PAGE with silver staining.

2.4. Equilibrium folding/unfolding

Bhx was diluted to a concentration of 0.35 mg/mL in 100 mM phosphate buffer, pH 7, containing either 0 or 6 M urea and incubated in a 10 °C water bath for 1 h. The bhx samples were then diluted ten-fold into final urea concentrations between 0 and 6 M urea with a final protein concentration of 0.035 mg/mL and incubated for 1 h before measuring the fluorescence spectra. Samples incubated for longer times (up to 20 h) did not show a change in behavior compared to the 1 h samples. Spectra were measured at approximately 16 °C in a PC1 spectrofluorometer (ISS, Illinois) using excitation and emission slit widths of 0.5 and 1.0 mm, respectively, and an excitation wavelength of 280 nm. The emission was measured from 300 to 400 nm, and the spectra were quantified using the spectral center of mass (COM) [24], defined as

$$\mathrm{COM} = \left(\frac{\sum I_\nu \cdot \nu}{\sum I_\nu}\right)^{-1} \tag{1}$$

where ν is the wavenumber, and I_ν is the intensity at a given wavenumber, ν.

2.5. Stopped-flow folding/unfolding

Kinetic folding and unfolding experiments were performed in a model SX.18MV-R stopped-flow fluorescence instrument (Applied Photophysics, Leatherhead, U.K.). Unfolding and refolding were initiated by rapid dilution of native or denatured bhx into varying final urea concentrations. All experiments were performed at 15 °C in 100 mM phosphate buffer, pH 7. Intrinsic fluorescence was measured with an excitation wavelength of 280 nm, and the integrated emission signal was measured above a 305 nm cutoff filter. Kinetic traces were fit to the equation:

$$I(t) = I_\infty + \sum_i A_i \cdot e^{-k_i \cdot t} \tag{2}$$

where I is the fluorescence intensity, A_i is the amplitude of the ith phase, k_i is the rate constant of the ith phase, t is time in seconds, and I_∞ is the signal at infinite time. Quality of fit was evaluated by magnitude and randomness of residuals.

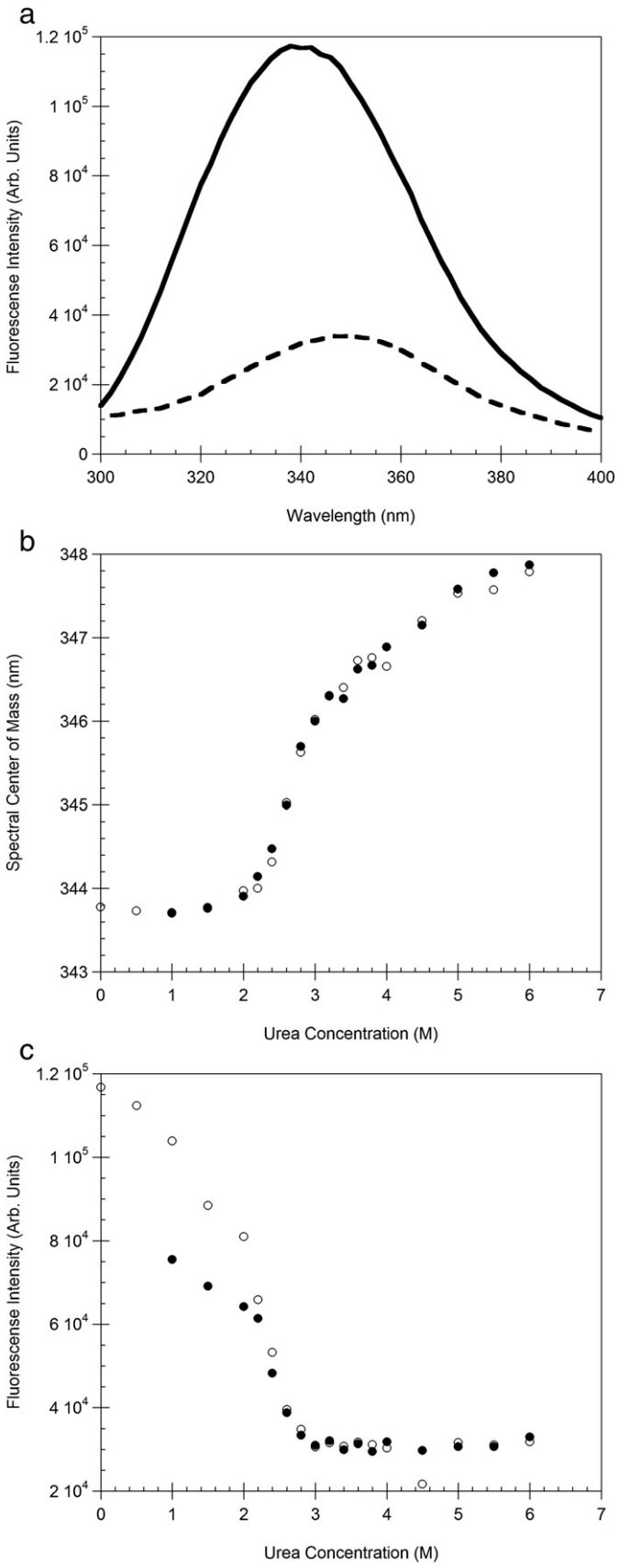


Fig. 2. Equilibrium folding of bhx. (a) Intrinsic fluorescence spectra of bhx in 0 M (_) and 6 M (…) urea. Denatured protein shows a decrease in intensity along with a small red shift. Spectral center of mass (b) and emission intensity at 340 nm (c) are shown for bhx unfolding (○) and refolding (●) as a function of final urea concentration.

2.6. Double-jump refolding assay

For double-jump (sequential mixing) refolding experiments, bhx samples were unfolded by dilution in 7 M urea, and were refolded by stopped-flow mixing after a measured delay time. For delay times shorter than 100 s, both mixing steps were performed using stopped-flow, and the final solution conditions were 0.05 mg/mL bhx in 1.17 M urea. For delay times longer than 100 s, the unfolding step was performed by manual dilution, and the final solution conditions were 0.05 mg/mL bhx in 0.64 M urea. Refolding data were fit to Eq. (2) as described for single-mixing stopped-flow experiments.

2.7. Size exclusion chromatography

Size exclusion chromatography was performed using a Superdex 200 column (GE Healthcare) in order to determine the oligomeric state of refolded samples. Samples were eluted isocratically with 50 mM Tris, pH 7.6, 250 mM NaCl, 2 mM EDTA. The high salt concentration in the mobile phase was used to minimize non-specific interactions between the resin and the sample. Peak identity in refolded samples was based on comparison with monomeric samples as well as the known void elution volume of the column.

2.8. Light scattering

Oligomeric states of refolding samples were also assayed using laser light scattering. Dynamic light scattering experiments were performed using a Brookhaven (Holtsville, NY) BI9000AT correlator and a BI200SM goniometer, and a Lexel (Fremont, CA) model 95 argon ion laser operating at 488 nm. The autocorrelation function was measured at 90°, and analyzed using a method of cumulants as described by Weiss et al. [25], to give values for the average hydrodynamic radius (R_h) and polydispersity index (p_2).

3. Results

3.1. Folding of bhx is not fully reversible

Equilibrium chemical denaturation experiments were used to assess the folding behavior of beta helix protein (bhx). The intrinsic fluorescence spectra of bhx in 0 and 6 M urea are shown in Fig. 2a. Upon unfolding, the spectrum shows a decrease in intensity as well as a red shift. Although the red shift upon denaturation is small, it is consistent with that observed for the fluorescence spectra of full length tailspike and results from high surface exposure of the tryptophans in the native structure. Equilibrium unfolding and refolding experiments were performed by dilution of bhx samples beginning from 0 and 6 M urea, respectively. All samples were incubated at the final urea concentration for at least 1 h prior to recording the spectra, and no change was observed upon further incubation.

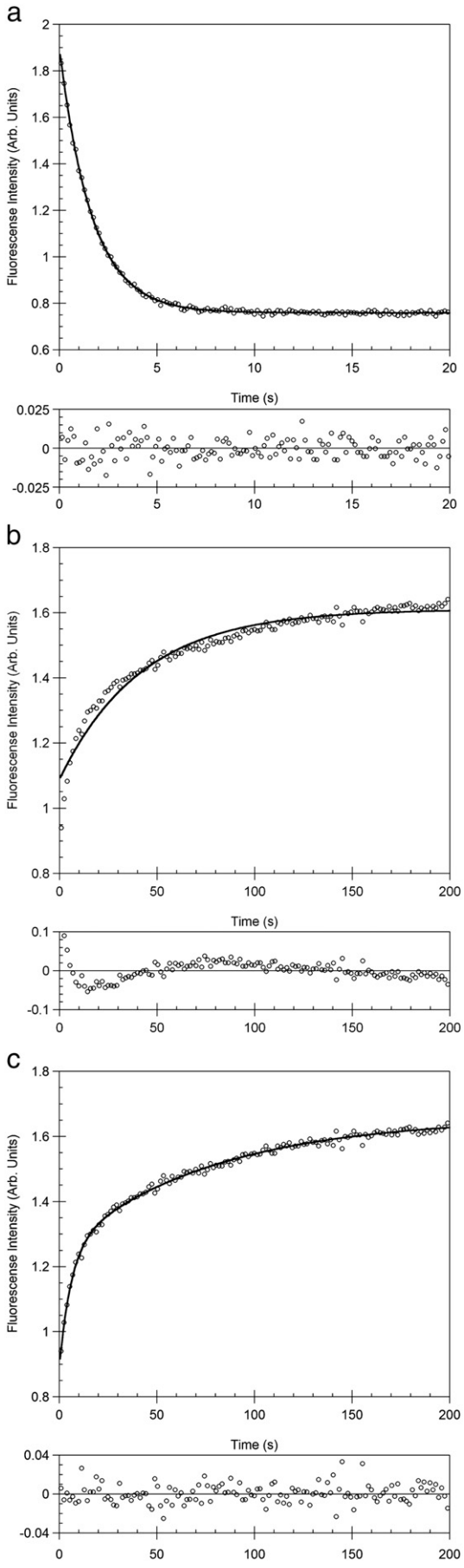
Table 1
Apparent thermodynamic[a] and kinetic[b] parameters for bhx unfolding.

Parameter	Calculated value[c]
ΔG^0_{app} (kcal mol^{-1})	7.4 ± 1.3
m_{eq} (kcal mol^{-1} M^{-1})	−2.9 ± 0.5
c_m (M urea)	2.6 ± 0.04
$k^0_u \times 10^6$ (s^{-1})	2.2 [0.2, 18]
m^1_u (M^{-1})	4.3 ± 0.99
m^2_u (M^{-1})	−0.35 ± 0.11

[a] Apparent free energy of unfolding in water(ΔG^0_{app}), denaturant dependence (m_{eq}), and midpoint unfolding concentration (c_m) calculated from regression of equilibrium COM data to a standard two-state model [46].

[b] Unfolding rate constant in water (k^0_u) and denaturant dependences (m^1_u, m^2_u) calculated from regression of kinetic unfolding data to Eq. (3).

[c] Uncertainties represent 95% confidence intervals from non-linear regression. For the unfolding rate constant, upper and lower bounds are given in brackets, as the confidence interval is not symmetric about the mean.



Denaturation curves are shown for both the spectral center of mass (Fig. 2b) and the fluorescence intensity at 340 nm (Fig. 2c). Although the curves show two-state behavior consistent with earlier reports [20], the fluorescence intensity was not fully recovered upon refolding, indicating partial misfolding or a loss in folding yield. Spectral center of mass data were fit to a two-state pseudo-equilibrium model, and the resulting parameters are listed in Table 1. It should be emphasized that these are only apparent equilibrium parameters, as the unfolding is clearly not fully reversible.

3.2. Kinetics of bhx folding and unfolding

The time-dependent folding behavior of bhx was further investigated through stopped-flow fluorescence experiments. For all conditions observed, unfolding data could be well fit to a single-exponential model (Fig. 3a), which suggests a single unfolding step. In contrast, refolding traces required two summed exponentials to accurately fit the data, as can be seen from the residuals of the one and two phase fits, shown in Fig. 3b and c, respectively. Inability to fit folding traces with a single exponential term indicates non-two-state behavior and could signify the presence of parallel folding paths or of folding intermediates along a sequential path. No improvement in quality of fit was observed by increasing the number of exponentials in the summation.

A chevron plot representation of the data from different (un) folding endpoint conditions is shown in Fig. 4. The unfolding rate constant shows a non-linear dependence on denaturant concentration. This type of chevron plot curvature has been observed for many protein systems including creatine kinase [26] and alpha-spectrin [27]. Suggested interpretations of chevron curvature include accumulation of kinetic intermediates and changes in the unfolding transition state as a function of denaturant concentration [28–30]. Unfolding data was fit with the following quadratic expression to extrapolate the value of k_u in the absence of denaturant:

$$\ln(k_u) = \ln\left(k_u^0\right) + m_u^1[c] + m_u^2[c]^2 \quad (3)$$

where k_u is the unfolding rate constant, k_u^0 is the unfolding rate constant without urea, and c is the concentration of urea (Table 1).

Perhaps more significantly, the rate constants of both folding phases show little dependence on denaturant concentration. The two phases have relaxation times of approximately 20 s and 100 s. As in the equilibrium experiments, the fluorescence intensity was not fully recovered upon refolding.

3.3. Refolding in presence of PPIase

Slow steps in protein folding can often be attributed to *cis–trans* isomerization around prolyl-peptide bonds [31]. In unordered polypeptides, the trans conformation is favored, and the ratio of trans:cis residues is approximately 70:30. In contrast, proline residues in folded protein structures typically adopt a specific isomeric state, based on steric constraints. In the bhx domain of tailspike, there are 16 proline residues, all of which adopt the *trans* conformation in the native state, according to the crystal structure [6–8].

To test for the significance of proline isomerization during the *in vitro* folding of bhx, time-dependent folding experiments were performed in the presence of the human peptidyl-prolyl isomerase, cyclophilin (PPIase). Representative refolding traces of bhx with and without the presence of 0.5 μM PPIase in the refolding buffer are

Fig. 3. Kinetics of (un)folding. Representative kinetic traces for unfolding (a) and refolding (b), (c) of bhx. Experimental data (○) and fits to a summed exponential expression (_) with one (a), (b) or two (c) terms are shown in the main plot. Residuals for the shown fit are plotted below. For clarity, plots show every third experimental point.

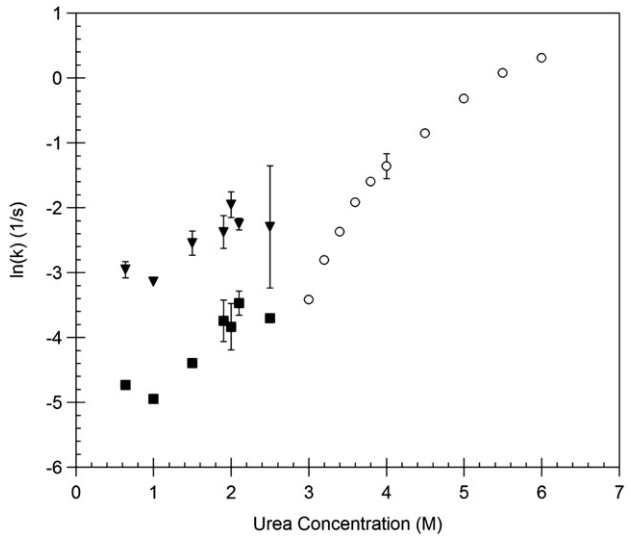


Fig. 4. Chevron plot representation of kinetic folding data. The urea dependence of the rate constants for unfolding (○) as well as the slow (■) and fast (▼) folding phases of bhx. Error bars represent the standard deviation from at least three experiments, and are smaller than the size of the symbol where not shown.

shown in Fig. 5. A fit of the data to Eq. (2) showed a 30% increase in the rate of the slower folding phase, consistent with catalysis of an isomerization step by the PPIase. No statistically significant change was observed in the rate constant of the fast folding phase.

3.4. Double-jump refolding

Double-jump refolding experiments were used to test for changes in the unfolded state over time. Because proline isomerization rates are slow relative to protein unfolding rates, the proline residues in a newly unfolded protein will initially be in the same conformation they had in the native state. Over time, an equilibrium will be established between *cis* and *trans* isomers in the unfolded protein. If proline isomerization is responsible for a slow folding step, this step should become more pronounced as a larger fraction of the unfolded protein population adopts the non-native isomer.

Double-jump refolding experiments were performed for both short delay times (5–100 s) and long delay times (>100 s). In both cases, the rate constants for the two folding phases did not change significantly over time (Fig. 6a and c) indicating that the mechanism is independent of delay time. With increasing delay times up to 100 s, the amplitude of the slow phase increased as that of the fast phase decreased (Fig. 6b). This is consistent with a slow phase linked to proline isomerization, where the population of the slow-folding isomer increases during the time the protein is unfolded. Interestingly, at delay times longer than 100 s the amplitudes of both phases decreased with increasing delay time (Fig. 6d) while the ratio of the amplitudes from the two phases remained constant. Thus the total amount of fluorescence recovery, which correlates with protein refolding yield, decreased over time. Additionally, the partitioning of unfolded protein into fast and slow folding populations was established by 100 s, and remained constant over longer times.

3.5. Irreversible loss due to aggregation

One key question is whether the loss of recovery of fluorescence intensity is due to misfolding of bhx to a state with lower fluorescence, or to a loss in overall refolding yield. To address this question, refolding samples were analyzed by size exclusion chromatography (Fig. 7). The chromatogram shows that the refolded sample consists of a mixture of monomer and a second species, which elutes at an earlier volume and is likely a small, soluble oligomer. The total peak area across both species remained constant, indicating that no protein is lost (i.e. becomes insoluble). The oligomer formed relatively quickly upon refolding, as it was observed in samples loaded onto the SEC column immediately following dilution into native-favoring conditions. In double-jump experiments analyzed by SEC, the monomer peak area decreased as a function of delay time, with a corresponding increase in the oligomer peak. If refolding was initiated following unfolding overnight, the resulting sample consisted entirely of oligomer, with no monomer observed in the chromatogram.

Dynamic light scattering (DLS) was performed to determine the approximate size and heterogeneity of the oligomeric species. The light scattering measurement was performed on a sample of native monomer (never unfolded) and a refolded sample consisting of only oligomer as confirmed by SEC. Analysis using the method of cumulants [25] gave a calculated hydrodynamic radius of 4.04 ± 0.07 nm for the monomer and 9.97 ± 0.15 nm for the oligomer. The polydispersity indexes for the monomer and oligomer were 0.15 ± 0.04 and 0.21 ± 0.03, respectively, indicating that both samples have relatively low polydispersity. From the combined SEC and DLS data, it is clear that the loss in fluorescence intensity upon refolding results from a loss of refolding yield due to formation of a small, misfolded oligomer.

4. Discussion

In this study, we examined the folding behavior of a truncation of P22 tailspike protein comprising the beta-helix domain as a model for beta-helix formation in the full-length protein. We found that although equilibrium unfolding occurred via a two-state pathway, refolding was not a simple two-state process. A proposed folding scheme for bhx that is consistent with all the data is shown in Fig. 8. The parentheses indicate steps that occur primarily in the unfolded state. The proposed mechanism shows an equilibrium between two unfolded species, U_{trans} and U_{cis}, which refold through parallel paths. U_{trans} refolds to the native species, N, in one step with rate constant k_1, while U_{cis} refolds in two steps through an intermediate, I. The first step, with rate constant k_3, is the major folding step, while the second, slower step, with rate constant k_2, represents an isomerization event. The scheme also shows an aggregation-prone unfolded state, U*, which forms during time in denatured-favoring conditions with rate constant, k_4. Aggregates or oligomers which are formed from U* are labeled A.

The ability of PPIase to catalyze the slow folding step of bhx is a strong indication that this step involves proline isomerization. Although the 30% increase in refolding rate is not dramatic, the activity of PPIase is often

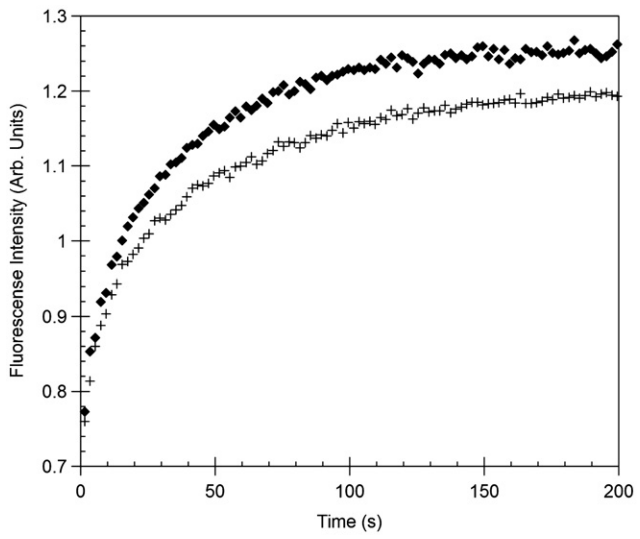


Fig. 5. Effect of PPIase. Refolding kinetic traces for bhx with (◆) and without (+) PPIase.

low, particularly in non-optimal solution conditions or in cases where the prolines are not easily accessible to the enzyme [31,32].

Changes in proline isomerization that occur in the unfolded state are generally not accompanied by a change in tryptophan fluorescence, and are therefore not directly observable using intrinsic fluorescence spectroscopy. Sequential mixing, or double jump experiments provide an indirect method to observe changes in the unfolded population based on observed differences in behavior when refolding is induced. Our double jump experiments with delay times up to 100 s showed the characteristic proline isomerization behavior, in which the amplitude of the slow folding phase increases relative to the fast phase with increasing delay times. The time scale at which proline isomerization effects were observed in double jump experiments varies. For some systems, changes in relative amplitude are observed at delay times of several minutes or longer [28], while for other systems the effect is primarily seen at shorter delay times [33]. For bhx, an equilibrium between fast and slow folding isomers was established within 100 s and no further change in the relative amplitudes of the two folding phases was observed after that time.

Experimentally, only two distinct kinetic folding phases are observed, although there are three total folding steps in the proposed mechanism. There are two mathematically possible explanations for this discrepancy: k_2 is very fast compared to k_3 such that the conversion of U_{cis} to I is rate limiting, or k_1 and k_3 have very similar values such that they are observed as one combined phase in the kinetic experiments. The first possibility is not consistent with our knowledge of the physical system, since k_2, the rate constant of the isomerization step, should be slow. In contrast, the latter possibility is easily explained, since k_1 and k_3 are both rate constants of major structural rearrangement steps in folding, and it is therefore reasonable for them to have similar values.

Proline isomerization has been identified as the slow folding step in another beta-helical protein, pectate lyase C (pelC) from *Erwinia chrysanthemi* [28,34]. There are 12 prolines in pelC; one of them has the *cis* confirmation in the native structure and is responsible for the slowest folding phase present. A second slow phase was also attributed to proline isomerization, although a comprehensive mutagenesis study was unable to identify the specific residue responsible for this phase [34]. The authors

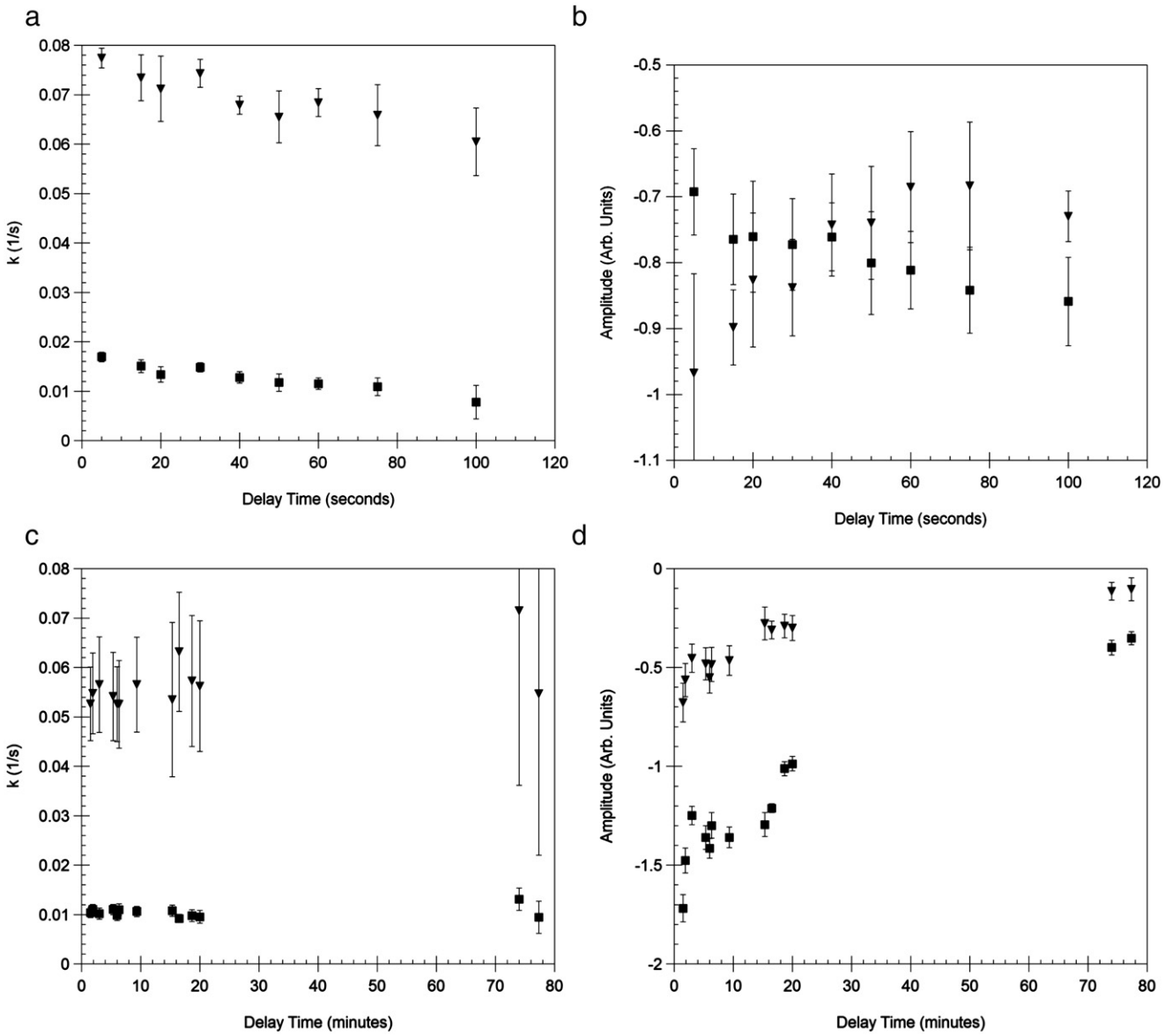


Fig. 6. Sequential-mixing refolding. Samples of bhx were allowed to unfold in high concentration urea solutions for varying delay times, before being refolded by dilution into low urea conditions. The rate constants (a,c) and relative amplitudes (b,d) of the fast (▼) and slow (■) refolding phases of bhx are plotted as a function of the refolding delay time. Error bars for data from long delay times (c,d) represent the 95% confidence intervals from the non-linear regression. For data with short delay times, where variability between runs was significant, error bars represent the standard deviation in the parameters from fits of at least 5 traces for each time point.

concluded that multiple proline residues contribute to this phase, and bhx folding behavior can likely be similarly explained.

The loss of overall folding yield with increased residence time in the denatured state is an unusual behavior in protein folding. The oligomer is seen in SEC samples injected immediately after dilution into native favoring conditions; therefore it forms relatively quickly. The development of aggregation propensity in the unfolded ensemble, however, is significantly slower than other folding steps, with a time constant of approximately 50 min. Attempts to refold bhx after long denaturation periods (overnight) resulted in formation of only the aggregated species with no detectable monomer, indicating that the formation of the aggregation prone species is effectively irreversible.

The formation of an aggregation prone unfolded state of bhx is likely to have implications for the folding of full-length tailspike. It has been shown that the early *in vitro* folding states of tailspike are significantly different from those observed *in vivo*, where the beta-helix structure is formed while the nascent polypeptide is still attached to the ribosome [35]. Decreased efficiency in formation of the beta-helix structure from polypeptides that have been unfolded in denaturant for significant periods of time may explain the significant loss to aggregation observed during *in vitro* refolding of tailspike. The current study conclusively demonstrates that folding of the monomeric beta-helix structure from tailspike, a critical first step in folding and assembly, is not a simple two-state process and may be responsible for the differences in folding efficiency *in vitro* and *in vivo*.

One possible physical explanation for the aggregation-prone unfolded state of bhx is isomerization of multiple proline residues. Perhaps a single proline in the *cis* conformation merely creates a slow folding form of bhx requiring an additional isomerization step to reach the native form, but multiple prolines in the *cis* conformation make it nearly impossible for the helix to form at all. It is also possible that some proline residues are critical for folding, while others are less important. Given the large number of proline residues in the bhx sequence, it is likely that both the number and position of *cis* prolines in the unfolded polypeptide contribute to its ability to productively refold. The resulting partially folded or misfolded intermediate species would be vulnerable to aggregation due to exposure of hydrophobic residues that are normally buried in the native structure. Refolding mechanisms with kinetic folding traps involving proline isomerization have also been reported for single chain antibodies [36,37].

Aggregation due to kinetic folding traps is consistent with the common view that protein aggregation occurs through partially unfolded,

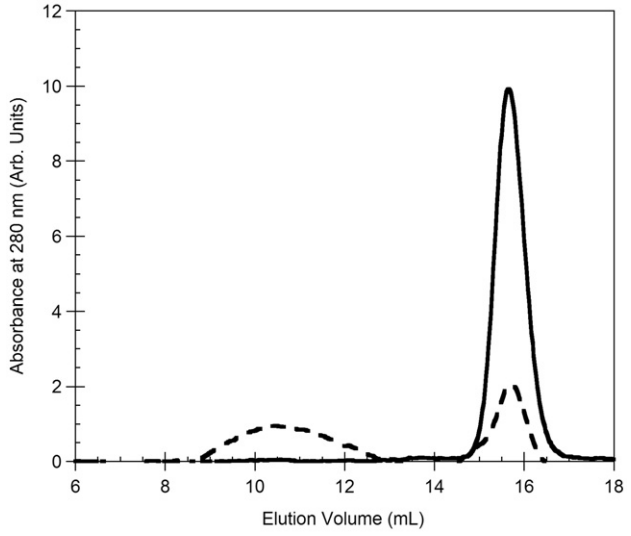


Fig. 7. Size-exclusion chromatogram of native bhx (_) and a refolded bhx sample (…). The chromatogram of bhx refolded after denaturation for approximately 2 h shows an additional peak with an earlier elution time, indicating formation of an oligomer.

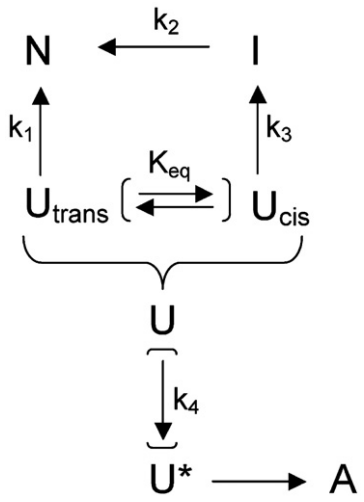


Fig. 8. Proposed folding mechanism for bhx showing a parallel refolding pathway which results from proline isomerization, as well as formation of an off-pathway oligomer. Parentheses indicate folding steps which occur primarily in the unfolded state.

molten-globule monomers, where early association steps result from relatively non-specific interactions [38,39]. In some cases, however, aggregation is driven by specific interactions. The presence of a specific conformation of protein, even in a small fraction of the population, can act as a seed for aggregation. These aggregate seeds or nuclei, which may consist of monomers or oligomers, are able to recruit additional monomer leading to formation of aggregates. An important example of this behavior is observed in prion diseases, where the aggregation-prone species is a conformationally altered monomeric form of prion protein [40,41]. Perhaps bhx adopts a similar, specific conformation under denaturing conditions that leads to the rapid formation of aggregate upon dilution into native favoring conditions.

The denatured state ensemble of proteins, once viewed as a random coil, is now recognized to contain significant structure and is being characterized with a variety of techniques [42–44]. Here, low levels of dimer (~3–10%) were observed in samples of unfolded bhx that had been chemically crosslinked using bis(maleimido)hexane prior to separation by SDS–PAGE (data not shown). It is possible that either these dimers or a specific conformation of monomers act as nuclei or seeds for further aggregation upon dilution into native favoring conditions. The low polydispersity of the aggregates formed during refolding of bhx is consistent with a mechanism in which fast growth occurs from preformed nuclei [45]. Further study of the denatured state of bhx could provide valuable insight into the nature of the aggregate nuclei.

Acknowledgements

The authors thank William F. Weiss IV for assistance with performing and analyzing data from DLS experiments, Dr. Eric Kaler for use of the light scattering equipment, and Dr. Carolyn Teschke at the University of Connecticut for assistance with sequential mixing stopped-flow experiments.

This research was supported by National Science Foundation IGERT 0221651 (MLS), National Science Foundation BES 99-84312 (ASR), and a Merck Faculty Fellowship (CJR).

References

[1] C.M. Dobson, Protein misfolding, evolution and disease, Trends Biochem. Sci. 24 (1999) 329–332.
[2] R. Wetzel, Ideas of order for amyloid fibril structure, Structure 10 (2002) 1031–1036.
[3] C. Govaerts, H. Wille, S.B. Prusiner, F.E. Cohen, Evidence for assembly of prions with left-handed {beta}-helices into trimers, Proc. Nat. Acad. Sci. U. S. A. 101 (2004) 8342–8347.

[4] P.R. Weigele, E. Scanlon, J. King, Homotrimeric, beta-stranded viral adhesins and tail proteins, J. Bacteriol. 185 (2003) 4022–4030.
[5] M. Junker, C.C. Schuster, A.V. McDonnell, K.A. Sorg, M.C. Finn, B. Berger, P.L. Clark, Pertactin beta-helix folding mechanism suggests common themes for the secretion and folding of autotransporter proteins, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 4918–4923.
[6] S. Steinbacher, U. Baxa, S. Miller, A. Weintraub, R. Seckler, R. Huber, Crystal structure of phage P22 tailspike protein complexed with *Salmonella* sp. O-antigen receptors, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 10584–10588.
[7] S. Steinbacher, S. Miller, U. Baxa, N. Budisa, A. Weintraub, R. Seckler, R. Huber, Phage P22 tailspike protein: crystal structure of the head-binding domain at 2.3 A, fully refined structure of the endorhamnosidase at 1.56 A resolution, and the molecular basis of O-antigen recognition and cleavage, J. Mol. Biol. 267 (1997) 865–880.
[8] S. Steinbacher, R. Seckler, S. Miller, B. Steipe, R. Huber, P. Reinemer, Crystal structure of P22 tailspike protein: interdigitated subunits in a thermostable trimer, Science 265 (1994) 383–386.
[9] D. Goldenberg, J. King, Trimeric intermediate in the *in vivo* folding and subunit assembly of the tailspike endorhamnosidase of bacteriophage P22, Proc. Nat. Acad. Sci. U. S. A. 79 (1982) 3403–3407.
[10] D.P. Goldenberg, D.H. Smith, J. King, Genetic analysis of the folding pathway for the tail spike protein of phage P22, Proc. Nat. Acad. Sci. U. S. A. 80 (1983) 7060–7064.
[11] J. King, M.H. Yu, Mutational analysis of protein folding pathways: the P22 tailspike endorhamnosidase, Methods Enzymol. 131 (1986) 250–266.
[12] C.A. Haase-Pettingell, J. King, Formation of aggregates from a thermolabile *in vivo* folding intermediate in P22 tailspike maturation. A model for inclusion body formation, J. Biol. Chem. 263 (1988) 4977–4983.
[13] R. Seckler, A. Fuchs, J. King, R. Jaenicke, Reconstitution of the thermostable trimeric phage P22 tailspike protein from denatured chains *in vitro*, J. Biol. Chem. 264 (1989) 11750–11753.
[14] A.S. Robinson, J. King, Disulphide-bonded intermediate on the folding and assembly pathway of a non-disulphide bonded protein, Nat. Struct. Biol. 4 (1997) 450–455.
[15] M.A. Speed, T. Morshead, D.I. Wang, J. King, Conformation of P22 tailspike folding and aggregation intermediates probed by monoclonal antibodies, Protein Sci. 6 (1997) 99–108.
[16] M. Beißinger, S.C. Lee, S. Steinbacher, P. Reinemer, R. Huber, M.-H. Yu, R. Seckler, Mutations that stabilize folding intermediates of phage P22 tailspike protein folding *in vivo* and *in vitro*, stability, and structural context, J. Mol. Biol. 249 (1995) 185–194.
[17] A. Fuchs, C. Seiderer, R. Seckler, *In vitro* folding pathway of phage P22 tailspike protein, Biochemistry 30 (1991) 6598–6604.
[18] D. Sargent, J.M. Benevides, M.-H. Yu, J. King, G.J. Thomas, Secondary structure and thermostability of the phage P22 tailspike: XX. Analysis by Raman spectroscopy of the wild-type protein and a temperature-sensitive folding mutant, J. Mol. Biol. 199 (1988) 491–502.
[19] R. Simkovsky, J. King, An elongated spine of buried core residues necessary for *in vivo* folding of the parallel beta-helix of P22 tailspike adhesin, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 3575–3580.
[20] S. Miller, B. Schuler, R. Seckler, A reversibly unfolding fragment of P22 tailspike protein with native structure: the isolated beta-helix domain, Biochemistry 37 (1998) 9160–9168.
[21] B. Schuler, R. Rachel, R. Seckler, Formation of fibrous aggregates from a non-native intermediate: the isolated P22 tailspike beta-helix domain, J. Biol. Chem. 274 (1999) 18589–18596.
[22] M.J. Gage, A.S. Robinson, C-terminal hydrophobic interactions play a critical role in oligomeric assembly of the P22 tailspike trimer, Protein Sci. 12 (2003) 2732–2747.
[23] C. Haase-Pettingell, S. Betts, S.W. Raso, L. Stuart, A. Robinson, J. King, Role for cysteine residues in the *in vivo* folding and assembly of the phage P22 tailspike, Protein Sci. 10 (2001) 397–410.
[24] J.L. Silva, E.W. Miles, G. Weber, Pressure dissociation and conformational drift of the beta dimer of tryptophan synthase, Biochemistry 25 (1986) 5780–5786.
[25] W.F. Weiss IV, T.K. Hodgdon, E.W. Kaler, A.M. Lenhoff, C.J. Roberts, Nonnative protein polymers: structure, morphology, and relation to nucleation and growth, Biophys. J. 93 (2007) 4392–4403.
[26] Y.X. Fan, J.M. Zhou, H. Kihara, C.L. Tsou, Unfolding and refolding of dimeric creatine kinase equilibrium and kinetic studies, Protein Sci. 7 (1998) 2631–2641.
[27] S. Batey, K.A. Scott, J. Clarke, Complex folding kinetics of a multidomain protein, Biophys. J. 90 (2006) 2120–2130.
[28] D.E. Kamen, R.W. Woody, Folding kinetics of the protein pectate lyase C reveal fast-forming intermediates and slow proline isomerization, Biochemistry 41 (2002) 4713–4723.
[29] D.E. Otzen, O. Kristensen, M. Proctor, M. Oliveberg, Structural changes in the transition state of protein folding: alternative interpretations of curved chevron plots, Biochemistry 38 (1999) 6499–6511.
[30] R. Kumar, A.K. Bhuyan, Two-state folding of horse ferrocytochrome c: analyses of linear free energy relationship, chevron curvature, and stopped-flow burst relaxation kinetics, Biochemistry 44 (2005) 3024–3033.
[31] J.F. Brandts, H.R. Halvorson, M. Brennan, Consideration of the possibility that the slow step in protein denaturation reactions is due to cis–trans isomerism of proline residues, Biochemistry 14 (1975) 4953–4963.
[32] K. Lang, F.X. Schmid, G. Fischer, Catalysis of protein folding by prolyl isomerase, Nature 329 (1987) 268–270.
[33] S.M. Doyle, O. Bilsel, C.M. Teschke, SecA folding kinetics: a large dimeric protein rapidly forms multiple native states, J. Mol. Biol. 341 (2004) 199–214.
[34] D.E. Kamen, R.W. Woody, Identification of proline residues responsible for the slow folding kinetics in pectate lyase C by mutagenesis, Biochemistry 41 (2002) 4724–4732.
[35] P.L. Clark, J. King, A newly synthesized, ribosome-bound polypeptide chain adopts conformations dissimilar from early *in vitro* refolding intermediates, J. Biol. Chem. 276 (2001) 25411–25420.
[36] H. Lilie, R. Rudolph, J. Buchner, Association of antibody chains at different stages of folding: prolyl isomerization occurs after formation of quaternary structure, J. Mol. Biol. 248 (1995) 190–201.
[37] M. Jäger, A. Plückthun, The rate-limiting steps for the folding of an antibody scFv fragment, FEBS Lett. 418 (1997) 106–110.
[38] R.M. Murphy, Peptide aggregation in neurodegenerative disease, Annu. Rev. Biomed. Eng. 4 (2002) 155.
[39] C.M. Dobson, Principles of protein folding, misfolding and aggregation, Semin. Cell Dev. Biol. 15 (2004) 3–16.
[40] C. Soto, J. Castilla, The controversial protein-only hypothesis of prion propagation, Nat. Med. 10 (2004) S63–67 Suppl.
[41] C. Soto, L. Estrada, J. Castilla, Amyloids, prions and the inherent infectious nature of misfolded protein aggregates, Trends Biochem. Sci. 31 (2006) 150–155.
[42] R.W. Alston, M. Lasagna, G.R. Grimsley, J.M. Scholtz, G.D. Reinhart, C.N. Pace, Tryptophan fluorescence reveals the presence of long-range interactions in the denatured state of ribonuclease Sa, Biophys. J. 94 (2008) 2288–2296.
[43] E. Sherman, A. Itkin, Y.Y. Kuttner, E. Rhoades, D. Amir, E. Haas, G. Haran, Using fluorescence correlation spectroscopy to study conformational changes in denatured proteins, Biophys. J. 94 (2008) 4819–4827.
[44] J.H. Cho, S. Sato, J.C. Horng, B. Anil, D.P. Raleigh, Electrostatic interactions in the denatured state ensemble: their effect upon protein folding and protein stability, Arch. Biochem. Biophys. 469 (2008) 20–28.
[45] J.M. Andrews, C.J. Roberts, A Lumry–Eyring nucleated polymerization model of protein aggregation kinetics: 1. Aggregation with pre-equilibrated unfolding, J. Phys. Chem. B 111 (2007) 7897–7913.
[46] C.N. Pace, Determination and analysis of urea and guanidine hydrochloride denaturation curves, Methods Enzymol. 131 (1986) 266–280.