Protein Science (1995), *4*:900-908. Cambridge University Press. Printed in the USA.

Multimeric intermediates in the pathway to the aggregated inclusion body state for P22 tailspike polypeptide chains

MARGARET A. SPEED, DANIEL I.C. WANG, AND JONATHAN KING

Biotechnology Process Engineering Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(RECEIVED December 6, 1994; ACCEPTED February 17, 1995)

Abstract

The failure of newly synthesized polypeptide chains to reach the native conformation due to their accumulation as inclusion bodies is a serious problem in biotechnology. The critical intermediate at the junction between the productive folding and the inclusion body pathway has been previously identified for the P22 tailspike endorhamnosidase. We have been able to trap subsequent intermediates in the in vitro pathway to the aggregated inclusion body state. Nondenaturing gel electrophoresis identified a sequential series of multimeric intermediates in the aggregation pathway. These represent discrete species formed from noncovalent association of partially folded intermediates rather than aggregation of native-like trimeric species. Monomer, dimer, trimer, tetramer, pentamer, and hexamer states of the partially folded species were populated in the initial stages of the aggregation reaction. This methodology of isolating early multimers along the aggregation pathway was applicable to other proteins, such as the P22 coat protein and carbonic anhydrase II.

Keywords: aggregation; bacteriophage P22; folding intermediates; in vitro refolding; inclusion body; polymerization; protein folding; tailspike protein

A major problem in biotechnology is the incorrect folding of newly synthesized polypeptide chains and formation of insoluble aggregates that are biologically inactive (Marston, 1986; Mitraki & King, 1989; DeBernardez-Clark & Georgiou, 1991; Wetzel, 1994). In vivo, this folding problem frequently arises with heterologous proteins overexpressed in *Escherichia coli*, which form inclusion bodies or amorphous aggregates within the cell. The analogous aggregation problem occurs in vitro via a similar association mechanism (Zettlmeissl et al., 1979; Colon & Kelly, 1992; Mitraki et al., 1993). Non-native multimerization subsequently leads irreversibly to the formation of large aggregates or inclusion bodies. A frequently employed strategy for minimizing in vitro aggregation is to refold protein under dilute conditions. The drawback of this practice is that it dilutes the final product, and the strategy does not work for oligomeric proteins that cannot assemble efficiently under dilute conditions (Jaenicke & Rudolph, 1986; Teschke & King, 1993). Therefore, further study of the aggregation reaction is essential in designing effective refolding strategies.

Analysis of the aggregation phenomenon has been limited by the available methodologies to study protein folding intermediates. Most analytical bioseparation techniques were designed to isolate stable native protein rather than metastable, hydrophobic species. Early aggregation studies used centrifugation to separate the aggregates in the pellet from the soluble protein. Due to the limitations in the analytical techniques, all multimers on the aggregation pathway have been grouped into one category, namely "aggregates." Understanding the mechanisms underlying the off-pathway aggregation reaction requires a description of the chain association pathway and multimeric intermediates.

In the cases where the competition between the refolding and aggregation reactions has been directly studied, the precursor to aggregation is a defined intermediate in the refolding pathway (Zettlmeissl et al., 1979; Brems, 1988; Hendrick & Hartl, 1993). Brems used a two-step dilution method to investigate the folding intermediates and aggregation process during refolding of bovine growth hormone (bGH). Turbidity data (450 nm) indicated that during refolding, the initial unfolding conditions affected the extent of aggregation and that only partially unfolded bGH led to aggregation. For numerous proteins, this class of folding intermediates is the species that is recognized by the GroE chaperonin, which assists in protein folding by binding to folding intermediates and preventing improper association of the polypeptide chains (Goloubinoff et al., 1989). Cleland and Wang (1991) have used quasi-elastic light scattering to determine the aggregation mechanism of bovine carbonic

Reprint requests to: Jonathan King, Massachusetts Institute of Technology, Department of Biology, Room 68-330, Cambridge, Massachusetts 02139; e-mail: jaking@mit.edu.

anhydrase II (CAB), a monomeric protein. In vitro refolding studies showed that a monomeric folding intermediate (I_1) is in equilibrium with a dimer with a K_{eq} of 1.3 mM^{-1} and association rate of 5.16×10^{-3} min^{-1}. After dimerization, trimerization proceeds rapidly (K_{eq} of 0.42 M^{-1} and association rate of 0.133 min^{-1}), which directly leads to irreversible aggregation (Cleland & Wang, 1991).

The in vivo intermediates in inclusion body formation have been characterized for the P22 tailspike endorhamnosidase (Haase-Pettingell & King, 1988). The tailspike is an oligomeric protein that binds phage P22 to the *Salmonella* cell by hydrolyzing the rhamnosylgalactose linkage in the *Salmonella* O-antigen. Analysis of the folding and aggregation mechanism of the tailspike protein have shown that aggregation occurs by the association of a partially folded monomeric intermediate rather than the native species. This thermolabile intermediate preferentially partitions onto the aggregation pathway as the temperature increases (Haase-Pettingell & King, 1988).

The in vivo folding and aggregation pathway of the P22 tailspike protein involves an unfolded nascent polypeptide chain coming off the ribosome and forming a partially folded species with significant secondary structure (Fig. 1A). This single-chain intermediate (I) can associate to form a protrimer (pt) of partially folded chains (Goldenberg et al., 1982; Haase-Pettingell & King, 1988). Although the protrimer is a transient species, it is metastable enough to be trapped on an acrylamide gel at low temperatures. The final folding step is the conversion of the protrimer to the stable native tailspike by a first order reaction. Because the single-chain folding intermediates are only partially folded species, there is no native monomer species.

Temperature-sensitive folding mutations further destabilize the early folding intermediate to shift from the productive folding pathway onto the aggregation pathway at restrictive temperatures of 38–42 °C (Goldenberg & King, 1981). Mutations at two sites suppress the aggregation reaction of the tailspike intermediates both in vivo and in vitro (Mitraki et al., 1991, 1993). The existence of single amino acid substitutions that suppress inclusion body formation suggested that the reaction was specific.

Recently, the structure of the tailspike has been solved by X-ray crystallography (Steinbacher et al., 1994). Figure 1 is a ribbon diagram depicting the structure of the tailspike monomer (Fig. 1B) and native trimer (Fig. 1C). The central region of each 72-kDa subunit of the native trimer has a β-spiral conformation similar to pectate lyase C (Yoder et al., 1994). The C-terminal regions interdigitate and form further β-sheet structures. The conformation of the partially folded intermediates is not known, but is likely to have substantial β-sheet content.

Tailspike refolding studies have directly compared in vitro folding and the aggregation reaction to the in vivo mechanism. The in vitro folding pathway of the tailspike protein corresponds to the in vivo folding reaction, having similar folding intermediates and folding kinetics (Fuchs et al., 1991; Danner & Seckler, 1993).

For most small monomeric proteins, hydrophobic collapse occurs on a timescale less than 1 ms, and the transition from a molten globule folding intermediate to the native state requires times on the order of seconds (Kim & Baldwin, 1990). For the tailspike protein, the transition of protrimer to native trimer is the rate-limiting step in the folding reaction. The half-times of tailspike refolding for both the in vivo and in vitro reactions are on the order of minutes rather than seconds (Danner & Seckler, 1993),

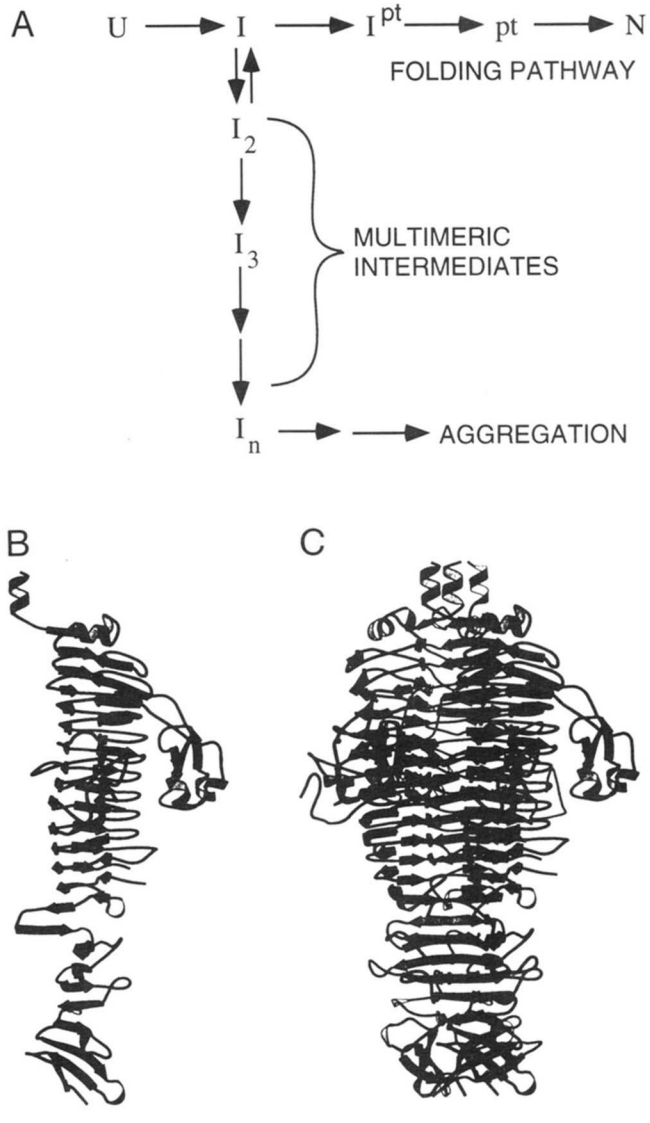

Fig. 1. A: Folding and aggregation pathway of the P22 tailspike protein. U, unfolded species; I, folding intermediate; pt, protrimer; N, native trimer; I_n, aggregation intermediates. Ribbon diagram of P22 tailspike monomer (**B**) and native trimer (**C**). Structure drawn by Per Kraulis with the MOLSCRIPT program using coordinates of Steinbacher et al. (1994) and reprinted with permission of Goldenberg and Creighton (1994).

and metastable intermediates can be trapped in the cold (Goldenberg & King, 1982). Half-times for the off-pathway aggregation reactions are similar to those for the productive pathway. This suggested that it might be possible to directly identify intermediates along the aggregation pathway.

Results

The in vitro partitioning of refolding intermediates between the productive and aggregation pathways had been established initially for denatured polypeptide chains diluted from concentrated solutions of guanidinium hydrochloride (Mitraki et al., 1993). To avoid problems associated with the presence of a charged denaturant, acid-urea (5 M urea, pH 3) was used as the

initial denaturant as described by Seckler et al. (1989). Denatured protein was diluted into buffer at pH 7.6 to a final concentration of 100 μg/mL. Under these conditions, approximately 10% of the chains refolded to the native trimer, which is SDS resistant, and the remainder of the chains associated to form small aggregates as detected by quasi-elastic light scattering. At the early stages of the multimerization reaction, the extent of aggregation under these conditions was too low to measure accurately by classical light scattering at 500 nm.

In the original studies of the in vivo folding pathway, nondenaturing polyacrylamide gel electrophoresis (PAGE) in the cold was used to identify the metastable protrimer intermediate (Goldenberg & King, 1982). It seemed reasonable that nondenaturing gel electrophoresis in the cold might also be useful in separating putative multimeric intermediates along the aggregation pathway.

For in vitro refolding studies, native P22 tailspike protein was denatured in 5 M urea at pH 3 for approximately 45 min. The refolding and aggregation reaction was initiated by rapid dilution (20-fold) with phosphate buffer containing 0.58 M urea, 1 mM EDTA, 1 mM β-mercaptoethanol, pH 7.6. The final aggregation conditions were 100 μg/mL protein at 20 °C and 0.8 M urea in sodium phosphate buffer. In following the aggregation kinetics, samples were taken at various times and rapidly transferred to preincubated tubes containing cold 3× sample buffer in an ice-water bath. Nondenaturing PAGE was performed using a discontinuous buffer system (Davis, 1964; Ornstein, 1964). The standard acrylamide concentration was 9–10% acrylamide to isolate tailspike multimers and proteolytic fragments, and the range of acrylamide concentration was 6.5–10% for the Ferguson plots. Gels were run at constant current (10 mA/gel) for 3–4 h at 4 °C and then silver stained (Sather & King, 1994).

Identification of aggregation intermediates

Figure 2 shows the pattern of protein bands as a function of time after initiation of the reaction. The starting native protein, prior to denaturation, migrated about 1/5 of the way down the gel as shown in the final lane. The lanes with samples renatured for various times after initiation showed multiple bands, representing multimers of tailspike chains. At the early time points (0.3–1.0 min), the distribution centered around species with higher electrophoretic mobility and presumably lower molecular weight than the native tailspike. The species with the highest mobility migrated approximately twice the distance as the native trimer. As the reaction proceeded, the multimer distribution shifted to species of decreasing mobility, consistent with increasing subunit composition. Polypeptide chains accumulated at the top of gel as the reaction proceeded, representing aggregates too large to enter the gel matrix.

The native tailspike that formed in the refolding reaction can be seen on the nondenaturing gels directly above the second band of multimers. Its time course of appearance was about 7 min with a final yield of less than 10%. The values for refolding kinetics reported in the literature (Danner & Seckler, 1993) were based on significantly different refolding conditions (30 μg/mL, 0.1 M urea, 25 °C, 20-h reconstitution). The native trimer is clearly a sharper band than that of the multimers. The fastest migrating band is presumably a conformer of the monomeric tailspike polypeptide chain. The relatively broad band may represent several species of folding or aggregation intermediates in different conformations. Comparison of the data to the native gels of Goldenberg and King (1982) suggests a possible candidate for the protrimer. The species migrating slightly slower than the trimer aggregate appears approximately on the same time scale as native trimer formation. This species is distinct but not predominant, and further analysis is necessary to confirm its identity.

Electrophoretic separation required 3–4 h in the cold. Because discrete species were observed, these species must have half-lives in the cold of hours rather than minutes. The isolation of a distinct ladder of multimer bands rather than a smear of protein indicated that aggregation reaction was retarded within the gel matrix during electrophoresis.

The earliest time point (0.3 min) already shows the presence of a distribution of multimers. One possibility was that the ag-

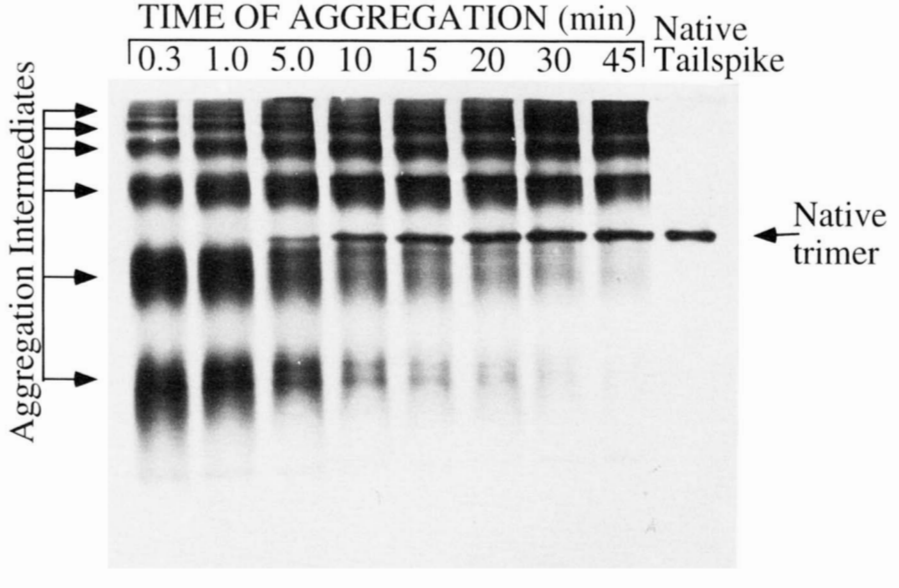

Fig. 2. In vitro aggregation kinetics of P22 tailspike protein. Denatured tailspike chains in 5 M urea, pH 3 were diluted 20-fold to 100 μg/mL protein at 20 °C and 0.8 M urea in 40 mM sodium phosphate buffer, pH 7.6. Samples were placed on ice and then electrophoresed through a 9% acrylamide native gel run for 4 h at 4 °C. The amount of sample loaded was 1.3 μg protein for the aggregation timepoints and 0.2 μg for the native tailspike control. The protein band were visualized by silver staining.

gregation reaction was not efficiently quenched by dilution and chilling. To examine this, an in vitro aggregation reaction was done at standard conditions of 100 μg/mL protein at 20 °C and 0.8 M urea in sodium phosphate buffer. The reaction mixture was placed on ice after 0.3 min, and samples from this mixture were electrophoresed at 1, 15, 30, and 60 min. Samples incubated on ice showed a distribution of multimers indistinguishable from that of the sample, which was electrophoresed immediately (data not shown). Although the aggregation reaction cannot be stopped completely, these controls suggest that the quenching conditions were adequate for experiments for which the samples incubate on ice for less than an hour.

In the productive refolding pathway, tryptophan fluorescence was quenched within milliseconds, even though the overall refolding reaction required much longer times (Fuchs et al., 1991). The presence of a distribution of multimers at early time points (0.3 min) suggested that these species formed from a very early kinetic intermediate. The multimers appeared to be stabilized within the gel matrix and were less susceptible to further aggregation in the gel than in solution.

In the initial development of this protocol to isolate aggregation intermediates, we were concerned that the polypeptide chains would aggregate further during the stacking step. The purpose of the upper gel is to concentrate the macromolecules into a narrow band, and this process could lead to aggregation, which could be occurring within the upper gel. In preliminary experiments, the upper gel was omitted to avoid any stacking effect. Comparison of the distribution of multimers obtained in gels with and without the upper gel indicated that the stacking effect in the upper gel as the dye became concentrated into a narrow band did not cause further aggregation. Additional controls indicated that the amount of protein loaded onto the gel and the extent of stacking did not affect the distribution of multimers.

Chain composition of the aggregation intermediates

Electrophoretic mobility in a nondenaturing gel depends on both size and charge (Chrambach & Rodbard, 1971). Unlike SDS gel electrophoresis, one cannot directly determine molecular weights from mobilities. In order to determine the number of chains in the various species, samples were electrophoresed through a set of nondenaturing gels at different percentages of acrylamide (Ferguson, 1964). The mobility of each multimeric species was then analyzed by the Ferguson method of plotting relative mobility versus the percent acrylamide.

The electrophoretic mobility of each multimer was measured in nondenaturing gels ranging from 6.5 to 10% acrylamide (Fig. 3A). To determine the molecular weight and size of multimers on a nondenaturing gel, the slope of the Ferguson plot $|K_r|$ (the retardation coefficient) was calculated. Assigning assumed molecular weights to the series of multimers and plotting this versus $|K_r|$ produced a linear correlation with $R = 0.994$ (Fig. 3B). Other nonsequential distributions (e.g., monomer, dimer, tetramer, octamer, etc.) did not produce a linear correlation with the slopes of the Ferguson plots. The results were consistent with the model that the ladder of distinct species separated on the nondenaturing gels was a sequential set of multimers of tailspike polypeptide chains.

The tailspike polypeptide chains have no disulfide bonds but contain eight cysteine residues. To avoid the formation of disulfide bonds, 1 mM dithiothreitol (DTT) was added to the standard nondenaturing sample buffer. To confirm that the species isolated on the nondenaturing gels were not a ladder of multimers having disulfide bonds, a lane was cut out of the nondenaturing gel and run in a second dimension on an SDS gel under nonreducing conditions. A sample of IgG (biotinylated polyclonal antibodies against the P22 coat protein) was used as a control to confirm that a disulfide-bonded species was reduced in an SDS gel in the presence of β-mercaptoethanol (βME) but remained intact when no βME was present. The multimers dissociated in the presence of SDS and comigrated with the monomeric tailspike control in the absence of βME (Fig. 4). This result proved that no disulfide bonds were present in the tailspike multimers, and the forces holding together the multimers could be disrupted by a strong detergent.

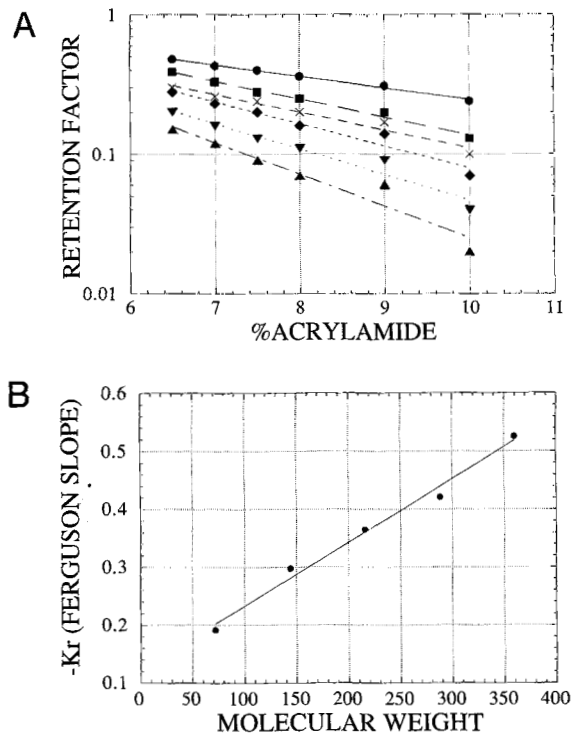

Fig. 3. A: Ferguson plot of P22 tailspike multimers. ●, monomer; ■, dimer; ×, native trimer; ◆, trimer aggregate; ▼, tetramer; ▲, pentamer. **B:** Molecular weight analysis of data from Ferguson plot. Linear regression determined to be the following: $-K_r = 0.122 + 0.00110(MW)$, $R = 0.994$.

Kinetics of the aggregation reaction

To assess more accurately the kinetic mechanism of the aggregation reaction, quantitative densitometry was used to determine the concentration profiles of tailspike multimers during the course of the aggregation reaction. Densitometry of the bands of multimers was quantified by recording the intensity of transmittance for the distribution of species and integrating the peaks. Because silver staining intensity is often not linearly related to the amount of protein, a set of dilutions of the multimers over a 5× range was used to confirm that the intensity of the silver staining was approximately linear within a limited range. To de-

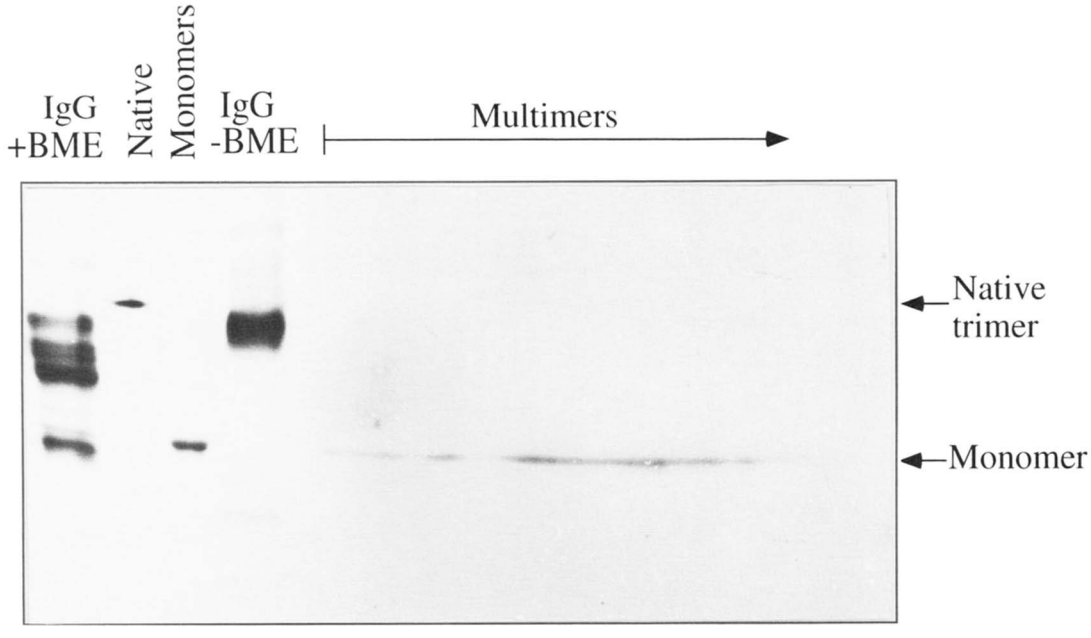

Fig. 4. Two-dimensional gel electrophoresis. Tailspike multimers isolated by nondenaturing gel electrophoresis and run in a second dimension on an SDS gel under nonreducing conditions. IgG + βME, polyclonal IgG against P22 coat protein + β-mercaptoethanol (βME). Native, native tailspike protein. Monomers, thermally denatured monomeric tailspike. IgG − βME, polyclonal IgG against P22 coat protein (no βME).

termine the relative amounts of protein in each band, the intensity of the silver stain was assumed to be proportional to the amount of polypeptide chain present. Therefore, a dimer was assumed to stain twice as much as a monomer, and the silver stain intensity of each *n*-mer was approximately *n* times that of the monomer. The resulting kinetic profile (Fig. 5) shows the relative amounts of monomers and dimers rapidly decreasing as the higher order aggregates increase as aggregation proceeded. This method of quantitative analysis may be used to determine the kinetic rate constants of each multimerization step along the aggregation pathway.

Extension of the methodology to other proteins

To determine if intermediates in other aggregation reactions could be trapped and fractionated by native gel electrophoresis, we examined the aggregation of CAB (Cleland & Wang, 1990) and the P22 coat protein (Teschke & King, 1993). Both proteins are monomeric and can be efficiently refolded in vitro from the fully denatured state. Off-pathway aggregation reactions compete with productive refolding by the association of a folding intermediate for both proteins. Native protein was denatured and refolding initiated by dilution following the protocol described in the Materials and methods. Multimers of CAB could be isolated by running a 10% acrylamide gel (Fig. 6A). Distinct multimers of the P22 coat protein were clearly separated by gel electrophoresis also (Fig. 6B). Both of these proteins have a native state that is monomeric, and the gels showed the presence of a strong monomeric band of native protein as well as a series of distinct multimers. The technique of nondenaturing PAGE to isolate early multimeric species during the initial stages of aggregation is clearly applicable to other protein systems.

The conformation of the aggregation intermediates

The native tailspike trimer is resistant to proteolysis, but both the monomer and protrimer intermediates are susceptible to proteolysis. If the multimeric species represent associated intermediates, they should be susceptible to proteolysis. To characterize the conformation of the multimers, a proteolysis experiment was performed on a solution of aggregating tailspike polypeptide chains. An in vitro refolding and aggregation study was done at the standard conditions of 100 μg/mL, 20 °C, and 0.8 M urea in 40 mM sodium phosphate buffer (pH 7.6). Samples were taken at 0.5 min and 30 min and treated with trypsin (4 μg/mL) for incubation periods of 0, 0.6, 1.3, 2.1, 5, and 15 min. After

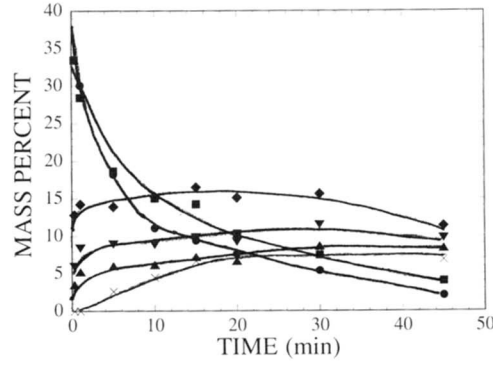

Fig. 5. Quantitative densitometry showing the time course of refolding and aggregation of P22 tailspike protein. ●, monomer; ■, dimer; ×, native trimer; ◆, trimer aggregate; ▼, tetramer; ▲, pentamer. Aggregation conditions were 100 μg/mL protein at 20 °C and 0.8 M urea in 40 mM sodium phosphate buffer, pH 7.6. Lines were hand drawn.

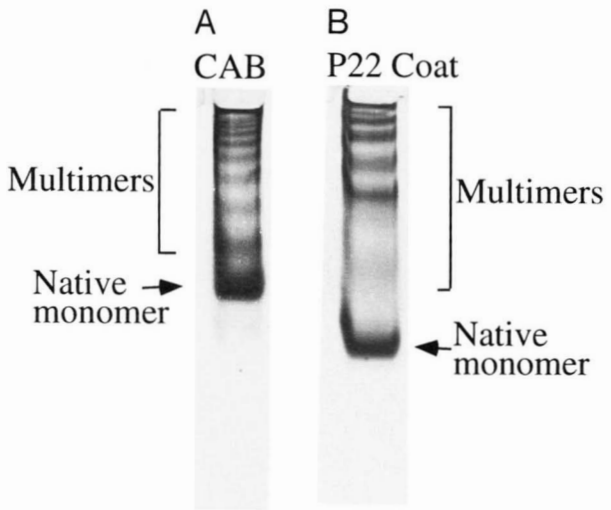

Fig. 6. Nondenaturing gel electrophoresis used to isolate multimers during in vitro aggregation. **A:** Carbonic anhydrase II (CAB) multimers formed by refolding at 1.0 mg/mL, 20 °C, 0.4 M urea in 40 mM sodium phosphate buffer, pH 7.6, for 0.3 min. Denaturation conditions were 5 M urea, pH 3, for ≥45 min at room temperature. **B:** P22 coat protein multimers aggregating at 100 μg/mL, 30 °C, 0.4 M urea in 40 mM sodium phosphate buffer, pH 7.6, for 0.3 min. Denaturation conditions were 6.75 M urea, pH 7.6, for ≥1 h at room temperature.

partial proteolysis for the given time interval, the trypsin was deactivated by combining the sample with bovine pancreatic trypsin inhibitor (200 μg/mL) for 0.5–1.0 min. Samples were then transferred to tubes on ice with aliquots of 3× sample buffer for nondenaturing and SDS gel electrophoresis. Nondenaturing gel electrophoresis was used to isolate tailspike multimers and the proteolytic fragments; SDS gel electrophoresis was used to determine the molecular weight distribution of the proteolytic fragments.

Although the multimers dissociated to form monomers in the presence of SDS, the SDS gel showed the pattern of proteolytic fragments clearer than the nondenaturing gel. Nondenaturing and SDS gel electrophoresis provided complementary data for the physical characterization of the multimers. Treatment of the reaction mixture with trypsin resulted in the appearance of shorter chain fragments under conditions in which native tailspike was not cleaved. All of the multimers were susceptible to trypsin digestion (Fig. 7). At high concentrations of trypsin, native tailspike that had the N-terminal end slightly unfolded (Chen & King, 1991) could be proteolytically clipped but not fully digested. The susceptibility of the multimers to protease digestion confirmed that these species had a non-native conformation. Note that large aggregates that were formed 30 min after the initial dilution were proteolyzed slower than the small multimeric species present early in the aggregation reaction.

Discussion

Despite the widespread occurrence of aggregation reactions competing effectively with the folding of pharmaceutically important proteins, there has been very limited analysis of the mechanisms of these association reactions. Part of the difficulty reflects the limited methodologies available to characterize the aggregated state; neither X-ray crystallography nor 2D-NMR is capable of providing structural information on the aggregation intermediates. The still soluble but larger aggregates presumed to be intermediates in such reactions have also been difficult to trap and characterize. However, the initial species in multimerization are likely to be critical in determining the yield in such pathways and in controlling the balance between productive folding and aggregation.

Nondenaturing PAGE effectively resolved multimeric species up to hexamers or heptamers of the tailspike partially folded chains. The results reported above identify discrete metastable polypeptide chain species that have the characteristics of early intermediates in the aggregation reaction. These species were not conformers of the native trimeric tailspike nor were they oligomers of the native state. The protease sensitivity indicates that these species were related to the folding intermediates along the productive pathway, which are also trypsin sensitive (Haase-Pettingell & King, 1988).

Characterization of aggregation intermediates

Under reducing conditions, in vitro tailspike aggregation occurs without the formation of intermolecular disulfide bonds. Dissociation of the multimers into monomers in the presence of SDS without reducing agents is consistent with hydrophobic forces as the underlying driving force of aggregation, but this does not exclude the importance of electrostatic interactions.

Proteolysis confirmed that the multimers on the aggregation pathway are non-native. Unlike small aggregates, the fully folded native tailspike protein is resistant to protease, SDS, and thermal denaturation (T_m = 88 °C). The species produced by partial proteolysis of the aggregates are a distinct distribution of small fragments. It is possible that the multimers have a defined structure and distinct exposed sites that are susceptible to proteolysis in a nonrandom manner. Aggregates may be less susceptible to proteolysis at later times in the aggregation reaction because large aggregates have a lower fraction of surface area exposed to the solvent. Protease molecules are prevented from reaching the inner core of the aggregate, and a significant fraction of the proteolytic cleavage sites may be buried.

Although aggregates form from folding intermediates and have non-native properties, there is evidence for various proteins that the folding intermediate involved in the association reaction has significant native-like secondary structure (Oberg et al., 1994). Raman spectroscopy data of β-lactamase inclusion bodies (Przybycien et al., 1994) and in vitro aggregates of P22 tailspike polypeptide chains (A. Mitraki & G.J. Thomas, unpubl. results) show the presence of a significant amount of β-sheet structure. Therefore, it is reasonable to hypothesize that the monomeric tailspike folding intermediate may have secondary structure similar to the subunit within the native trimer. If the folding intermediate has a β-spiral motif, the aggregation process may be analogous to the folding reaction. In tailspike folding, each monomeric subunit forms a β-spiral structure with significant hydrophobic interactions and intrasubunit β-sheet contacts between rungs of the helix. The complex is stabilized by hydrophilic contacts between the three β-spirals. A late folding step involves intersubunit winding of the C-terminal ends. By analogy, possible models for chain aggregation include the following: (1) inappropriate intersubunit β-sheet formation;

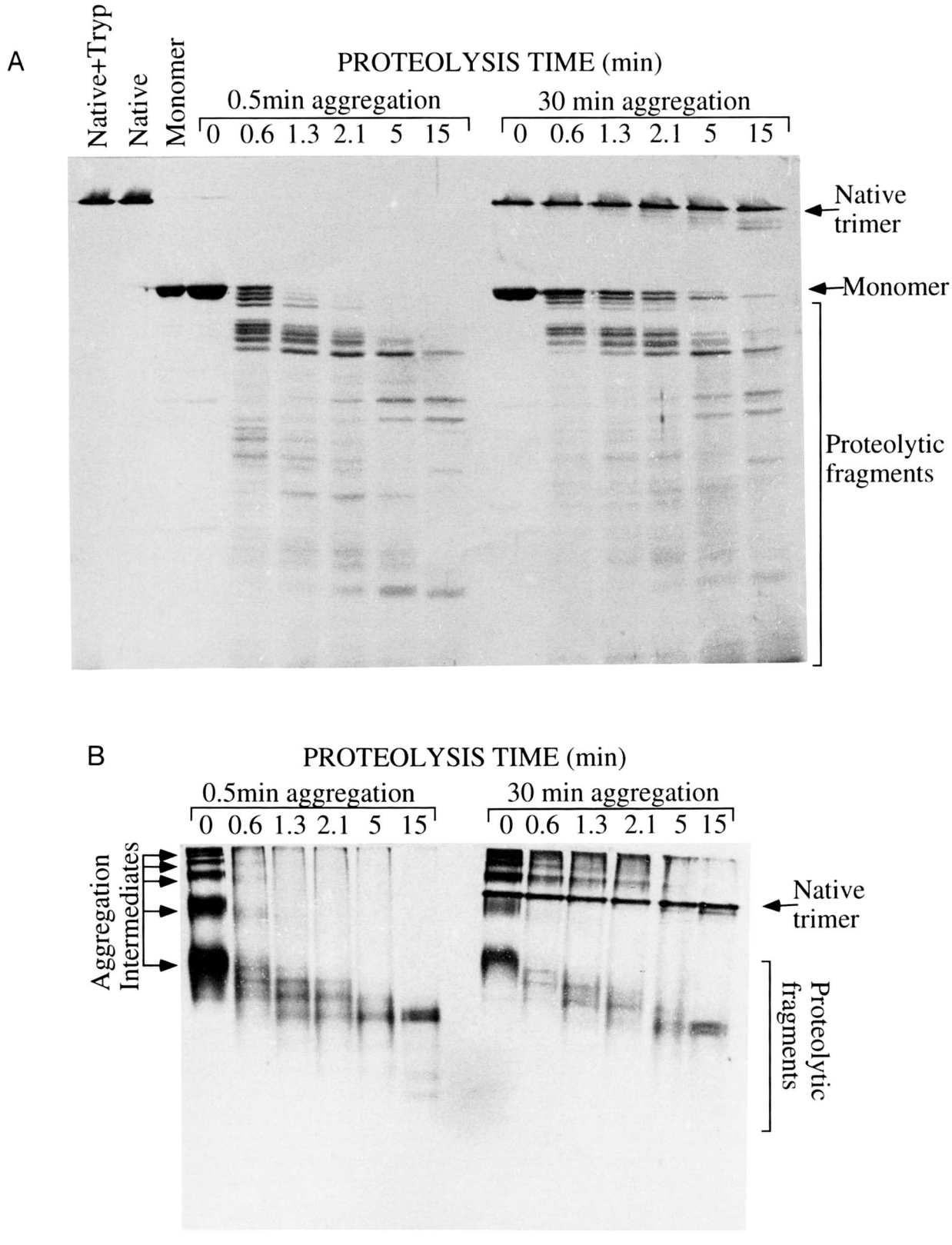

Fig. 7. Trypsin digestion of tailspike aggregates. During in vitro refolding and aggregation (100 μg/mL, 20 °C, and 0.8 M urea in 40 mM sodium phosphate buffer, pH 7.6), samples were taken at 0.5 min and 30 min and treated with trypsin (4 μg/mL) for 0, 0.6, 1.3, 2.1, 5, and 15 min. After partial proteolysis, the trypsin was deactivated by combining each sample with bovine pancreatic trypsin inhibitor (200 μg/mL) for 0.5–1.0 min. Tailspike polypeptide chains and proteolytic fragments were isolated by (**A**) SDS gel electrophoresis and (**B**) nondenaturing PAGE. To conserve space, certain control lanes of MW markers, bovine pancreatic trypsin inhibitor, trypsin, and native controls are not shown.

(2) misalignment of hydrophilic β-spiral contacts; and (3) incorrect intersubunit contacts at the C-terminus. Although all of these models are consistent with the Raman spectroscopy data, evidence that the folding intermediate responsible for aggregation is monomeric rules out the third model, which involves association of the protrimer species.

Determination of the aggregation pathway

In the tailspike aggregation pathway, the critical species is the monomeric folding intermediate (I) at the junction between productive folding and aggregation. It is unknown whether the first species along the aggregation pathway is monomeric (I*, in the model of Haase-Pettingell & King [1988]) or dimeric, similar to CAB (Cleland & Wang, 1991). The monomeric species isolated by nondenaturing gel electrophoresis may be I, I*, or a mixture of both species. Close inspection of the gels reveals that the broad monomer band contains several distinct monomeric species, which may correspond to I and I*. Although the gel electrophoresis data provide no evidence for the reversibility of the aggregation reaction, we cannot rule out the possibility of an equilibrium existing between the first two aggregation intermediates or even with higher order multimeric species.

The presence of the dimer aggregate under nonproductive refolding conditions provides evidence that aggregation involves the association of a monomeric folding intermediate to form discrete populations of multimers. During productive folding, no stable dimer folding intermediate is detectable in the folding pathway between the monomeric intermediate and the protrimer (Goldenberg & King, 1982). Therefore, the dimer identified in these experiments is one of the early species along the aggregation pathway.

The mechanism of polymerization

Interaction of monomeric polypeptide chains is required for the formation of a dimer and subsequent addition of a monomer to generate a trimer. Within the sensitivity of silver staining, it appears that the aggregation reaction continues after the monomeric species have been depleted. This makes it unlikely that sequential addition of monomers continues to be the major association pathway. Aggregation can proceed without a pool of monomers presumably by the association of multimers of various sizes by the "cluster–cluster" mechanism (Hemker & Frank, 1990). The polymerization process is not restricted to a sequential addition of monomeric units to the growing aggregate, or "particle–cluster" aggregation. These data suggest that multimeric polymerization rather than sequential polymerization can occur at later stages of aggregation.

The question of whether aggregation is a reaction under kinetic or thermodynamic control has been difficult to resolve. Because aggregation occurs by the association of a folding intermediate, the extent of aggregation is not a function of the thermodynamic stability of the native state. The kinetic competition between folding and aggregation may be predicted and controlled by focusing on the critical folding intermediate and the relative rates of the folding and aggregation reactions at that junction. Although it is evident that the aggregation process is essentially irreversible once large aggregates have formed, the early steps of multimerization may be reversible. For example, CAB forms a dimeric species that is in thermodynamic equilibrium with the monomeric folding intermediate (Cleland & Wang, 1991). For the tailspike protein, it is possible that early multimers along the aggregation pathway may be in equilibrium with each other. Controlling the early steps of multimerization is essential in pursuit of an effective strategy to balance the relative rates of folding versus aggregation. Application of this methodology of nondenaturing gel electrophoresis to study aggregation intermediates is feasible for other proteins of potential pharmaceutical interest.

Materials and methods

Materials

P22 tailspike protein was produced by infecting *Salmonella typhimurium* with P22 phage as described (King & Yu, 1986; Fuchs et al., 1991). The tailspike was then purified to approximately 95% purity, and its purity was verified by silver-stained SDS gels. The protein was stored as an ammonium sulfate precipitate and then dialyzed against Tris buffer when required. All experiments with the P22 tailspike used protein from the same preparation. The P22 coat protein was produced and purified as described (Teschke & King, 1993). Carbonic anhydrase II (CAB) from bovine erythrocytes was purchased from Sigma (product C2522).

In vitro aggregation kinetics

Native P22 tailspike protein was denatured for approximately 45 min in 5 M urea at pH 3. The refolding and aggregation reaction was initiated by rapid dilution (20-fold) with 40 mM sodium phosphate buffer and 0.58 M urea (1 mM EDTA, 1 mM βME, pH 7.6). To ensure rapid mixing for the aggregation kinetics experiment, 342 μL of the dilution buffer was added to a tube containing 18 μL of denatured protein, and the mixture was pipetted vigorously several times. The aggregation conditions were 100 μg/mL protein at 20 °C and 0.8 M urea in 40 mM sodium phosphate buffer (pH 7.6). In following the aggregation kinetics, aliquots of sample were taken at various time points and rapidly transferred to preincubated tubes containing cold 3× sample buffer (0.015 M Tris, 0.12 M glycine, 3 mM DTT, 30% glycerol, bromophenol blue) in an ice-water bath.

Gel electrophoresis

Nondenaturing PAGE was performed using a discontinuous buffer system (Davis, 1964; Ornstein, 1964). The resolving gel contained 0.37 M Tris buffer, pH 8.9, with 3.8 mM TEMED, 3.0 mM ammonium persulfate, and a specified percent acrylamide. A stacking gel was used initially, but omitting the stacking gel gave the same distribution of species. The upper gel contained 0.07 M Tris buffer, pH 6.7, with 4.3% acrylamide, 7.5 mM TEMED, and 2.5 mM ammonium persulfate. The standard acrylamide concentration was 9–10% acrylamide to isolate multimers and proteolytic fragments, and the range of acrylamide concentration was 6.5–10% for the Ferguson plots (Ferguson, 1964). A reducing agent (1 mM DTT) was added to the sample buffer to ensure that disulfide bonds did not form. An aliquot of 20 μL of protein solution in sample buffer (1.3 μg protein) was loaded onto the gel for each sample. The gels were run

at constant current (10 mA/gel) for 3-4 h at 4 °C and then silver stained as described by Sather and King (1994).

Two-dimensional SDS gel electrophoresis was performed by cutting a lane of multimers isolated on a nondenaturing gel (9% acrylamide) and soaking the lane in SDS sample buffer without any reducing agent. The lane was then placed horizontally across the top of a 1.5-mm-thick SDS gel (7.5% acrylamide) and run in a second dimension. SDS gel electrophoresis was done in a discontinuous buffer system with 7.5% acrylamide gels.

Densitometry was performed using a Hoefer densitometry and recording the intensity of transmittance. Integration of the densitometry peaks for each multimer was done by fitting the peaks using Hoefer software.

Proteolysis

Proteolysis experiments were performed using trypsin. Tailspike polypeptide chains and proteolytic fragments were isolated using both nondenaturing and SDS gel electrophoresis (10% acrylamide gel). A sample of refolding and aggregating tailspike was taken 0.5 min and 30 min after dilution (100 μg/mL, 0.8 M urea, 20 °C) and was proteolyzed using trypsin (4 μg/mL) for given incubation times (0, 0.6, 1.3, 2.1, 5, 15 min). Samples were taken at these designated times and combined with bovine pancreatic trypsin inhibitor (200 μg/mL) for 0.5-1.0 min to quench the proteolysis reaction. Aliquots of the proteolytic fragments were then combined with 3× sample buffer on ice for both SDS and nondenaturing gel electrophoresis. The nondenaturing gel was run at constant current (10 mA) and the SDS gel at constant voltage (150 V) until the dye front reached the end of the gel.

Acknowledgments

We acknowledge financial support from the National Science Foundation's Engineering Research Center Initiative under the Cooperative Agreement CDR-8803014. M.A.S. received support from the NIH Biotechnology Training Grant (NIGMS-2-T32-GM08334). The tailspike protein used in these studies was purified by A. Mitraki and C. Haase-Pettingell.

References

Brems DN. 1988. Solubility of different folding conformers of bovine growth hormone. *Biochemistry 27*:4541-4546.

Brunschier R, Danner M, Seckler R. 1993. Interactions of phage P22 tailspike protein with GroE molecular chaperones during refolding in vitro. *J Biol Chem 268*:2767-2772.

Chen BL, King J. 1991. Thermal unfolding pathway for the thermostable P22 tailspike endorhamnosidase. *Biochemistry 30*:6260-6269.

Chrambach A, Rodbard D. 1971. Polyacrylamide gel electrophoresis. *Science 172*:440-451.

Cleland JL, Wang DIC. 1990. Refolding and aggregation of bovine carbonic anhydrase B: Quasi-elastic light scattering analysis. *Biochemistry 29*:11072-11078.

Cleland JL, Wang DIC. 1991. Equilibrium association of a molten globule intermediate in the refolding of bovine carbonic anhydrase. In: Georgiou G, DeBernardez-Clark E, eds. *Protein refolding*. Washington, D.C.: American Chemical Society Symposium Series No. 470. pp 169-179.

Colon W, Kelly JW. 1992. Partial denaturation of transthyretin is sufficient for amyloid fibril formation in vitro. *Biochemistry 31*:8654-8660.

Danner M, Seckler R. 1993. Mechanism of phage P22 tailspike protein folding mutations. *Protein Sci 2*:1869-1881.

Davis BJ. 1964. Disk electrophoresis—II. Methods and application to human serum proteins. *Ann NY Acad Sci 121*:404-427.

DeBernardez-Clark E, Georgiou G. 1991. Inclusion bodies and recovery of proteins from the aggregated state. In: Georgiou G, DeBernardez-Clark E, eds. *Protein refolding*. Washington, D.C.: American Chemical Society Symposium Series No. 470. pp 1-20.

Ferguson KA. 1964. Starch-gel electrophoresis—Application to the classification of pituitary proteins and polypeptides. *Metabolism 13*:985-1002.

Fuchs A, Seiderer C, Seckler R. 1991. In vitro folding pathway of phage P22 tailspike protein. *Biochemistry 30*:6598-6604.

Goldenberg DP, Creighton TE. 1994. A fishy tale of protein folding. *Curr Biol 4*:1026-1029.

Goldenberg D, King J. 1981. Temperature-sensitive mutants blocked in the folding of subunit assembly of the bacteriophage P22 tail spike protein. II. Active mutant proteins matured at 30 °C. *J Mol Biol 145*:633-651.

Goldenberg D, King J. 1982. Trimeric intermediate in the in vivo folding and subunit assembly of the tail spike endorhamnosidase of bacteriophage P22. *Proc Natl Acad Sci USA 79*:3403-3407.

Goldenberg D, Smith DH, King J. 1983. Genetic analysis of the folding pathway for the tail spike protein of phage P22. *Proc Natl Acad Sci USA 80*:7060-7064.

Goloubinoff P, Christeller JT, Gatenby AA, Lorimer GH. 1989. Reconstitution of active dimeric ribulose bisphosphate carboxylase from an unfolded state depends on two chaperonin proteins and Mg-ATP. *Nature 342*:884-889.

Haase-Pettingell C, King J. 1988. Formation of aggregates from a thermolabile in vivo folding intermediate in P22 tailspike maturation. *J Biol Chem 263*:4977-4983.

Hemker DJ, Frank CW. 1990. Dynamic light-scattering studies of the fractal aggregation of poly(methacrylic acid) and poly(ethylene glycol). *Macromolecules 23*:4404-4410.

Hendrick JP, Hartl FU. 1993. Molecular chaperone functions of heat-shock proteins. *Annu Rev Biochem 62*:349-384.

Jaenicke R, Rudolph R. 1986. Refolding and association of oligomeric proteins. *Methods Enzymol 131*:218-250.

Kim PS, Baldwin RL. 1990. Intermediates in the folding reactions of small proteins. *Annu Rev Biochem 59*:631-660.

King J, Yu MH. 1986. Mutational analysis of protein folding pathways: The P22 tailspike endorhamnosidase. *Methods Enzymol 131*:250-266.

Marston FAO. 1986. The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. *Biochem J 240*:1-12.

Mitraki A, Betton JM, Desmadril M, Yon JM. 1987. Quasi-irreversibility in the unfolding-refolding transition of phosphoglycerate kinase induced by guanidine hydrochloride. *Eur J Biochem 163*:29-34.

Mitraki A, Danner M, King J, Seckler R. 1993. Temperature-sensitive mutations and second-site suppressor substitutions affect folding of the P22 tailspike protein in vitro. *J Biol Chem 268*:20071-20075.

Mitraki A, Fane B, Haase-Pettingell C, Sturtevant J, King J. 1991. Global suppression of protein folding defects and inclusion body formation. *Science 253*:54-58.

Mitraki A, King J. 1989. Protein folding intermediates and inclusion body formation. *Bio/Technology 7*:690-697.

Oberg K, Chrunyk BA, Wetzel R, Fink AL. 1994. Nativelike secondary structure in interleukin-1β inclusion bodies by attenuated total reflectance FTIR. *Biochemistry 33*:2628-2634.

Ornstein L. 1964. Disc electrophoresis—I. Background and theory. *Ann NY Acad Sci 121*:321-349.

Przybycien TM, Dunn JP, Valax P, Georgiou G. 1994. Secondary structure characterization of β-lactamase inclusion bodies. *Protein Eng 7*:131-136.

Sather S, King J. 1994. Intracellular trapping of a cytoplasmic folding intermediate of the phage P22 tailspike using iodoacetamide. *J Biol Chem 269*:25268-25276.

Seckler R, Fuchs A, King J, Jaenicke R. 1989. Reconstitution of the thermostable trimeric phage P22 tailspike protein from denatured chains in vitro. *J Biol Chem 264*:11750-11753.

Steinbacher S, Seckler R, Miller S, Steipe B, Huber R, Reinemer P. 1994. Crystal structure of P22 tailspike protein: Interdigitated subunits in a thermostable trimer. *Science 265*:383-386.

Teschke CM, King J. 1993. Folding of the phage P22 coat protein in vitro. *Biochemistry 32*:10839-10847.

Wetzel R. 1994. Mutations and off-pathway aggregation of proteins. *Trends Biotechnol 12*:193-198.

Yoder MD, Keen NT, Jurnak F. 1994. New domain motif: The structure of pectate lyase C, a secreted plant virulence factor. *Science 260*:1503-1507.

Zettlmeissl G, Rudolph R, Jaenicke R. 1979. Reconstitution of lactic dehydrogenase. Noncovalent aggregation vs. reactivation. 1. Physical properties and kinetics of aggregation. *Biochemistry 18*:5567-5571.